



Técnicas de Agrupamento de Dados Multi-Grupo e Multi-Padrão

HUGO FERNANDO FREITAS PEREIRA

Outubro de 2019

Técnicas de Agrupamento de Dados Multi-Grupo e Multi-Padrão

Hugo Fernando Freitas Pereira

**Dissertação para obtenção do Grau de Mestre em
Engenharia Informática, Área de Especialização em
Sistemas Computacionais**

Orientador: Fernando Jorge Ferreira Duarte

Co-orientador: Ana Luísa Nobre Fred, Instituto Superior Técnico

Júri:

Presidente:

Vogais:

Resumo

Em qualquer processo de aprendizagem automática, é importante saber quais os algoritmos que produzem melhores resultados na análise de um determinado conjunto de dados. Os algoritmos de agrupamento de dados são avaliados, principalmente, pela qualidade da exatidão (precisão) do agrupamento produzido quando comparado com a divisão real do conjunto de dados, pelo seu tempo de execução e pelos recursos utilizados (p.e., tempo de CPU). Com o conhecimento prévio de quais os algoritmos têm um melhor desempenho, o utilizador tem a oportunidade de usar esses algoritmos e obter uma melhor partição do conjunto de dados.

O objetivo desta dissertação consiste no estudo de técnicas de agrupamento de dados associadas aos conceitos multi-grupo (cada objeto pode pertencer a mais do que um grupo) e multi-padrão (cada objeto pode ser caracterizados por mais do que um atributo, atributo esse que pode ser responsável pela classificação do objeto).

Nesta dissertação são apresentados vários algoritmos de agrupamento multi-grupo e multi-padrão e várias métricas de avaliação da qualidade da performance desses mesmos algoritmos.

Esta dissertação apresenta dois estudos comparativos. Um entre algoritmos de agrupamento multi-grupo (LIFT, CLMLC, CBLML e HOMER) e outro entre algoritmos de agrupamento multi-padrão (MCIL e BARTMIP). Os estudos comparam e avaliam a performance dos vários algoritmos para diferentes conjuntos de dados, dando indicações aos utilizadores de quais os algoritmos que obtiveram melhores resultados na performance do agrupamento de dados obtido.

Palavras-chave: Agrupamento de dados, Aprendizagem Multi-Grupo, Aprendizagem Multi-Padrão, métricas de avaliação

Abstract

In any machine learning process, it is important to know which algorithms produce the best results when analysing a given dataset. Clustering algorithms are mainly evaluated by the accuracy and precision of the produced clusters when compared with the real divisions in the dataset, its execution time and the computing resources used (e.g., CPU time). With prior knowledge of which algorithms perform best, the user has the opportunity to use these algorithms and obtain a better partition of the dataset.

This thesis objective is to study data clustering techniques associated with the multi-label concept (each object can belong to more than one group) and the multi-instance concept (each object can be characterized by more than one attribute, this attribute may be solely responsible by the object's classification).

In this thesis, several multi-label clustering and multi-instance clustering algorithms are presents, as well as several performance quality evaluation metrics of these same algorithms.

This thesis presents two comparative studies. One a study between multi-label clustering algorithms (LIFT, CLMLC, CBLML and HOMER) and another between multi-instance algorithms (MCIL and BARTMIP). The studies compare and evaluate the performance of the various algorithms for different datasets, giving users an indication of which algorithms achieved the best results in these performance evaluations.

Keywords: Data clustering, Multi-Label Learning, MLC, Multi-Instance Learning, MIC, evaluation metrics

Índice

1	Introdução	21
2	Estado de Arte	23
2.1	Agrupamento de Dados	23
2.2	Aprendizagem Multi-Padrão	28
2.2.1	Aprendizagem supervisionada Multi-Padrão	28
2.2.2	Aprendizagem não-supervisionada Multi-Padrão.....	30
2.3	Aprendizagem Multi-Grupo	32
2.3.1	Aprendizagem supervisionada Multi-Grupo	32
2.3.2	Aprendizagem não supervisionada Multi-Grupo.....	34
2.4	Aprendizagem Multi-Grupo e Multi-Padrão simultânea	35
2.4.1	MIMLBOOST.....	37
2.4.2	MIMLSVM	37
2.4.3	En-MIMLSVM	37
2.5	Outros conceitos relacionados com processamento de dados	38
2.5.1	Cross Validation	38
2.5.2	k-fold.....	38
3	Algoritmos de Agrupamento Multi-Padrão.....	39
3.1	Agrupamento Multi-Padrão	39
3.1.1	BAMIC	39
3.1.2	BARTMIP	41
3.1.3	M ³ IC-MBM	43
3.1.4	MCIL	45
3.1.5	UC-kNN.....	46
3.1.6	COSMIC.....	47
3.2	Conjuntos de dados Multi-Padrão	47
3.2.1	Musk.....	47
3.2.2	Fox, Tiger e Elephant.....	48
4	Algoritmos de Agrupamento Multi-Grupo.....	49
4.1	Agrupamento Multi-Grupo	49
4.1.1	LIFT.....	49
4.1.2	CBMLC.....	51
4.1.3	CLMLC.....	51
4.1.4	HOMER	52
4.2	Conjuntos de dados Multi-Grupo	52
4.2.1	Bibtex	53
4.2.2	Birds.....	53
4.2.3	Corel5k.....	53
4.2.4	Emotions.....	53
4.2.5	Genbase	53

4.2.6	Medical	53
4.2.7	Yeast	53
5	Avaliação dos Algoritmos.....	54
5.1	Desenho da solução	54
5.2	Avaliação da solução	56
5.3	Métricas de Avaliação	57
5.3.1	Tempo de Execução (tExe)	57
5.3.2	Tempo de CPU (tCPU)	58
5.3.3	Matriz de confusão.....	59
5.3.4	Exatidão (Accuracy, A)	59
5.3.5	Precisão (Precision, P).....	59
5.3.6	Revocação (Recall, R)	60
5.3.7	Área abaixo da curva ROC (AUC).....	60
5.3.8	F-Measure (F)	60
5.3.9	Exact Match Ratio (MR).....	60
5.3.10	Hamming Loss (HL).....	60
5.3.11	F1-macro	61
5.3.12	F1-micro.....	61
5.4	Recursos de Computação	62
6	Análise dos Algoritmos Multi-Padrão.....	63
6.1	Testes de Normalidade	64
6.2	Teste de Hipóteses 1	65
6.3	Teste de Hipóteses 2	66
7	Análise dos Algoritmos Multi-Grupo.....	67
7.1	Testes de Normalidade.....	69
7.2	Teste de Hipóteses 1	70
7.3	Teste de Hipóteses 2	73
8	Conclusão	75
A	Anexo A - Processamento dos Algoritmos Multi-Padrão.....	83
A.1	Resultados e Testes de Normalidade	83
A.1.1	MCIL.....	83
A.1.2	BARTMIP	86
A.2	Teste de Hipóteses 1	89
A.3	Teste de Hipóteses 2	104
B	Anexo B - Processamento dos Algoritmos Multi-Grupo.....	107
B.1	Resultados e Testes de Normalidade	107
B.1.1	LIFT	107
B.1.2	CLMLC.....	111

B.1.3	CBMLC.....	114
B.1.4	HOMER	118
B.2	Teste de Hipóteses 1	122
B.2.1	Tempo de Execução.....	122
B.2.2	Tempo de CPU	136
B.2.3	Exact Match Ratio	150
B.2.4	Hamming Loss.....	164
B.2.5	F1-macro	178
B.2.6	F1-micro	192
B.3	Teste de Hipóteses 2	206

Lista de Figuras

Figura 1 – Exemplo de agrupamento de dados (Jain, 2010)	24
Figura 2 – Fases do agrupamento de dados	26
Figura 3 – Aprendizagem Supervisionada	28
Figura 4 – Aprendizagem Multi-Padrão	29
Figura 5 – Exemplo de instâncias de um objeto multi-padrão	30
Figura 6 – Comparação dos componentes da aprendizagem MIML	35
Figura 7 – Exemplo de aprendizagem MIML em classificação de imagens (Feng and Xu, 2010)	36

Lista de Tabelas

Tabela 1 – Pseudocódigo descrevendo o algoritmo BAMIC	40
Tabela 2 – Pseudocódigo descrevendo o algoritmo BARTMIP	41
Tabela 3 – Pseudocódigo do algoritmo M ³ IC-MBM.....	43
Tabela 4 – Pseudocódigo do algoritmo MICL.....	45
Tabela 5 – Descrição de conjuntos de dados multi-padrão usados.....	47
Tabela 6 – Pseudocódigo do algoritmo LIFT	50
Tabela 7 – Descrição de conjuntos de dados multi-grupo.....	52
Tabela 8 – Matriz de confusão	59
Tabela 9 – Resumo com médias dos resultados gerados pelos algoritmos multi-padrão.....	64
Tabela 10 – Resultados do teste de hipóteses 1.....	65
Tabela 11 – Resultados do teste de hipóteses 2.....	66
Tabela 12 – Resumo com médias dos resultados gerados pelo algoritmo LIFT	68
Tabela 13 – Resumo com médias dos resultados gerados pelo algoritmo CLMLC.....	68
Tabela 14 – Resumo com médias dos resultados gerados pelo algoritmo CBMLC	68
Tabela 15 – Resumo com médias dos resultados gerados pelo algoritmo HOMER	69
Tabela 16 – Resultados do teste de hipóteses 1 – Teste ANOVA	70
Tabela 17 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – Tempo de Execução.....	70
Tabela 18 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – Tempo de CPU	71
Tabela 19 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – Exact Match Ratio.....	71
Tabela 20 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – Hamming Loss.....	71
Tabela 21 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – F1-macro.....	72
Tabela 22 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – F1-micro.....	72
Tabela 23 – Resultados do teste de hipóteses 2 – teste de Friedman.....	73
Tabela 24 – Resultados do teste de hipóteses 2 – teste de Conover	73
Tabela 25 – Resultados e teste Shapiro-Wilk para o conjunto de dados Musk1.....	83
Tabela 26 – Resultados e teste Shapiro-Wilk para o conjunto de dados Musk2.....	84
Tabela 27 – Resultados e teste Shapiro-Wilk para o conjunto de dados Fox.....	84
Tabela 28 – Resultados e teste Shapiro-Wilk para o conjunto de dados Tiger.....	85
Tabela 29 – Resultados e teste Shapiro-Wilk para o conjunto de dados Elephant	85
Tabela 30 – Resultados e teste Shapiro-Wilk para o conjunto de dados Musk1.....	86
Tabela 31 – Resultados e teste Shapiro-Wilk para o conjunto de dados Musk2.....	86
Tabela 32 – Resultados e teste Shapiro-Wilk para o conjunto de dados Fox.....	87
Tabela 33 – Resultados e teste Shapiro-Wilk para o conjunto de dados Tiger.....	87
Tabela 34 – Resultados e teste Shapiro-Wilk para o conjunto de dados Elephant	88
Tabela 35 – Teste T de Student para o tempo de execução - conjunto de dados Musk1	89
Tabela 36 – Teste T de Student para o tempo de execução - conjunto de dados Musk2	90
Tabela 37 – Teste T de Student para o tempo de execução - conjunto de dados Fox	91
Tabela 38 – Teste T de Student para o tempo de execução - conjunto de dados Tiger.....	92
Tabela 39 – Teste T de Student para o tempo de execução - conjunto de dados Elephant.....	93
Tabela 40 – Teste T de Student para a exatidão - conjunto de dados Musk1	94

Tabela 41 – Teste T de Student para a exatidão - conjunto de dados Musk2	95
Tabela 42 – Teste T de Student para a exatidão - conjunto de dados Fox	96
Tabela 43 – Teste T de Student para a exatidão - conjunto de dados Tiger	97
Tabela 44 – Teste T de Student para a exatidão - conjunto de dados Elephant.....	98
Tabela 45 – Teste T de Student para AUC - conjunto de dados Musk1	99
Tabela 46 – Teste T de Student para AUC - conjunto de dados Musk2	100
Tabela 47 – Teste T de Student para AUC - conjunto de dados Fox	101
Tabela 48 – Teste T de Student para AUC - conjunto de dados Tiger.....	102
Tabela 49 – Teste T de Student para AUC - conjunto de dados Elephant.....	103
Tabela 50 – Teste de Wilcoxon para tempo de execução.....	104
Tabela 51 – Teste de Wilcoxon para a exatidão.....	105
Tabela 52 – Teste de Wilcoxon para AUC	106
Tabela 53 – Resultados e teste Shapiro-Wilk para o conjunto de dados Bibtex.....	107
Tabela 54 – Resultados e teste Shapiro-Wilk para o conjunto de dados Birds.....	108
Tabela 55 – Resultados e teste Shapiro-Wilk para o conjunto de dados Corel5k.....	108
Tabela 56 – Resultados e teste Shapiro-Wilk para o conjunto de dados Emotions.....	109
Tabela 57 – Resultados e teste Shapiro-Wilk para o conjunto de dados Genbase.....	109
Tabela 58 – Resultados e teste Shapiro-Wilk para o conjunto de dados Medical	110
Tabela 59 – Resultados e teste Shapiro-Wilk para o conjunto de dados Yeast	110
Tabela 60 – Resultados e teste Shapiro-Wilk para o conjunto de dados Bibtex.....	111
Tabela 61 – Resultados e teste Shapiro-Wilk para o conjunto de dados Birds.....	111
Tabela 62 – Resultados e teste Shapiro-Wilk para o conjunto de dados Corel5k.....	112
Tabela 63 – Resultados e teste Shapiro-Wilk para o conjunto de dados Emotions.....	112
Tabela 64 – Resultados e teste Shapiro-Wilk para o conjunto de dados Genbase.....	113
Tabela 65 – Resultados e teste Shapiro-Wilk para o conjunto de dados Medical	113
Tabela 66 – Resultados e teste Shapiro-Wilk para o conjunto de dados Yeast	114
Tabela 67 – Resultados e teste Shapiro-Wilk para o conjunto de dados Bibtex.....	114
Tabela 68 – Resultados e teste Shapiro-Wilk para o conjunto de dados Birds.....	115
Tabela 69 – Resultados e teste Shapiro-Wilk para o conjunto de dados Corel5k.....	115
Tabela 70 – Resultados e teste Shapiro-Wilk para o conjunto de dados Emotions.....	116
Tabela 71 – Resultados e teste Shapiro-Wilk para o conjunto de dados Genbase.....	116
Tabela 72 – Resultados e teste Shapiro-Wilk para o conjunto de dados Medical	117
Tabela 73 – Resultados e teste Shapiro-Wilk para o conjunto de dados Yeast	117
Tabela 74 – Resultados e teste Shapiro-Wilk para o conjunto de dados Bibtex.....	118
Tabela 75 – Resultados e teste Shapiro-Wilk para o conjunto de dados Birds.....	118
Tabela 76 – Resultados e teste Shapiro-Wilk para o conjunto de dados Corel5k.....	119
Tabela 77 – Resultados e teste Shapiro-Wilk para o conjunto de dados Emotions.....	119
Tabela 78 – Resultados e teste Shapiro-Wilk para o conjunto de dados Genbase.....	120
Tabela 79 – Resultados e teste Shapiro-Wilk para o conjunto de dados Medical	120
Tabela 80 – Resultados e teste Shapiro-Wilk para o conjunto de dados Yeast	121
Tabela 81 – Teste ANOVA para o tempo de execução - conjunto de dados Bibtex	122
Tabela 82 – Teste Tukey HSD para o tempo de execução - conjunto de dados Bibtex	123
Tabela 83 – Teste ANOVA para o tempo de execução - conjunto de dados Birds.....	124

Tabela 84 – Teste Tukey HSD para o tempo de execução - conjunto de dados Birds	125
Tabela 85 – Teste ANOVA para o tempo de execução - conjunto de dados Corel5k	126
Tabela 86 – Teste Tukey HSD para o tempo de execução - conjunto de dados Corel5k.....	127
Tabela 87 – Teste ANOVA para o tempo de execução - conjunto de dados Emotions	128
Tabela 88 – Teste Tukey HSD para o tempo de execução - conjunto de dados Emotions.....	129
Tabela 89 – Teste ANOVA para o tempo de execução - conjunto de dados Genbase	130
Tabela 90 – Teste Tukey HSD para o tempo de execução - conjunto de dados Genbase.....	131
Tabela 91 – Teste ANOVA para o tempo de execução - conjunto de dados Medical.....	132
Tabela 92 – Teste Tukey HSD para o tempo de execução - conjunto de dados Medical	133
Tabela 93 – Teste ANOVA para o tempo de execução - conjunto de dados Yeast.....	134
Tabela 94 – Teste Tukey HSD para o tempo de execução - conjunto de dados Yeast	135
Tabela 95 – Teste ANOVA para o tempo de CPU - conjunto de dados Bibtex.....	136
Tabela 96 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Bibtex	137
Tabela 97 – Teste ANOVA para o tempo de CPU - conjunto de dados Birds.....	138
Tabela 98 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Birds.....	139
Tabela 99 – Teste ANOVA para o tempo de CPU - conjunto de dados Corel5k	140
Tabela 100 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Corel5k	141
Tabela 101 – Teste ANOVA para o tempo de CPU - conjunto de dados Emotions.....	142
Tabela 102 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Emotions	143
Tabela 103 – Teste ANOVA para o tempo de CPU - conjunto de dados Genbase.....	144
Tabela 104 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Genbase	145
Tabela 105 – Teste ANOVA para o tempo de CPU - conjunto de dados Medical	146
Tabela 106 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Medical.....	147
Tabela 107 – Teste ANOVA para o tempo de CPU - conjunto de dados Yeast	148
Tabela 108 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Yeast.....	149
Tabela 109 – Teste ANOVA para o exact match ratio - conjunto de dados Bibtex.....	150
Tabela 110 – Teste Tukey HSD para o exact match ratio - conjunto de dados Bibtex	151
Tabela 111 – Teste ANOVA para o exact match ratio - conjunto de dados Birds.....	152
Tabela 112 – Teste Tukey HSD para o exact match ratio - conjunto de dados Birds	153
Tabela 113 – Teste ANOVA para o exact match ratio - conjunto de dados Corel5k.....	154
Tabela 114 – Teste Tukey HSD para o exact match ratio - conjunto de dados Corel5k	155
Tabela 115 – Teste ANOVA para o exact match ratio - conjunto de dados Emotions.....	156
Tabela 116 – Teste Tukey HSD para o exact match ratio - conjunto de dados Emotions	157
Tabela 117 – Teste ANOVA para o exact match ratio - conjunto de dados Genbase.....	158
Tabela 118 – Teste Tukey HSD para o exact match ratio - conjunto de dados Genbase	159
Tabela 119 – Teste ANOVA para o exact match ratio - conjunto de dados Medical.....	160
Tabela 120 – Teste Tukey HSD para o exact match ratio - conjunto de dados Medical.....	161
Tabela 121 – Teste ANOVA para o exact match ratio - conjunto de dados Yeast	162
Tabela 122 – Teste Tukey HSD para o exact match ratio - conjunto de dados Yeast.....	163
Tabela 123 – Teste ANOVA para o hamming loss - conjunto de dados Bibtex.....	164
Tabela 124 – Teste Tukey HSD para o hamming loss - conjunto de dados Bibtex	165
Tabela 125 – Teste ANOVA para o hamming loss - conjunto de dados Birds.....	166
Tabela 126 – Teste Tukey HSD para o hamming loss - conjunto de dados Birds.....	167

Tabela 127 – Teste ANOVA para o hamming loss - conjunto de dados Corel5k.....	168
Tabela 128 – Teste Tukey HSD para o hamming loss - conjunto de dados Corel5k.....	169
Tabela 129 – Teste ANOVA para o hamming loss - conjunto de dados Emotions.....	170
Tabela 130 – Teste Tukey HSD para o hamming loss - conjunto de dados Emotions.....	171
Tabela 131 – Teste ANOVA para o hamming loss - conjunto de dados Genbase.....	172
Tabela 132 – Teste Tukey HSD para o hamming loss - conjunto de dados Genbase.....	173
Tabela 133 – Teste ANOVA para o hamming loss - conjunto de dados Medical.....	174
Tabela 134 – Teste Tukey HSD para o hamming loss - conjunto de dados Medical.....	175
Tabela 135 – Teste ANOVA para o hamming loss - conjunto de dados Yeast.....	176
Tabela 136 – Teste Tukey HSD para o hamming loss - conjunto de dados Yeast.....	177
Tabela 137 – Teste ANOVA para o f1-macro - conjunto de dados Bibtex.....	178
Tabela 138 – Teste Tukey HSD para o f1-macro - conjunto de dados Bibtex.....	179
Tabela 139 – Teste ANOVA para o f1-macro - conjunto de dados Birds.....	180
Tabela 140 – Teste Tukey HSD para o f1-macro - conjunto de dados Birds.....	181
Tabela 141 – Teste ANOVA para o f1-macro - conjunto de dados Corel5k.....	182
Tabela 142 – Teste Tukey HSD para o f1-macro - conjunto de dados Corel5k.....	183
Tabela 143 – Teste ANOVA para o f1-macro - conjunto de dados Emotions.....	184
Tabela 144 – Teste Tukey HSD para o f1-macro - conjunto de dados Emotions.....	185
Tabela 145 – Teste ANOVA para o f1-macro - conjunto de dados Genbase.....	186
Tabela 146 – Teste Tukey HSD para o f1-macro - conjunto de dados Genbase.....	187
Tabela 147 – Teste ANOVA para o f1-macro - conjunto de dados Medical.....	188
Tabela 148 – Teste Tukey HSD para o f1-macro - conjunto de dados Medical.....	189
Tabela 149 – Teste ANOVA para o f1-macro - conjunto de dados Yeast.....	190
Tabela 150 – Teste Tukey HSD para o f1-macro - conjunto de dados Yeast.....	191
Tabela 151 – Teste ANOVA para o f1-micro - conjunto de dados Bibtex.....	192
Tabela 152 – Teste Tukey HSD para o f1-macro - conjunto de dados Bibtex.....	193
Tabela 153 – Teste ANOVA para o f1-micro - conjunto de dados Birds.....	194
Tabela 154 – Teste Tukey HSD para o f1-macro - conjunto de dados Birds.....	195
Tabela 155 – Teste ANOVA para o f1-micro - conjunto de dados Corel5k.....	196
Tabela 156 – Teste Tukey HSD para o f1-macro - conjunto de dados Corel5k.....	197
Tabela 157 – Teste ANOVA para o f1-micro - conjunto de dados Emotions.....	198
Tabela 158 – Teste Tukey HSD para o f1-macro - conjunto de dados Emotions.....	199
Tabela 159 – Teste ANOVA para o f1-micro - conjunto de dados Genbase.....	200
Tabela 160 – Teste Tukey HSD para o f1-macro - conjunto de dados Genbase.....	201
Tabela 161 – Teste ANOVA para o f1-micro - conjunto de dados Medical.....	202
Tabela 162 – Teste Tukey HSD para o f1-macro - conjunto de dados Medical.....	203
Tabela 163 – Teste ANOVA para o f1-micro - conjunto de dados Yeast.....	204
Tabela 164 – Teste Tukey HSD para o f1-macro - conjunto de dados Yeast.....	205
Tabela 165 – Teste de Friedman para tempo de execução.....	206
Tabela 166 – Teste de Conover para tempo de execução.....	207
Tabela 167 – Teste de Friedman para tempo de CPU.....	208
Tabela 168 – Teste de Conover para tempo de CPU.....	209
Tabela 169 – Teste de Friedman para exact match ratio.....	210

Tabela 170 – Teste de Conover para exact match ratio	211
Tabela 171 – Teste de Friedman para hamming loss.....	212
Tabela 172 – Teste de Friedman para f1-macro	213
Tabela 173 – Teste de Friedman para f1-micro	214

Acrónimos e Símbolos

Lista de Acrónimos

ARN	Ácido Ribonucleico
AUC	<i>Area Under the ROC Curve</i>
BAMIC	<i>BAG-level Multi-Instance Clustering</i>
BARTMIP	<i>BAG-level Representation Transformation for Multi-Instance Prediction</i>
CBMLC	<i>Clustering-Based Multi-Label Classification</i>
CLMLC	<i>Clustering-based Local Multi-Label Classification</i>
COSMIC	<i>COnceptual Specified Multi-Instance Clusters</i>
CPU	<i>Central Processing Unit</i>
EM	<i>Exact Match Ratio</i>
GPU	<i>Graphic Processing Unit</i>
HL	<i>Hamming Loss</i>
HOMER	<i>Hierarchy Of Multilabel classifiERs</i>
HTML	<i>HyperText Markup Language</i>
IC	Índice de Consistência
LIFT	<i>multi-label learning with Label specific FeaTures</i>
MCIL	<i>Multiple Clustered Instance Learning</i>
MIC	<i>Multiple Instance Clustering</i>
MIL	<i>Multiple Instance Learning</i>
MIML	<i>Multiple Instance Multiple Label</i>
M3IC-MBM	<i>Maximum Margin Multiple Instance Clustering - Modified Bag Margin</i>
RAM	<i>Random Access Memory</i>
RC	Razão de Consistência
ROC	<i>Receiver Operator Characteristic Curve</i>

SSD	<i>Solid-State Drive</i>
SVM	Support Vector Machine
UC-kNN	<i>Unsupervised Citation k-Nearest Neighbor</i>

1 Introdução

As técnicas de agrupamento de dados são úteis em vários problemas de análise exploratória de dados, tomada de decisão, aprendizagem automática, estruturação de documentos e segmentação de imagem, entre outros. Têm como objetivo a divisão de um conjunto de dados em grupos sem uso de informação *a priori*, em que dados semelhantes são colocados no mesmo grupo e dados dissemelhantes são colocados em grupos diferentes (Duarte, 2008).

Os algoritmos de agrupamento de dados consideram, tradicionalmente, objetos compostos por um único padrão (instância), pertencendo cada objeto a um único grupo. No entanto, em muitas aplicações do mundo real, os objetos podem ser compostos por um conjunto de padrões e/ou pertencer a vários grupos ou classes. Isto verifica-se especialmente quando as classes naturais de dados não são disjuntas nem imprecisas, mas claramente se sobrepõem (Cleuziou, 2007).

Para fazer face a estas questões apareceram algumas técnicas de aprendizagem multi-grupo, onde o paradigma tradicional de aprendizagem é alterado para que um objeto possa ser associado a vários grupos. Esta situação ocorre em áreas de aplicação importantes, tais como, a recuperação de informação (temáticas diferentes para um único documento), a biologia (várias funções metabólicas para um gene), os *websites* e correspondência eletrónica (documentos de texto associados a várias categorias)(Sorower, 2010).

Existem ainda domínios de aplicação em que um objeto é muitas vezes descrito com um conjunto de vetores de atributos, isto é, podem existir várias instâncias de um único objeto. Neste caso, estamos na presença de um objeto multi-padrão (Dietterich, Lathrop and Lozano-Pérez, 1997). A título de exemplo, uma molécula pode ser representada por um conjunto de vetores de atributos, onde cada vetor descreve uma conformação espacial, ou um *website* pode ser analisado como um conjunto de vetores de palavras correspondentes aos seus documentos HTML (Andrews, Tsochantaridis and Hofmann, 2003). Como resultado, a comunidade científica começou a desenvolver técnicas para aprendizagem multi-padrão.

Nesta dissertação pretende-se estender o estudo do agrupamento de dados ao paradigma multi-padrão e multi-grupo, onde cada objeto a agrupar pode ser composto por vários padrões e/ou pertencer, simultaneamente, a vários grupos. Tratando-se de um problema bastante complexo, é sem dúvida vantajoso explorar as diferentes abordagens existentes. Para tal, será feito um levantamento do estado de arte focado neste paradigma. Seguidamente, serão analisados e implementados alguns dos algoritmos de agrupamento de dados multi-grupo e multi-padrão encontrados. Só serão utilizados algoritmos cuja sua implementação seja aberta e acessível publicamente. Por fim, os desempenhos dos algoritmos estudados serão comparados a fim de identificar os melhores algoritmos para cada situação.

2 Estado de Arte

Este capítulo apresenta os conceitos e aspetos básicos relacionados com esta dissertação, nomeadamente o agrupamento de dados, a aprendizagem multi-padrão, a aprendizagem multi-grupo e a aprendizagem multi-grupo e multi-padrão simultânea. Por fim, apresenta a análise de valor desta dissertação.

2.1 Agrupamento de Dados

O agrupamento de dados é uma técnica de aprendizagem não supervisionada que cria grupos de objetos, de tal forma que objetos pertencentes ao mesmo grupo são similares entre si, enquanto objetos pertencentes a diferentes grupos são dissimilares. Neste tipo de abordagens, não se conhece à partida o número de grupos existentes no conjunto de objetos (a maior parte das vezes), não existem grupos predefinidos nem conjuntos de dados de treino previamente classificados. Esta característica coloca o agrupamento de dados num nível de dificuldade acima dos algoritmos de aprendizagem supervisionada (Hand, 2008).

O agrupamento de dados é um problema difícil pela variedade de situações em que pode ser aplicado e pela complexidade e formas possíveis que os grupos podem assumir. A Figura 1 apresenta um exemplo de um agrupamento de dados. A imagem (a) apresenta sete conjuntos de dados, assinalados com cores em (b). De realçar que, apesar dos grupos em (b) serem claros, nenhum algoritmo de agrupamento conhecido consegue detetar todos esses grupos simultaneamente (Jain, 2010).

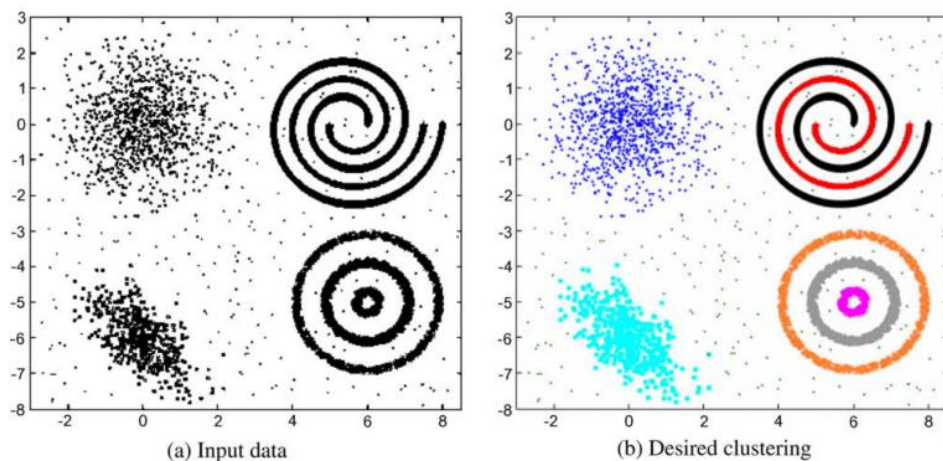


Figura 1 – Exemplo de agrupamento de dados (Jain, 2010)

Apesar das suas raízes estarem mais relacionadas com a estatística, a aprendizagem automática e a biologia, o agrupamento de dados tem sido aplicado a várias áreas como (Duarte, 2008):

- Genética: para encontrar similaridades em genes de determinadas populações ou características de diferentes genes (Liao *et al.*, 2014);
- Marketing: para organizar os clientes em grupos de acordo com os seus interesses, de forma a direcionar produtos específicos ao conjunto de clientes adequado;
- Processamento de Imagem: para agrupar imagens com características semelhantes, por exemplo para classificação de imagens de células cancerígenas (Xu *et al.*, 2012);
- Geografia: na determinação de similaridades entre diferentes pontos da Terra; No planeamento das cidades pela determinação das diferentes zonas económicas de acordo com as características geográficas;
- Astronomia: para identificação de diferentes tipos de astros;
- Sismologia: para identificar falhas geográficas de acordo com o epicentro dos sismos;
- Psicologia: Para identificar grupos de pessoas que possam precisar de tratamentos específicos;
- Informática: Na classificação de documentos ou artigos web de acordo com o seu conteúdo (Kriegel and Schubert, 2004);
- *Data Mining*: Na extração de conhecimento de grandes conjuntos de dados, encontrando relações novas entre objetos ou grupos.

Com a evolução tecnológica que se tem vindo a registar nos últimos anos, a quantidade, variedade e complexidade dos dados disponíveis para análise aumentou exponencialmente.

Como consequência, existe uma necessidade de aperfeiçoar ou encontrar novas técnicas de extração de conhecimento, de forma a melhorar a sua a eficácia e a eficiência.

Acompanhando este requisito, o desenvolvimento de algoritmos de agrupamento de dados tem procurado melhorar os tempos de processamento e a qualidade da resposta para estes conjuntos de dados mais extensos e com mais atributos. Sendo assim, o algoritmo deve ter as seguintes características (Duarte, 2008):

- Escalabilidade: a qualidade da resposta do algoritmo não deve ser alterada pelo tamanho do conjunto de dados a analisar;
- Alta dimensionalidade: o algoritmo deve ser capaz de tratar conjuntos de dados com muitas dimensões ou atributos, sem perder desempenho;
- Ter a capacidade de agrupar dados em grupos de diferentes tamanhos, caso isso seja adequado ao conjunto analisado;
- Estar apto para tratar diferentes tipos de atributos, sejam eles binários, nominais, ordinais, contínuos ou conjuntos de diferentes tipos;
- Robustez e sensibilidade ao ruído: muitas vezes os conjuntos de dados contêm valores isolados, dados com erros ou desconhecidos. O algoritmo deve ser capaz de excluir estas situações, de forma a não introduzir erro no seu processamento;
- Estabilidade: a resposta do algoritmo deve ser a mesma, independentemente da ordem dos dados de entrada;
- Agrupamento baseado em restrições: possibilitar a inserção de regras que, por exemplo, simulam um comportamento do mundo real, mantendo a sua qualidade de resposta;
- Interpretabilidade e utilidade: a apresentação dos resultados do algoritmo deve facilitar a sua interpretação por aplicações externas ou mesmo por humanos.

Um desenvolvedor de algoritmos de agrupamento de dados deve ter em conta todas estas características aquando do desenho do mesmo. No entanto, o conjunto de algoritmos de agrupamento disponíveis atualmente mostra que pode ser difícil satisfazer todas as características num só algoritmo.

Normalmente, um processo de agrupamento de dados apresenta os seguintes passos (Jain, Murty and Flynn, 1999):

1. Representação dos objetos de dados (opcionalmente pode incluir extração e/ou seleção de atributos);
2. Definição de uma medida de similaridade entre objetos apropriada ao domínio de dados;
3. Agrupamento de dados;
4. Abstração de dados, se necessário;
5. Avaliação dos grupos obtidos, se necessário.

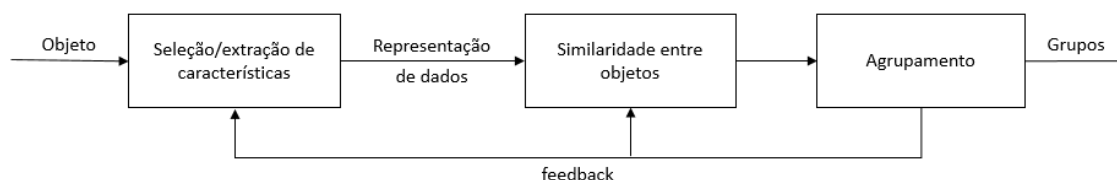


Figura 2 – Fases do agrupamento de dados

A fase de representação dos objetos de dados define um conjunto de parâmetros que os caracterizam, como número de objetos, variedade e quantidade de atributos ou até o número de grupos previstos. De notar que o utilizador pode não ter controlo sobre estes parâmetros. Os processos de extração e seleção de atributos podem ser pré-aplicados ao conjunto de dados com o propósito de melhorar o resultado produzido pelo algoritmo. A extração de atributos consiste na transformação de atributos originais, de forma a gerar novos atributos. A seleção de atributos consiste na identificação do subconjunto dos atributos originais que melhor descrevem o conjunto de dados.

A similaridade entre objetos, também conhecida por medida de proximidade ou dissimilaridade, é normalmente definida por uma função de cálculo de distâncias entre pares de objetos com o intuito de determinar quão similares são esses objetos. O tipo de função utilizada varia com o tipo de atributos caracterizados (contínuos, binários, nominais, ordinais ou uma combinação destes). Seguidamente são enumeradas algumas das medidas mais usadas para determinar a proximidade entre objetos (Irani, Pise and Phatak, 2016):

- Distância Euclidiana: distância entre dois pontos, em linha reta. Caminho mais perto entre dois pontos;
- Distância de Manhattan: soma das diferenças absolutas entre as coordenadas dos dois pontos;

- Distância de Hausdorff: maior distância entre os pontos mais próximos de dois subconjuntos de dados;
- Distância de Chebyshev: maior das diferenças entre as coordenadas de dois pontos.

Existem várias classificações propostas para os algoritmos de agrupamento. Geralmente, estes podem ser divididos em duas grandes categorias (Hand, 2008):

- Agrupamento exclusivo/rígido: um objeto pertence a um e um só grupo;
- Agrupamento difuso/*fuzzy*: um objeto pode pertencer a um ou mais grupos com determinada probabilidade.

O agrupamento exclusivo pode ainda ser dividido em subcategorias. As mais comuns são (Hand, 2008):

- Algoritmos de partição: k-Médias (MacQueen, 1967), k-Medóides (Kaufman and Rousseeuw, 1987);
- Algoritmos hierárquicos: CURE (Guha, Rastogi and Shim, 1998);
- Algoritmos baseados em densidade: DBSCAN (Ester *et al.*, 1996);
- Algoritmos baseados em grelha: STING (Wang, Yang and Muntz, 1997);
- Algoritmos baseados em modelos: EM (Dempster, Laird and Rubin, 1977).

A abstração de dados é o processo de adaptação dos resultados do agrupamento de forma que estes possam ser apresentados de forma rápida e intuitiva para elementos externos, sejam estes máquinas ou humanos.

A avaliação de um agrupamento de dados tem duas abordagens. Uma delas consiste em determinar o número de grupos existentes no conjunto de dados. Dados que não contenham grupos não devem ser submetidos a agrupamento. No entanto, há algoritmos de agrupamento que não necessitam de saber à partida o número de grupos existentes nos dados. A segunda abordagem consiste na validação dos resultados obtidos, seja pela qualidade do agrupamento, pela performance computacional do algoritmo ou por outros fatores relacionados com o conjunto analisado.

2.2 Aprendizagem Multi-Padrão

Este subcapítulo apresenta o conceito de aprendizagem multi-padrão. Esta pode ser supervisionada ou não supervisionada.

2.2.1 Aprendizagem supervisionada Multi-Padrão

A aprendizagem supervisionada Multi-Padrão (Multi-Instance Learning ou MIL) foi introduzida por (Dietterich, Lathrop and Lozano-Pérez, 1997) como uma variação à aprendizagem supervisionada tradicional, com o intuito de analisar situações em que o sistema de aprendizagem tem um conhecimento parcial ou incompleto sobre os dados de treino. Na aprendizagem supervisionada tradicional (Figura 3), um conjunto de dados de treino constituído por pares objeto-grupo é usado para construir uma função de classificação que permita prever a que grupo pertence um novo objeto. Os pares objeto-grupo são, muitas vezes, de difícil obtenção em certas áreas.

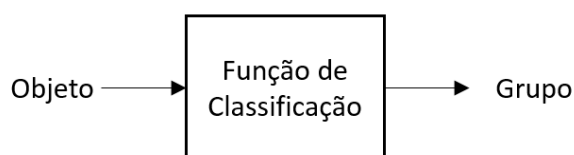


Figura 3 – Aprendizagem Supervisionada

No paradigma de aprendizagem supervisionada multi-padrão (Figura 4), o conjunto de dados de treino consiste em “sacos” classificados, cada um contendo um conjunto de padrões (também conhecido por instâncias ou vetores de atributos) não classificados. O algoritmo determina quais os padrões responsáveis pela classificação do saco. Na sua forma tradicional, o saco é marcado como pertencendo a um grupo quando pelo menos um dos seus padrões pertence ao grupo. O saco é marcado como não pertencendo a um grupo se nenhum dos seus padrões pertencer ao grupo (Babenko, 2008). Assim, o objetivo da aprendizagem supervisionada multi-padrão é prever a que grupo pertence um novo saco e que instâncias são responsáveis por essa classificação.

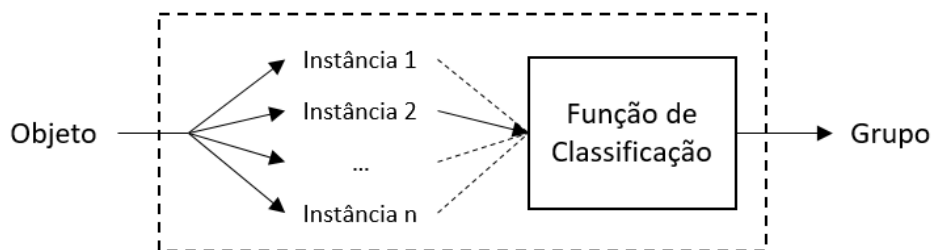


Figura 4 – Aprendizado Multi-Padrão

Os autores de (Dietterich, Lathrop and Lozano-Pérez, 1997) exemplificaram este paradigma numa situação real. Imagine-se que existem várias pessoas, cada uma delas com um porta-chaves com algumas chaves. Sabe-se apenas quais as pessoas que conseguem abrir a porta de um certo quarto. A tarefa de aprendizagem seria prever que chave e, por consequência, que porta-chaves conseguem abrir a porta do quarto. O processo de aprendizagem passaria por descobrir vetores de atributos comuns (chaves) nos diferentes objetos (porta-chaves) classificados de forma positiva (consegue abrir a porta), criando assim uma função de classificação. O algoritmo poderá então classificar um porta-chaves novo, analisando as suas instâncias (chaves).

Esta forma de previsão tradicional utiliza métodos de classificação binários com a intenção de qualificar os sacos positiva ou negativamente. Novas formas de previsão generalizada foram estudadas por (Weidmann, Frank and Pfahringer, 2003) e (Scott, Zhang and Brown, 2005), onde um saco é considerado positivo quando satisfaz um conjunto de restrições mais sofisticado do que apenas uma das suas instâncias pertencer a um grupo. Estas generalizações permitem a aplicação deste conceito a um grupo mais alargado de situações (Zhang and Zhou, 2009).

A aprendizagem supervisionada multi-padrão foi inicialmente introduzida no estudo das características e efeitos de fármacos (Dietterich, Lathrop and Lozano-Pérez, 1997), onde cada molécula pode assumir várias conformações espaciais e o algoritmo tenta associar essas conformações a certas características dos fármacos. Outras aplicações são a classificação de imagens (Xu *et al.*, 2012), a classificação de *websites* (Zhou, Jiang and Li, 2005), o *data mining* (Kriegel, Pryakhin and Schubert, 2006) e a genética (Liao *et al.*, 2014).

Seguidamente, são enumerados alguns algoritmos que aplicam o paradigma de aprendizagem supervisionada multi-padrão:

- MI-SVM (Andrews, Tsochantaridis and Hofmann, 2003);
- Citation-kNN (Wang and Zucker, 2000);
- Diverse Density (Maron and Lozano-Pérez, 1997);
- MILBoosting (Auer and Ortner, 2004).

2.2.2 Aprendizagem não-supervisionada Multi-Padrão

Em (Zhang and Zhou, 2009), o conceito de aprendizagem não supervisionada multi-padrão foi explorado pela primeira vez, concretamente com a técnica de agrupamento multi-padrão (Multiple Instance Clustering ou MIC). Nesta abordagem, não se tem conhecimento do grupo a que cada objeto multi-padrão pertence. Existem duas razões para usar a aprendizagem não supervisionada multi-padrão em vez da supervisionada multi-padrão:

1. Em algumas situações, o processo de obtenção de classificação dos objetos é muito trabalhoso ou demasiado dispendioso. Usando o exemplo dos fármacos, podem-se facilmente desenhar novas moléculas, cada uma representada por um saco de padrões correspondentes às diferentes conformidades espaciais. No entanto, o processo de classificação dos padrões requer a realização de experiências bioquímicas possivelmente complexas e demorosas, podendo os custos associados serem muito elevados.
2. Apesar do processo de aprendizagem supervisionada multi-padrão já relacionar um saco de vetores de atributos com um grupo, ao adicionar-se um processo de agrupamento é produzida informação sobre os dados que poderá influenciar o procedimento de predição. Assim, uma das vantagens da aprendizagem não supervisionada consiste na possibilidade de ajudar a encontrar relações novas entre as estruturas analisadas.

Alguns problemas tradicionais de agrupamento podem ser melhor formulados usando uma abordagem multi-padrão. Por exemplo, em agrupamento de imagens, existe sempre um determinado grau de ambiguidade associado à escolha da parte da imagem que serve como ponto comum. Essa imagem (Figura 5) pode ser tratada como um saco em que cada uma das suas instâncias corresponde a uma região da imagem. Seguidamente, pode ser aplicado um algoritmo de agrupamento multi-padrão para particionar esses sacos.



Figura 5 – Exemplo de instâncias de um objeto multi-padrão

Os algoritmos disponíveis para este contexto utilizam uma medida de similaridade entre sacos para agrupar sacos semelhantes, idêntica às que muitos algoritmos de agrupamento tradicionais utilizam. No entanto, em problemas multi-padrão, pelo facto de existirem várias instâncias que caracterizam um único objeto, o problema de encontrar a medida de similaridade correta torna-se mais complexo, especialmente porque a relação entre instâncias do mesmo objeto pode corresponder a uma classificação diferente (Herrera *et al.*, 2016). Para que seja gerada uma boa classificação, os objetos não podem ser vistos como simples coleções de instâncias independentes, devendo se ter em consideração as características das relações entre essas instâncias.

Seguidamente, são enumerados alguns algoritmos que aplicam o paradigma de aprendizagem não supervisionada multi-padrão:

- BAGIC (Zhang and Zhou, 2009);
- BARTMIP (Zhang and Zhou, 2009);
- M³IC-MBM (Zhang *et al.*, 2009);
- MCIL (Xu *et al.*, 2012);
- UC-kNN (Henegar, Clément and Zucker, 2006);
- COSMIC (Kriegel *et al.*, 2006).

2.3 Aprendizagem Multi-Grupo

Este subcapítulo apresenta o conceito de aprendizagem multi-grupo. Esta pode ser supervisionada ou não supervisionada.

2.3.1 Aprendizagem supervisionada Multi-Grupo

Uma das principais abordagens de aprendizagem supervisionada é a classificação tradicional, em que um objeto é caracterizado por um vetor de atributos e está associado a um e um só grupo. O número de grupos existentes no conjunto de dados é conhecido à partida. Se o número de grupos nesse conjunto for dois, essa é uma classificação binária. Se o número de grupos for superior a dois, então dá-se o nome de classificação multi-classe (Tsoumakas and Katakis, 2009).

Apesar da classificação tradicional ser uma das técnicas de aprendizagem automática mais estudada e com melhores resultados, existem situações em que esta não se adequa, nomeadamente na classificação de objetos que podem ser associados a vários grupos. Por exemplo, o filme “O Código de Da Vinci” pode ser caracterizado como religioso, arte e ficção, simultaneamente. Na categorização de músicas, uma sinfonia pode ser associada a piano, música clássica, Mozart e Áustria.

Para representar corretamente estes e outros exemplos de objetos do mundo real que podem assumir diferentes interpretações, surgiu o conceito de aprendizagem supervisionada multi-grupo. Ao contrário da aprendizagem supervisionada tradicional, o objetivo da aprendizagem supervisionada multi-grupo é associar objetos a um ou mais grupos pertencentes a um determinado conjunto de grupos (Sorower, 2010).

Inicialmente, a aprendizagem supervisionada multi-grupo esteve principalmente ligada à categorização de textos (McCallum, 1999). Mais tarde, o conceito foi ganhando força nas comunidades relacionadas com a aprendizagem automática, estando agora associado a diversos problemas como *websites* (Tang, Rajan and Narayanan, 2009), bioinformática (Liao *et al.*, 2014), *data mining* (Tsoumakas, Katakis and Vlahavas, 2009), ou classificação automática de conteúdos multimédia, como por exemplo imagens (Nasierding, Tsoumakas and Kouzani, 2009).

Um dos problemas que pode surgir com a aprendizagem multi-grupo consiste no enorme número de resultados (combinações de grupos) que o algoritmo pode ter (Zhang and Zhou, 2014). Por exemplo, para um conjunto de 20 grupos, o número de combinações únicas de grupos possível excede o milhão (2^{20}), o que pode ter grande impacto no tempo de execução do algoritmo. Para lidar com este aumento exponencial de resultados, os algoritmos exploram a correlação ou dependência entre grupos. Por exemplo, a probabilidade de uma imagem receber a anotação “Portugal” poderá ser maior se se souber que esta pertence aos grupos “praia” e “futebol”. Existem algumas estratégias que exploram a relação entre grupos, sendo

geralmente organizadas em três famílias segundo a ordem de correlação usada (Tsoumakas and Katakis, 2009):

- Estratégia de primeira ordem: As tarefas de aprendizagem multi-grupo são feitas grupo a grupo, ignorando a existência de grupos relacionados, decompondo assim o problema em questão num conjunto de classificações binárias independentes, uma por grupo. Esta estratégia é conceptualmente simples e eficiente, mas peca na qualidade dos resultados;
- Estratégia de segunda ordem: São analisadas relações entre pares de grupos, atingindo melhor qualidade de resultados;
- Estratégia de alta ordem: Considera as relações entre grupos de forma a que se analise a influência de um grupo em todos os outros grupos, ou apenas relações entre conjuntos aleatórios de grupos. Naturalmente, esta estratégia exige maiores recursos operacionais, mas apresenta melhores resultados.

Os algoritmos de aprendizagem supervisionada multi-grupo podem ser divididos em duas categorias (Zhang and Zhou, 2014):

1. Algoritmos por transformação do problema: Um problema multi-grupo é transformado num ou mais problemas de classificação tradicional ou regressão. Alguns exemplos são:
 - a. Primeira Ordem: Binary Relevance (Boutell *et al.*, 2004);
 - b. Segunda Ordem: Calibrated Label Ranking (Fürnkranz *et al.*, 2008);
 - c. Alta Ordem: Classifier Chains (Read *et al.*, 2009) e Random k-labelsets (Tsoumakas and Vlahavas, 2007).
2. Algoritmos por adaptação do algoritmo: Estendem algoritmos de aprendizagem conhecidos de forma a estes poderem lidar com um problema multi-grupo. Alguns exemplos são:
 - a. Primeira Ordem: ML-kNN que adapta o conceito de aprendizagem preguiçosa (Zhang and Zhou, 2007) e ML-DT que adapta as árvores de decisão (Clare and King, 2001);
 - b. Segunda Ordem: Rank-SVM que adapta técnicas kernel (Elisseeff and Weston, 2001) e CML que adapta técnicas de informação teórica (Ghamrawi and McCallum, 2005).

2.3.2 Aprendizagem não supervisionada Multi-Grupo

O agrupamento de dados também tem sido recentemente explorado com a vertente de aprendizagem multi-grupo (Nasierding, Tsoumakas and Kouzani, 2009; Megano *et al.*, 2015; Zhang and Wu, 2015). Numa primeira fase, os dados são agrupados com uma técnica de agrupamento conhecida (por exemplo k-médias (MacQueen, 1967) ou EM (Dempster, Laird and Rubin, 1977)), sendo seguidamente classificados de acordo com o conceito multi-grupo. Esta abordagem tem sido aplicada em vários problemas de aprendizagem, por exemplo:

- Atribuição automática de anotações a imagens;
- Reconhecimento facial;
- Categorização de páginas Web;
- Determinação das funcionalidades das proteínas;
- Categorização de músicas.

Seguidamente, são enumerados alguns algoritmos que aplicam o paradigma de aprendizagem não supervisionada Multi-Grupo:

- LIFT (Zhang and Wu, 2015);
- CBMLC (Nasierding, Tsoumakas and Kouzani, 2009);
- CLMLC (Sun, Kudo and Kimura, 2016);
- HOMER (Tsoumakas, Katakis and Vlahavas, 2008).

2.4 Aprendizagem Multi-Grupo e Multi-Padrão simultânea

O conceito de aprendizagem supervisionada multi-grupo e multi-padrão simultânea (*Multi-Instance Multi-Label* ou MIML) foi introduzido em (Zhou and Zhang, 2006) como alternativa à aprendizagem supervisionada tradicional. Enquanto a aprendizagem multi-padrão explora a ambiguidade nos dados de entrada e a aprendizagem multi-grupo explora a ambiguidade nos dados de saída, o paradigma MIML explora ambas as ambiguidades de entrada e de saída. Sendo assim, um objeto descrito por um saco de vetores de atributos (ou padrões) está associado a um ou mais grupos. A Figura 6 ilustra as diferenças entre a aprendizagem supervisionada tradicional, multi-padrão, multi-grupo e multi-grupo e multi-padrão simultânea.

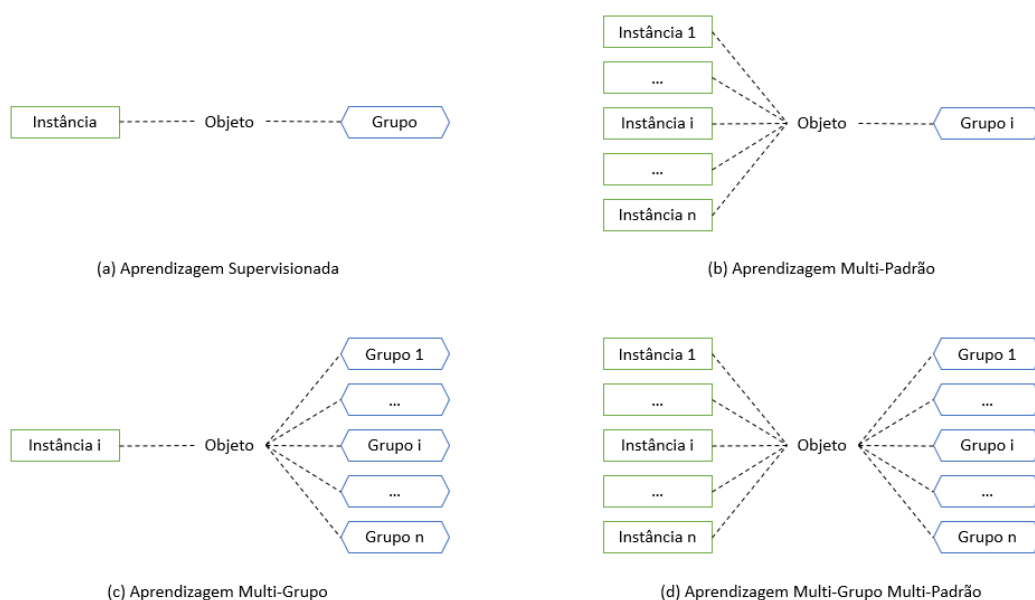


Figura 6 – Comparação dos componentes da aprendizagem MIML

Como se pode depreender da Figura 6, a aprendizagem supervisionada tradicional pode ser vista como uma forma simplificada da aprendizagem supervisionada multi-grupo e da aprendizagem supervisionada multi-padrão. Isto porque na aprendizagem supervisionada tradicional um objeto só é constituído por uma instância e só pertence a um grupo, enquanto que na aprendizagem supervisionada multi-grupo um objeto pode pertencer a mais do que um grupo e na aprendizagem supervisionada multi-padrão um objeto pode ser constituído por várias instâncias. Por sua vez, estas duas últimas abordagens podem ser vistas como versões simplificadas da aprendizagem supervisionada multi-grupo e multi-padrão simultânea. Uma das formas de resolver um problema de aprendizagem supervisionada multi-grupo e multi-padrão simultânea é transformá-lo num problema de aprendizagem supervisionada tradicional usando a aprendizagem multi-grupo ou a aprendizagem multi-padrão como ponte.

A aprendizagem MIML foi demonstrada com sucesso na classificação de imagens e na caracterização de textos, onde os objetos analisados podem assumir estruturas de dados ambíguas (Zhou and Zhang, 2006; Feng and Xu, 2010) Mais tarde, a sua utilização foi alargada à anotação automática de vídeo e áudio (Xu, Xue and Zhou, 2011).

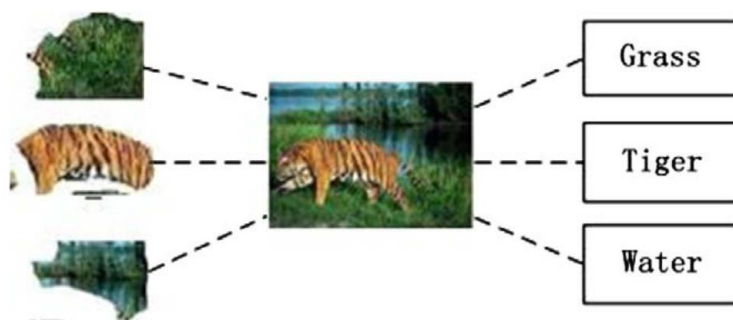


Figura 7 – Exemplo de aprendizagem MIML em classificação de imagens (Feng and Xu, 2010)

De seguida, são enumerados alguns algoritmos supervisionados que aplicam o paradigma de aprendizagem multi-grupo e multi-padrão simultaneamente:

- MIMLBOOST (Zhou and Zhang, 2006);
- MIMLSVM (Zhou and Zhang, 2006);
- M³MIML (Zhang and Zhou, 2008)
- En-MIMLSVM (Xu, Xue and Zhou, 2011);
- MIMLNN(Zhou *et al.*, 2012).
- MIML DML (Rong Jin, Shijun Wang and Zhi-Hua Zhou, 2009).
- MIHC (Liao *et al.*, 2014);

Não foram encontrados algoritmos de aprendizagem multi-grupo e multi-padrão simultânea não supervisionados. Seguidamente são apresentados alguns dos algoritmos supervisionados acima mencionados.

2.4.1 MIMLBOOST

O MIMLBOOST é um algoritmo de aprendizagem supervisionada multi-grupo e multi-padrão desenvolvido por (Zhou and Zhang, 2006). Este algoritmo transforma um problema MIML em vários problemas de aprendizagem supervisionada tradicional usando aprendizagem multi-padrão como ponte.

Numa primeira fase, o MIMLBOOST usa *category wise decomposition* para transformar a componente multi-grupo numa componente de grupo único. Seguidamente, usa o algoritmo MIBOOSTING para analisar a componente multi-padrão resultante.

2.4.2 MIMLSVM

O MIMLSVM é um algoritmo de aprendizagem supervisionada multi-grupo e multi-padrão também desenvolvido por (Zhou and Zhang, 2006). Tal como o MIMLBOOST, este algoritmo transforma um problema MIML em vários problemas de aprendizagem supervisionada tradicional, mas usando aprendizagem multi-grupo como ponte.

Numa primeira fase, o MIMLSVM usa *clustering-based representation transformation* para transformar a componente multi-padrão numa componente de padrão único. Seguidamente, usa o algoritmo MLSVM para analisar a componente multi-grupo resultante.

Mais tarde, (Ying-Xin Li *et al.*, 2012) propuseram duas extensões deste algoritmo, o MIMLSVM+ e o E-MIMLSVM+.

2.4.3 En-MIMLSVM

O En-MIMLSVM é um algoritmo de aprendizagem multi-grupo e multi-padrão criado por (Xu, Xue and Zhou, 2011), baseado no algoritmo MIMLSVM. Tem a particularidade de utilizar métodos de aprendizagem *ensemble*.

Inicialmente, o algoritmo constrói pequenas amostras do conjunto de dados original. Posteriormente, treina vários classificadores usando essas amostras do conjunto de dados. Por fim, todos os classificadores são combinados de forma a obter a classificação final.

2.5 Outros conceitos relacionados com processamento de dados

Nesta secção estão apresentados sucintamente conceitos que, apesar de não estarem diretamente ligados ao âmbito desta dissertação, são importantes para a boa compreensão dos processos aplicados e dos seus resultados.

2.5.1 Cross Validation

Cross Validation, também conhecida em português como validação cruzada, é uma técnica utilizada para a avaliação de modelos de dados a partir de um conjunto de dados. Esta técnica consiste na divisão do conjunto de dados em dados de treino e dados de teste. Os dados de treino são usados para treinar o modelo, enquanto que os dados de teste são usados para validar o modelo (Kohavi and Kohavi, 1995).

2.5.2 k-fold

O *k-fold* é um dos métodos de *cross validation* e consiste em dividir um conjunto de dados em k conjuntos. A cada uma das k iterações, um dos conjuntos é utilizado para teste, enquanto que os outros $k-1$ conjuntos são utilizados para treino (Kohavi and Kohavi, 1995).

3 Algoritmos de Agrupamento Multi-Padrão

Este capítulo descreve os algoritmos de agrupamento de dados multi-padrão que foram encontrados na literatura.

3.1 Agrupamento Multi-Padrão

3.1.1 BAMIC

O BAMIC (*B*Ag-level *M*ulti-*I*nstance *C*lustering) é um algoritmo de agrupamento de dados multi-padrão criado por (Zhang and Zhou, 2009) para a resolução de problemas de aprendizagem não-supervisionada multi-padrão, onde não se conhece *a priori* os grupos a que os “sacos” pertencem.

Nesta abordagem, o problema de agrupamento de dados multi-padrão não pode ser simplesmente tomado como uma extensão dos problemas de agrupamento de dados tradicionais, onde os objetos a serem agrupados são agora conjuntos de instâncias em vez de apenas uma instância. As relações entre instâncias e as suas características individuais não podem ser desprezadas, especialmente porque estas podem exibir diferentes funcionalidades. Por exemplo, na classificação de medicamentos, apenas uma ou algumas das conformidades espaciais da molécula seriam responsáveis por essa classificação.

O BAMIC tenta dividir os sacos não classificados em k grupos disjuntos de sacos, usando uma medida de distância para calcular a relação entre os sacos, seguindo um algoritmo de agrupamento de dados tradicional para encontrar esses k grupos disjuntos. As medidas de distância usadas são a distância máxima de Hausdorff, a distância mínima de Hausdorff e a

distância média de Hausdorff. O processo de agrupamento é feito pelo conhecido algoritmo *k-Medóides* (Kaufman and Rousseeuw, 1987). A Tabela 1 apresenta uma descrição completa do algoritmo BAMIC.

Tabela 1 – Pseudocódigo descrevendo o algoritmo BAMIC

<ul style="list-style-type: none"> • Dados de Entrada: <ul style="list-style-type: none"> ○ X – conjunto de dados multi-padrão não classificado $\{X_1, X_2, \dots, X_N\}$ ○ k – número de grupos ○ <i>Bag_dist</i> – medida de distância • Dados de Saída: <ul style="list-style-type: none"> ○ Grupos – Agrupamento de dados final $\{G_1, G_2, \dots, G_N\}$ ○ <i>Medóides</i> – Medóides dos grupos encontrados • Procedimento: <ol style="list-style-type: none"> 1. Selecionar aleatoriamente k sacos de treino como medóides iniciais C_j ($1 \leq j \leq k$); 2. Repetir <ol style="list-style-type: none"> a. Para $j \in \{1, 2, \dots, k\}$ fazer <ol style="list-style-type: none"> i. $G_j = \{C_j\}$; b. Para $i \in \{1, 2, \dots, N\}$ fazer <ol style="list-style-type: none"> i. $index = \arg \min_{j \in \{1, 2, \dots, k\}} Bag_dist(X_i, C_j)$; ii. $G_{index} = G_{index} \cup \{X_i\}$; c. Para $j \in \{1, 2, \dots, k\}$ fazer <ol style="list-style-type: none"> i. $C_j = \arg \min_{X_k \in G_j} (\sum_{B \in G_j} Bag_dist(X_i, X_k) / G_j)$; <p>Até {o resultado do agrupamento não mudar};</p> 3. $Grupos = \{G_j 1 \leq j \leq k\}$; $Medóides = \{C_j 1 \leq j \leq k\}$;

3.1.2 BARTMIP

O BARTMIP (BAg-level Representation Transformation for Multi-Instance Prediction) é um algoritmo de agrupamento de dados multi-padrão criado por (Zhang and Zhou, 2009) que assume que os grupos dos sacos multi-padrão estão relacionados com a distância entre sacos.

O algoritmo à a possibilidade de calcular a distância entre sacos com a medida de distância máxima de *Hausdorff*, com a distância mínima de *Hausdorff* e com a distância máxima de *Hausdorff*.

O algoritmo BARTMIP aplica inicialmente um agrupamento dos sacos de instâncias, aplicandos o algoritmo BAGIC (uma variação do algoritmo *k-Medóides* (Kaufman and Rousseeuw, 1987) para conjuntos de dados multi-padrão) com uma das medidas de distância de *Hausdorff*. Seguidamente, a informação dos grupos originais é adicionada aos sacos já agrupados. Pro fim, toda esta informação passa por um algoritmo de aprendizagem supervisionada tradicional que cria o modelo de classificação do algoritmo. A Tabela 2 apresenta uma descrição completa do algoritmo BARTMIP.

Tabela 2 – Pseudocódigo descrevendo o algoritmo BARTMIP

$Label = BARTMIP(D, k, Bag_{dist}, TBag, Learner)$

- Dados de Entrada:
 - D – conjunto de dados multi-padrão de treino
 $\{(X_1, L_{X_1}), (X_2, L_{X_2}), \dots, (X_N, L_{X_N})\}, (X_i \subseteq X, L_{X_i} \in \{0, 1\} \text{ ou } [0, 1])$
 - k – número de grupos
 - Bag_{dist} – medida de distância de Hausdorff (máxima, mínima, ou média)
 - $TBag$ – conjunto de dados de teste ($TBag \subseteq X$)
 - $Learner$ – algoritmo de aprendizagem supervisionada
- Dados de Saída:
 - $Label$ – Grupos previstos para o conjunto de dados de teste, $Label \in \{0, 1\} \text{ ou } [0, 1]$
- Procedimento:
 1. Definir $U = \{X_1, X_2, \dots, X_N\}$

2. $[\{G_1, G_2, \dots, G_k\}, \{C_1, C_2, \dots, C_k\}] = \text{Bamic}(U, k, \text{Bag dist});$
 3. Definir $Tr = 0$
 4. Para $i \in \{1, 2, \dots, N\}$ fazer
 - a. Para $j \in \{1, 2, \dots, k\}$ fazer
 - i. $y_j^i = \text{Bag dist}(X_i, C_j)$
 - b. $Tr = Tr \cup \{(\langle y_1^i, y_2^i, \dots, y_k^i \rangle, L_{X_i})\}$
 5. $\text{Predictor}(\cdot) = \text{Learner}(Tr)$
 6. Para $j \in \{1, 2, \dots, k\}$ fazer
 - a. $y_i = \text{Bag dist}(T\text{Bag}, C_j)$
 7. $\text{Label} = \text{Predictor}(\langle y_1, y_2, \dots, y_k \rangle)$
-

3.1.3 M³IC-MBM

O M³IC-MBM (*Maximum Margin Multiple Instance Clustering - Modified Bag Margin*) é um algoritmo de agrupamento de dados multi-padrão criado por (Zhang *et al.*, 2009).

Este algoritmo utiliza os conceitos de *Constrained Concave-Convex Procedure (CCCP)* e *Cutting Plane* para relacionar os diferentes sacos de instância. A Tabela 3 apresenta uma descrição completa do algoritmo M³IC-MBM.

Tabela 3 – Pseudocódigo do algoritmo M³IC-MBM

-
- Dados de Entrada:
 - $\{X_1, X_2, \dots, X_N\}$ - Conjunto de dados multi-padrão não classificados ("sacos de instâncias")
 - C – Constance de Regularização
 - ϵ_1 – Precisão da solução CCCP
 - ϵ_2 – Precisão da solução *cutting plane*
 - k – Número de grupos
 - l – Equilíbrio do tamanho dos grupos
 - Dados de Saída:
 - Vetor de atribuição de grupos f
 - Procedimento:
 1. Contruir $\tilde{\beta} = \{X_{ij(r)}\}$
 2. Inicializar \tilde{w}^0 , $t = 0$, $\Delta J = 10^{-3}$, $J^{-1} = 10^{-3}$
 3. Enquanto $\Delta J / J^{t-1} > \epsilon_1$ fazer
 - a. Derivar Problema $P = \min_{\tilde{w}, \xi \geq 0} \frac{1}{2} \|\tilde{W}\|^2 + C\xi$
 - b. Definir $\Omega = \emptyset$, $\forall 1 \leq i \leq n$, $ci = 0$, $s = -1$
 - c. Enquanto H^{t_s} é verdadeiro, fazer
 - i. $s = s + 1$

ii. Obter $(\tilde{w}^{(t_s)}, \xi^{(t_s)})$ ao resolver P com Ω

iii. Encontrar os sacos mais violados $c_i^{t_s}$

$$c_i^{t_s} = \begin{cases} 1, & \text{se } (\tilde{w}^{(t_s)})^T \frac{\partial f(\tilde{w}, i)}{\partial \tilde{w}} \Big|_{\tilde{w} = \tilde{w}^{(t)}} \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

iv. Atualizar $\Omega = \Omega \cup c^{t_s}$

Fim Enquanto

d. $t = t + 1$

e. $\tilde{w}^{(t)} = \tilde{w}^{(t-1)_s}$

f. $\Delta J = J^{t-1} - J^t$

Fim Enquanto

4. Agrupamento: Para um saco X_i , $f_i = \arg \max_p (\tilde{w}^{(t)})^T X_{ij^*(p)}$, em que

$$j^* = \arg \max_{j \in X_i} (\max_u (\tilde{w}^{(t)})^T X_{ij(u)} - \text{mean}_v ((\tilde{w}^{(t)})^T X_{ij(v)}))$$

3.1.4 MCIL

O MCIL (*Multiple Clustered Instance Learning*) é um algoritmo de agrupamento de dados multi-padrão criado por (Xu *et al.*, 2012), ligado à classificação de tipos de tecidos cancerígenos em imagens histopatológicas. Apesar de ter sido usado neste domínio, o MCIL pode ser aplicado a outros problemas de agrupamento multi-padrão.

O algoritmo foi desenvolvido com dois objetivos: classificar imagens como “cancro” ou “sem cancro” e, posteriormente, identificar subclasses dos vários tipos de tecido cancerígeno.

O primeiro objetivo constitui um problema de aprendizagem supervisionada, uma vez que o conjunto de sacos (imagens) de entrada chega já classificado como “cancro” ou “sem cancro”. No entanto, é desconhecida a relação entre os padrões de cada saco e a sua classificação, sendo o primeiro procedimento do algoritmo encontrar esta relação.

O segundo objetivo constitui um problema de aprendizagem não supervisionada, uma vez que se desconhecem as subclasses de tecido cancerígenos que existem no conjunto de sacos de instâncias, sendo que os sacos podem pertencer a várias subclasses ao mesmo tempo. Para tal, o algoritmo define k funções de classificação h para indicarem a probabilidade de cada padrão dos sacos pertencer ao grupo. A Figura 7 apresenta uma descrição completa do algoritmo MICL.

Tabela 4 – Pseudocódigo do algoritmo MICL

-
- Dados de Entrada:
 - $\{X_1, X_2, \dots, X_N\}, X_i = \{x_{i1}, \dots, x_{im}\}$ - Conjunto de dados multi-padrão não classificados (“sacos de instâncias”)
 - $\{y_1, y_2, \dots, y_N\}, y_i \in Y = \{-1, 1\}$ - Conjunto de Grupos
 - K – Número de grupos (subclasses)
 - T – Número de iterações
 - Dados de Saída:
 - Classificadores h^1, h^2, \dots, h^k
 - Procedimento:
 1. Para $t = 1 \rightarrow T$ fazer
 - a. Para $k = 1 \rightarrow K$ fazer

- i. Calcular pesos $w_{ij}^k = -\frac{\partial \mathcal{L}}{\partial h_{ij}^k}$
- ii. Treinar classificadores fracos h_t^k usando pesos $|w_{ij}^k|$
- iii. $h_t^k = \arg \min_h \sum_{ij} 1(h(x_{ij}^k) \neq y_i | w_{ij}^k|)$
- iv. $\alpha_t^k = \arg \min_{\alpha} \mathcal{L}(\cdot, h^k + \alpha h_t^k, \cdot)$
- v. Atualizar classificadores fortes $h^k \leftarrow h^k + \alpha_t^k h_t^k$

Fim Para

Fim Para

3.1.5 UC-kNN

O UC-kNN (*Unsupervised Citation k-Nearest Neighbor*) é um algoritmo de agrupamento de dados multi-padrão criado por (Henegar, Clément and Zucker, 2006). O algoritmo tem como base o método de agrupamento hierárquico de dados Citation-kNN, desenvolvido por (Wang and Zucker, 2000).

O seu desenvolvimento foi motivado pela intenção dos autores encontrarem soluções para problemas complexos de aprendizagem não supervisionada, utilizando um esquema de dados multi-padrão que consideram simples e robusto. Era também objetivo dar a possibilidade de resolver problemas de agrupamento complexos, visando descobrir comportamentos desconhecidos entre objetos caracterizados por vários tipos de atributos diferentes (p. e., numéricos, simbólicos, binários).

Para demonstrar a funcionalidade do UC-kNN, os autores aplicaram-no ao ramo da genética, utilizando representações multi-padrão das moléculas de ARN, com o objetivo concreto de identificar padrões biológicos envolvidos em processos celulares.

3.1.6 COSMIC

O COSMIC (*CO*nceptual *Specified Multi-Instance Clusters*) é um algoritmo de agrupamento de dados multi-padrão baseado em densidade criado por (Kriegel *et al.*, 2006). Este tipo de algoritmos tem como principal vantagem a possibilidade de encontrar grupos de objetos organizados numa forma não esférica. Sendo assim, este algoritmo tem a capacidade de distinguir zonas com agrupamento denso de zonas com agrupamento mais disperso.

O COSMIC consiste em dois grandes passos. Primeiro, o algoritmo executa um agrupamento baseado em densidade e obtém um conjunto de grupos de todas as instâncias (padrões) dos objetos. No segundo passo, o COSMIC utiliza informações específicas dos algoritmos baseados em densidade, obtidas no primeiro passo, para caracterizar todos os grupos de objetos encontrados no conjunto de dados analisado.

3.2 Conjuntos de dados Multi-Padrão

Nesta secção estão descritos alguns dos conjuntos de dados mais utilizados em demonstrações de performance de algoritmos multi-padrão e que foram usados neste trabalho no agrupamento multi-padrão. A Tabela 5 apresenta um resumo das características desses conjuntos de dados.

Tabela 5 – Descrição de conjuntos de dados multi-padrão usados

Dataset	Atributos	Sacos Pos.	Sacos Neg.	Total Sacos	Instâncias
Musk1	166	47	45	92	476
Musk2	166	39	63	102	6598
Fox	230	100	100	200	1320
Tiger	230	100	100	200	1391
Elephant	230	100	100	200	1220

3.2.1 Musk

Este conjunto de dados (inserido no ramo da biologia) é usado para tentar prever se uma molécula tem um cheiro almiscarado (musky em inglês) ou não (Dietterich, Lathrop and Lozano-Pérez, 1997). Este é o conjunto de dados mais usado em testes de performance de algoritmos multi-padrão. Uma molécula é descrita pelas suas diferentes conformações. O Musk assume que se pelo menos uma das conformações for responsável pelo facto da molécula ter um cheiro almiscarado, a molécula é considerada positiva para a classe musky. Se nenhuma das conformações puder causar um cheiro almiscarado, então a molécula é considerada negativa.

Existem dois conjuntos de dados associados a este problema, o Musk1 e o Musk2. A diferença entre os dois é que o Musk1 é mais pequeno, tanto em termos de número de moléculas (sacos), como também em número de instâncias por molécula.

3.2.2 Fox, Tiger e Elephant

Inseridos no ramo da classificação de imagens, estes três conjuntos de dados foram introduzidos por (Andrews, Tsochantaridis and Hofmann, 2003) e são frequentemente usados como referências para a avaliação da performance de algoritmos multi-padrão. O objetivo é identificar se uma determinada imagem contém um determinado animal. Uma imagem é considerada positiva se pelo menos um dos seus segmentos contém o animal, caso contrário é considerada negativa.

4 Algoritmos de Agrupamento Multi-Grupo

Este capítulo descreve os algoritmos de agrupamento de dados multi-grupo que foram encontrados na literatura.

4.1 Agrupamento Multi-Grupo

4.1.1 LIFT

O LIFT (*multi-label learning with Label specific FeaTures*) é um algoritmo de aprendizagem de dados multi-grupo criado por (Zhang and Wu, 2015). Este isola as características mais influentes em cada grupo e utiliza-as no processo de classificação.

Inicialmente, para cada grupo, é feita uma análise das características mais pertinentes e descritivas do grupo tanto nas suas instâncias positivas como nas negativas. Seguidamente, são gerados vários classificadores, cada um associado às características específicas de um grupo em vez das suas características originais. A Tabela 6 apresenta uma descrição completa do algoritmo LIFT.

Tabela 6 – Pseudocódigo do algoritmo LIFT

-
- Dados de Entrada:
 - D – Conjunto de dados de treino multi-grupo $\{(x_i, Y_i) | 1 \leq i \leq m\}$

$(x_i \in X, Y_i \subseteq Y, X = \mathbb{R}^d, Y = \{l_1, l_2, \dots, l_q\})$

 - r – Rácio do número de grupos mantidos, $r \in [0,1]$
 - \mathcal{L} – Algoritmo de aprendizagem binário
 - u – novo objeto desconhecido
 - Dados de Saída:
 - Y – Grupo previsto para u
 - Procedimento:
 1. Para $k = 1 \rightarrow q$ fazer
 - a. Definir $P_k = \{x_i | (x_i, Y_i) \in D, l_k \in Y_i\}$;
 - b. Definir $N_k = \{x_i | (x_i, Y_i) \in D, l_k \notin Y_i\}$;
 - c. Executar agrupamento k-médias $m_k = \lceil r \cdot \min(|P_k|, |N_k|) \rceil$;
 - d. $\phi_k(x) = [d(x, p_1^k), \dots, d(x, p_{m_k}^k), d(x, n_1^k), \dots, d(x, n_{m_k}^k)]$;

Fim Para
 2. Para $k = 1 \rightarrow q$ fazer
 - a. $B_k = \{(\phi_k(x), Y_i(k)) | (x_i, Y_i) \in D\}$ em que $Y_i(k) = +1$ se $l_k \in Y_i$; senão $Y_i(k) = -1$;
 - b. Induzir g_k ao evocar \mathcal{L} , $g_k \leftarrow \mathcal{L}(B_k)$;

Fim Para
 3. Retornar $Y = \{l_k | g_k(\phi_k(u)) > 0, 1 \leq k \leq q\}$;
-

4.1.2 CBMLC

O CBMLC (*Clustering-Based Multi-Label Classification*) é um algoritmo de aprendizagem multi-grupo criado por (Nasierding, Tsoumakas and Kouzani, 2009). É constituído por um algoritmo de agrupamento de dados e um algoritmo de classificação multi-grupo.

O processo de treino do CBMLC consiste em dois passos:

1. Agrupar o conjunto de dados de treino em k grupos, sendo que k é dado pelo utilizador à partida. Nesta fase, os grupos presentes no conjunto de dados de treino não são considerados.
2. Usar um algoritmo multi-grupo nos dados de cada grupo e produzir k modelos de classificação multi-grupo.

Para classificar um novo objeto, o CBMLC encontra o grupo mais próximo desse objeto e depois usa o modelo multi-grupo para o classificar.

Como algoritmos de agrupamento de dados, o CBMLC usa o k -médias e o algoritmo EM (*Expectation Maximization*). Para a componente de classificação multi-grupo, o CBMLC usa os algoritmos BR, RAKEL, MLkNN e BPMLL.

4.1.3 CLMLC

O CLMLC (*Clustering-based Local Multi-Label Classification*) é um algoritmo de agrupamento de dados multi-grupo criado por (Sun, Kudo and Kimura, 2016).

Para o desenvolvimento deste algoritmo, os autores definiram duas suposições sobre a classificação multi-grupo, baseadas na observação de várias situações reais:

1. A existência de meta-grupos, ou seja, combinações de grupos que existem implicitamente no conjunto de grupos;
2. Apenas uma fração dos atributos são relevantes para o meta-grupo.

Assim, o algoritmo tenta decompor o conjunto de dados original em vários conjuntos de dados menores, cada um relevante apenas para alguns meta-grupos. Uma fração desses pequenos conjuntos de dados é usado como treino.

4.1.4 HOMER

O HOMER (Hierarchy Of Multilabel classifiERs) é um algoritmo de agrupamento de dados multi-grupo criado por (Tsoumakas, Katakis and Vlahavas, 2008). Este algoritmo utiliza uma hierarquia de dados em árvore e foi inicialmente desenvolvido especialmente para o tratamento de conjuntos de dados com um grande número de grupos.

Seguidamente são descritos os passos de processamento do algoritmo HOMER:

1. Organizar todos os grupos numa hierarquia em árvore, onde cada nóculo contém um conjunto de grupos que é substancialmente mais pequeno do que o conjunto total de grupos;
2. Um classificador multi-grupo é construído para cada nóculo da árvore que não seja uma folha;
3. A classificação multi-grupo é aplicada recursivamente, a começar a partir da raiz da árvore e avançando para os seus filhos apenas se os grupos contidos nesses filhos estiverem entre os previstos pelo classificador do nóculo pai.

Uma das principais características do algoritmo HOMER é que os seus autores desenvolveram um método (balanced k-means) para dividirem o conjunto de todos os grupos em subconjuntos, onde grupos similares são colocados no mesmo subconjunto e grupos dissimilares são colocados em subconjuntos diferentes.

4.2 Conjuntos de dados Multi-Grupo

Nesta secção estão descritos alguns dos conjuntos de dados mais utilizados em demonstrações de performance de algoritmos multi-grupo. A Tabela 7 apresenta um resumo das características de todos dos conjuntos de dados apresentados. Cardinalidade corresponde ao número médio de grupos por objeto. Densidade corresponde é igual à cardinalidade a dividir pelo número de grupos. Distintos corresponde ao número de combinações diferentes de grupos.

Tabela 7 – Descrição de conjuntos de dados multi-grupo

Dataset	Objetos	Atributos	Grupos	Cardinalidade	Densidade	Distintos	Ramo
Bibtex	7395	1836	159	2,402	0,015	2856	Texto
Birds	645	260	19	1,014	0,053	133	Áudio
Corel5k	5000	499	374	3,522	0,009	1453	Imagem
Emotions	593	1185	6	1,869	0,311	27	Áudio
Genbase	662	1186	27	1,252	0,046	32	Biologia
Medical	978	1449	45	1,245	0,028	94	Texto
Yeast	2417	103	14	4,237	0,303	198	Biologia

4.2.1 Bibtex

Inserido no ramo da classificação de texto, este conjunto de dados contém 7395 referências de publicações de artigos da Bibtex (Katakis, Tsoumakas and Vlahavas, 2008), classificadas pelas anotações dos seus utilizadores (p.e. datamining, estatística, quantum, etc.).

4.2.2 Birds

Inserido no ramo da classificação acústica, este conjunto de dados (Sun, Kudo and Kimura, 2016) contém 645 sons produzidos por pássaros, classificados pelas espécies que os produzem.

4.2.3 Corel5k

Inserido no ramo da classificação de imagens, o conjunto de dados Corel5k (Duygulu *et al.*, 2002) inclui imagens 5000 associadas a 4-5 palavras-chave que descrevem o conteúdo da imagem.

4.2.4 Emotions

Inserido no ramo da classificação acústica, neste conjunto de dados cada objeto corresponde a uma música, classificada pelo grupo a que pertence (p.e., clássica, metal, rock, etc.) (Trohidis *et al.*, 2011). O conjunto de dados Emotions é constituído por 593 músicas, cada uma classificada por um de seis grupos.

4.2.5 Genbase

Inserido no ramo da biomédica, neste conjunto de dados cada objeto corresponde a uma proteína, num total de 662. Estas são classificadas por 27 grupos que representam classes de proteínas (Diplaris *et al.*, 2005).

4.2.6 Medical

Inserido no ramo da classificação de texto, o conjunto de dados Medical contém 978 relatórios clínicos em texto, classificados por um ou mais de 45 códigos de doenças.

4.2.7 Yeast

Inserido no ramo da biomédica, neste conjunto de dados cada objeto corresponde a um gene de levedura (*yeast gene*) descrito pela concatenação de dados de microarranjos e perfil filogenético. Neste conjunto de dados introduzido em (Elisseff and Weston, 2001), cada um dos 2417 genes está associado a uma ou mais de 14 classes funcionais diferentes.

5 Avaliação dos Algoritmos

Este capítulo descreve os processos de desenho e avaliação aplicados nesta dissertação. Apresenta ainda as métricas de avaliação de algoritmos e as características do *hardware* e *software* utilizados nos testes de avaliação de performance dos algoritmos multi-grupo e multi-padrão.

5.1 Desenho da solução

Como já foi referido anteriormente no documento, o objetivo desta dissertação consiste no estudo de técnicas de agrupamento de dados associadas aos conceitos multi-grupo e multi-padrão. Sendo assim, a dissertação deve apresentar os algoritmos associados a estes conceitos, expondo as vantagens da sua utilização bem como as fraquezas dessas abordagens.

Apesar de, durante o estudo do estado de arte, ter sido encontrado o conceito de aprendizagem multi-grupo e multi-padrão em simultâneo, não foram encontrados algoritmos de agrupamento associados a este conceito. Por esse motivo, a aprendizagem multi-grupo e multi-padrão em simultâneo não será analisada nesta dissertação.

Esta dissertação analisará dois conceitos diferentes:

- Agrupamento de dados Multi-Padrão;
- Agrupamento de dados Multi-Grupo.

Estas técnicas não podem ser relacionadas entre si porque analisam problemas de agrupamento distintos, os seja, um conjunto de dados multi-padrão é fundamentalmente diferente de um conjunto de dados multi-grupo. Apesar disso, a solução apresentada para o estudo dos dois conceitos. Os conjuntos de dados e algoritmos analisados dependem do tipo de conceito a que estas pertencem.

Para se proceder ao estudo dos conceitos mencionados, seguiram-se os seguintes passos:

1. Fazer um levantamento dos algoritmos existentes. O resultado deste levantamento foi apresentado neste documento, nos capítulos de “Algoritmos de Agrupamento Multi-Grupo” e “Algoritmos de Agrupamento Multi-Padrão”.
2. Elaborar uma descrição formal dos algoritmos a implementar. Este passo é fulcral para um bom entendimento do funcionamento do algoritmo. Estas descrições estão também presentes nos capítulos de “Algoritmos de Agrupamento Multi-Grupo” e “Algoritmos de Agrupamento Multi-Padrão”.
3. Implementar (i.e., utilizar o código dos algoritmos disponibilizados pelos seus autores e adaptar sempre que necessário ao ambiente de desenvolvimento e ferramentas contemporâneas utilizadas) um conjunto de algoritmos associados ao tema numa ferramenta de processamento e cálculo numérico. A ferramenta a ser utilizada tem de ser comum a todos os algoritmos, de forma a evitar erros provocados por diferenças nos seus motores de processamento. Dos vários ambientes de desenvolvimento disponíveis, podiam ser escolhidos o R ou o Matlab, devido à sua relevância na área e à experiência do aluno.

Após a recolha dos algoritmos, verificou-se que a melhor ferramenta de processamento e cálculo numérico para executar tanto os algoritmos multi-grupo como os algoritmos multi-padrão seria o Matlab, isto porque os próprios algoritmos encontrados estavam já implementados em Matlab. No que diz respeito aos conjuntos de dados, também existiam mais fontes para Matlab, principalmente para conjuntos de dados multi-padrão.

Foi necessário adaptar algumas das implementações dos algoritmos, de forma a que estes pudessem tratar os mesmos conjuntos e para que pudessem ser executados com as ferramentas de cálculo e bibliotecas mais recentes. Esta adaptação foi longa e iterativa, tendo ocupado grande parte do tempo de implementação do projeto.

4. Avaliar os algoritmos implementados através da comparação da sua performance em diferentes conjuntos de dados. Uma descrição detalhada sobre o processo de avaliação e comparação dos algoritmos pode ser encontrada no capítulo seguinte. Os resultados são apresentados nos capítulos “Análise dos Algoritmos Multi-Grupo” e “Análise dos Algoritmos Multi-Padrão”.

Foi ainda ponderada a criação de um sistema para inclusão de algoritmos multi-grupo e multi-padrão, de forma a estes poderem ser facilmente configurados e avaliado o seu desempenho para diferentes conjuntos de dados. No entanto, esta ideia foi deixada de parte pois alteraria o foco da dissertação para algo completamente distinto do objetivo inicialmente idealizado, que consiste no estudo das técnicas de agrupamento de dados multi-grupo e multi-padrão e a comparação do desempenho das mesmas.

5.2 Avaliação da solução

O processo de avaliação do desempenho dos diferentes algoritmos deve ter em conta vários parâmetros, de forma a essa validação ser corretamente realizada.

O primeiro parâmetro a ter em conta são os conjuntos de dados utilizados. Para que os resultados dos algoritmos não sejam influenciados por uma eventual falta de qualidade dos conjuntos de dados, é importante utilizar vários conjuntos de dados diferentes e que estes sejam considerados fidedignos pela comunidade de aprendizagem. Estes conjuntos de dados podem ser encontrados *online* facilmente em vários repositórios públicos.

O paradigma que o algoritmo analisa também influencia o tipo de conjunto de dados utilizados. Por exemplo, um algoritmo de agrupamento de dados multi-padrão deve utilizar conjuntos de dados especializados para o problema multi-padrão.

O *hardware* e o *software* utilizados têm também influência no processo de avaliação. Como tal, deve ser sempre usado o mesmo computador e o mesmo motor de processamento matemático.

Os parâmetros de avaliação dos algoritmos foram obtidos através de várias execuções dos mesmos, com os recursos de *hardware*, *software* e conjuntos de dados acima descritos. O número de execuções foi de pelo menos 10, e maior (30+) nos casos que se considerou necessário para confirmar a normalidade dos resultados.

Na literatura, geralmente, a performance dos algoritmos de aprendizagem multi-padrão é avaliada segundo as métricas de exatidão, precisão, revocação e área abaixo da curva ROC. Por sua vez, a performance dos algoritmos de aprendizagem multi-grupo é frequentemente avaliada usando as métricas de *hamming loss*, f1-micro e f1-macro. Adicionalmente, também é importante, no âmbito desta dissertação, avaliar a performance dos algoritmos segundo os recursos que utilizam, como o tempo de execução e tempo de processador.

A comparação da performance dos algoritmos é o grande foco desta dissertação. Uma das hipóteses que se pretende verificar é se existe algum algoritmo que apresente melhor performance que os outros, em cada um dos parâmetros acima mencionados, para cada conjunto de dados. A mesma hipótese será também colocada para os resultados agregados de todos os conjuntos de dados, ou seja, tentar encontrar o melhor algoritmo para cada parâmetro de avaliação em todos os conjuntos de dados.

Afirmações como as que um determinado algoritmo é melhor do que os outros, ou até que não existem diferenças significativas na performance dos algoritmos, têm de ser suportadas por uma análise estatística. Em seguida, estão descritos quais os métodos estatísticos utilizados para comparar a performance dos algoritmos analisados:

- Espera-se que os resultados de N execuções de um algoritmo para um conjunto de dados tenda para uma distribuição normal. Para confirmar a normalidade dos resultados, será utilizado o teste de Shapiro-Wilk.

- Uma vez que os modelos de dados são criados com recurso ao método *k-fold cross validation*, e que não é garantido que os *folds* são iguais nas execuções dos diferentes algoritmos, os indicadores de performance dos algoritmos são variáveis independentes. Assim, para testar a hipótese se algum dos algoritmos é melhor para um determinado conjunto de dados, será realizado o teste ANOVA quando se comparam mais do que dois algoritmos, seguido do teste Tukey HSD se a hipótese nula é rejeitada. Nas situações em que só existem dois algoritmos a comparar, é usado o teste t de Student para amostras independentes.
- Para testar a hipótese de haver um algoritmo que apresente melhor performance do que os outros para a generalidade dos conjuntos de dados, para cada um dos parâmetros de performance, será utilizado o teste de Friedman, uma vez que os resultados dos diferentes algoritmos constituem variáveis emparelhadas (dependentes do conjunto de dados) e não normais, seguido do teste de Conover quando a hipótese nula é rejeitada. Nas situações onde só existem dois algoritmos a comparar, é usado o teste de Wilcoxon para amostras emparelhadas.

5.3 Métricas de Avaliação

A avaliação do desempenho de um determinado algoritmo pode seguir diferentes perspetivas. Pode atender ao tempo de execução de um algoritmo, ao tempo de CPU de um algoritmo ou uma medição do acerto (exatidão/precisão) entre a classificação ou agrupamento obtido pela aplicação do algoritmo a um conjunto de dados e a classificação ou o agrupamento de dados real. As métricas de avaliação do desempenho dos algoritmos variam de acordo com o tipo de algoritmo analisado e com o paradigma estudado.

Abaixo estão descritas as métricas de avaliação do desempenho dos algoritmos, mais utilizadas, para os paradigmas multi-grupo e multi-padrão. Nem todas as métricas apresentadas são utilizadas na avaliação do desempenho dos algoritmos desta dissertação, seja por não se adequarem ao paradigma estudado, ou por não se ter considerado relevante a sua implementação.

5.3.1 Tempo de Execução (tExe)

O tempo de execução de um algoritmo corresponde ao tempo, em segundos, que um algoritmo demora a gerar um resultado. Pode incluir o tempo de treino e o tempo de teste, ou apenas um deles.

Apesar de não avaliar a qualidade dos resultados produzidos por um algoritmo, esta é uma das características mais relevantes para aqueles que vão utilizar os algoritmos.

O tempo de execução será utilizado como métrica de avaliação de todos os algoritmos analisados nesta dissertação.

5.3.2 Tempo de CPU (tCPU)

O tempo de CPU de um algoritmo é uma medida do tempo, em segundos, que o algoritmo ocupa o processador do computador.

Esta medida tem em atenção que grande parte dos computadores modernos têm processadores com vários núcleos, capazes de executar partes do processamento a ser realizado em paralelo. Nestes casos, cada núcleo conta como se fosse um processador independente, ou seja, numa situação em que o algoritmo utilize 100% de um processador com quatro núcleos, o tempo de CPU deve ser quatro vezes superior ao tempo de execução.

Algoritmos de processamento mais antigos tendem a não utilizar as capacidades de paralelismo dos computadores modernos. Nos casos em que um algoritmo é executado num processador com vários núcleos e o tempo de execução é idêntico ao tempo de CPU, podemos quase sempre concluir que o algoritmo pode ser melhorado, de forma a tirar partido do paralelismo. No entanto, este tipo de melhorias não deve ter influência na qualidade dos resultados do acerto algoritmo.

O tempo de CPU será utilizado como métrica de avaliação de todos os algoritmos analisados nesta dissertação.

5.3.3 Matriz de confusão

A matriz de confusão ajuda a perceber algumas métricas de avaliação apresentadas.

A Tabela 8 apresenta a matriz de confusão, em que em que TP (*True Positive*) corresponde aos verdadeiros positivos, TN (*True Negative*) aos verdadeiros negativos, FP (*False Positive*) aos falsos positivos e FN (*False Negative*) aos falsos negativos (Sorower, 2010).

Tabela 8 – Matriz de confusão

Classe Real	Classe Prevista	
	Positivo	Negativo
Positivo	TP	FN
Negativo	FP	TN
	P=TP+FN	N=FP+TN

5.3.4 Exatidão (Accuracy, A)

Exatidão dos resultados de um algoritmo refere-se à relação entre o número de objetos bem classificados (TP + TN) e o número total de objetos do conjunto de dados testado (P+N=TP+FN+FP+TN) (Sorower, 2010). A fórmula de cálculo da exatidão é:

$$A = \frac{TP + TN}{P + N}$$

A exatidão será utilizada como métrica de avaliação dos algoritmos multi-padrão analisados nesta dissertação. Nos algoritmos multi-grupo, a exatidão não é utilizada por não ter em conta que um objeto pode pertencer a vários grupos, a atribuição de um objeto a um grupo pode ser apenas parcialmente correta (Sorower, 2010). A qualidade dos resultados do acerto dos algoritmos multi-grupo será avaliada por métricas com *exact match ratio* ou *hamming loss*.

5.3.5 Precisão (Precision, P)

Precisão dos resultados de um algoritmo refere-se à relação entre o número de objetos bem classificados como positivos (TP) e total dos objetos classificados como positivos (TP + FP) (Sorower, 2010). A fórmula de cálculo da precisão é:

$$P = \frac{TP}{TP + FP}$$

5.3.6 Revocação (Recall, R)

Revocação (também conhecida por sensibilidade) dos resultados de um algoritmo refere-se à relação entre o número de objetos classificados como positivos (TP) e o total real de objetos positivos ($P = TP + FN$) (Sorower, 2010). A fórmula de cálculo da revocação é:

$$R = \frac{TP}{P}$$

5.3.7 Área abaixo da curva ROC (AUC)

A área abaixo da curva ROC (Receiver Operator Characteristic) é um indicador da possibilidade de um algoritmo atribuir uma classificação positiva verdadeira em vez de atribuir uma classificação positiva falsa (Hanley and McNeil, 1982). A curva ROC é uma relação entre o rácio de verdadeiros positivos (recall) e o rácio de falsos positivos. Quanto maior a área abaixo da curva ROC, maior a probabilidade de uma classificação positiva atribuída pelo algoritmo ser verdadeira em vez de falsa.

Esta métrica é usada para a avaliação da performance dos algoritmos multi-padrão.

5.3.8 F-Measure (F)

F-measure (também conhecida por f1 score ou f-score) é um indicador da exatidão de um algoritmo. Corresponde à média harmónica da precisão e da revocação (Sorower, 2010). A fórmula de cálculo da F-measure é:

$$F = \frac{2}{\text{precisão}^{-1} + \text{revocação}^{-1}} = \frac{2TP}{2TP + FP + FN}$$

5.3.9 Exact Match Ratio (MR)

Exact match ratio é um indicador de exatidão para algoritmos multi-grupo. Esta métrica é uma extensão do método tradicional de exatidão que considera apenas como corretas as classificações que acertem em todos os grupos reais, declarando as classificações parcialmente corretas como incorretas (Sorower, 2010). Esta métrica é usada na avaliação da qualidade do acerto dos algoritmos multi-grupo.

5.3.10 Hamming Loss (HL)

Hamming Loss é um indicador de performance de algoritmo multi-grupo. Corresponde à relação entre os grupos mal atribuídos (sejam eles grupos que deviam ter sido atribuídos mas não foram ou grupos que foram atribuídos mas não deviam ter sido) e o número total de grupos (Sorower, 2010). Idealmente, o valor obtido de HL seria 0, o que indicaria que nenhum erro

tinha ocorrido. Esta métrica será usada na avaliação da qualidade do acerto dos algoritmos multi-grupo.

5.3.11 F1-macro

F1-macro é um indicador de performance de algoritmos multi-grupo que deriva do F-measure tradicional. Neste caso, os cálculos são efetuados para cada grupo individualmente e depois é feita a média de todos os grupos. Esta métrica é mais sensível aos grupos com menos exemplos e é usada na avaliação da qualidade do acerto dos algoritmos multi-grupo.

5.3.12 F1-micro

F1-micro é um indicador de performance de algoritmos multi-grupo que deriva do F-measure tradicional. Neste caso, os cálculos são efetuados globalmente para todos os objetos e grupos. Esta métrica é mais sensível aos grupos com mais exemplos e é usada na avaliação da qualidade do acerto dos algoritmos multi-grupo.

5.4 Recursos de Computação

Todas as análises dos algoritmos foram executadas no mesmo computador, com as seguintes características:

- CPU: Intel Core i5-6600K @ 4.4GHz, quatro núcleos;
- RAM: 16GB DDR4-2400;
- Disco: Samsung 850 EVO 512GB SSD;
- GPU: Nvidia GeForce GTX 1060 6GB;
- Sistema Operativo: Windows 10 Pro x64.

Adicionalmente, todos os algoritmos foram executados utilizando as seguintes ferramentas de *software*:

- Matlab R2017b x64
- XRealStats (Real Statistics Resource Pack for Excel) ¹
- Bibliotecas auxiliares:
 - PRTools (Pattern Recognition Toolbox), versão 5.1 ²
 - dd_tools (Data Description Toolbox), versão 2.1.2 ³
 - mil_tools (Multiple Instance Learning Toolbox), versão 1.2.1 ⁴
 - libsvm (Library for Support Vector Machines), versão 3.24 ⁵

¹ <http://www.real-statistics.com/free-download/real-statistics-resource-pack/>

² <http://prtools.tudelft.nl/>

³ <https://www.tudelft.nl/ewi/over-de-faculteit/afdelingen/intelligent-systems/pattern-recognition-bioinformatics/pattern-recognition-laboratory/data-and-software/dd-tools/>

⁴ <https://www.tudelft.nl/ewi/over-de-faculteit/afdelingen/intelligent-systems/pattern-recognition-bioinformatics/pattern-recognition-laboratory/data-and-software/mil-tools/>

⁵ <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>

6 Análise dos Algoritmos Multi-Padrão

Cada algoritmo foi testado utilizando o método *10-fold cross validation*, ou seja, com cada conjunto de dados dividido em 10 partes. Cada valor apresentado representa a média dos resultados obtidos para estes 10 *folds*. Por sua vez, este processo foi repetido 10 vezes, querendo isto dizer que cada algoritmo é foi executado 100 vezes.

Os algoritmos utilizados na análise dos algoritmos multi-padrão foram o MCIL e o BARTMIP, ambos com as configurações por omissão definidas pelos autores. O algoritmo MCIL foi executado com o número de grupos $K = 5$. O algoritmo BARTMIP também foi executado com o número de grupos $K = 5$, utilizando como medida de distância a média das distâncias mínimas de Hausdorff entre instâncias e tendo SVM como algoritmo de aprendizagem supervisionada no espaço da dissimilaridade.

Apenas foram analisados estes dois algoritmos porque, durante a fase de pesquisa, não foi encontrada a implementação de nenhum outro algoritmo, mesmo após tentar entrar em contacto com os autores dos algoritmos descritos no capítulo 3.1. Grande parte dos algoritmos descritos foram apresentados há mais de 10 anos, uma altura em que ainda não existia a cultura e facilidade de partilha de informação e recursos que existe hoje. Por outro lado, os autores podem já não estar contactáveis pelos emails institucionais presentes nos seus artigos.

Os conjuntos de dados utilizados para a realização deste teste foram o Musk1, Musk2, Fox, Tiger e Elephant. A performance dos algoritmos foi avaliada de acordo com o tempo de execução, tempo de CPU, exatidão e precisão.

A Tabela 9 apresenta as médias dos resultados gerados pelos algoritmos de agrupamento multi-padrão analisados. Os valores de tempo de execução e tempo de CPU são apresentados em segundos. Os restantes valores são apresentados no intervalo $[0, 1]$.

Tabela 9 – Resumo com médias dos resultados gerados pelos algoritmos multi-padrão

Dataset	MCIL				BARTMIP			
	tExe (s)	tCPU (s)	Exatidão	AUC	tExe (s)	tCPU (s)	Exatidão	AUC
Musk1	2,59	0,31	0,78	0,87	8,31	9,14	0,78	0,81
Musk2	30,51	2,91	0,79	0,87	11,77	35,84	0,82	0,84
Fox	9,31	0,71	0,59	0,60	40,91	40,99	0,56	0,55
Tiger	6,99	0,68	0,76	0,83	42,58	42,06	0,66	0,67
Elephant	8,87	1,02	0,77	0,88	38,88	40,06	0,52	0,52

O algoritmo BARTMIP não utiliza todos os recursos de processamento. Para o conjunto de dados Musk2, utilizou, no máximo, cerca de 80% do CPU. Nos restantes conjuntos de dados, utilizou no máximo cerca de 40% do CPU. Este facto tem influência direta no tempo de execução do algoritmo.

O algoritmo MCIL esteve sempre próximo dos 100% de utilização do CPU, uma vez que este implementa paralelismo para poder processar vários folds ao mesmo tempo. No entanto, devido a uma falha do Matlab, que não consegue contar o tempo de CPU utilizado por processos externos criados por ele próprio (p.e. executar ficheiros .exe por linha de comandos do sistema e esperar pelos resultados), o tempo de CPU do algoritmo MCIL não pode ser contabilizado corretamente. Por este motivo, o tempo de CPU não foi utilizado como métrica de avaliação dos testes de hipóteses 1 e 2.

As métricas obtidas de todas as execuções dos algoritmos podem ser consultadas no Anexo A, secção Resultados e Testes de Normalidade.

6.1 Testes de Normalidade

Tal como esperado, todas as métricas analisadas de todos os algoritmos tendem para a normalidade, confirmadas pelo teste Shapiro-Wilk (com nível de significância $\alpha = 0.05$). O tempo de CPU do algoritmo BARTMIP não foi tido em conta porque o Matlab não permite a sua contabilização de forma correta.

Todos os testes de Normalidade podem ser consultados no Anexo A, secção Resultados e Testes de Normalidade, incluindo o *p-value* de todos os testes Shapiro-Wilk.

6.2 Teste de Hipóteses 1

Pretende-se testar a hipótese de algum dos algoritmos ter melhor performance num determinado conjunto de dados, para cada métrica de avaliação.

A Tabela 10 apresenta os resultados do teste de hipóteses 1. Aplicou-se o teste T de Student por se tratar de análises a dois conjuntos de dados independentes (os conjuntos criados pela *cross validation* eram diferentes a cada execução) e normais. Para as situações em que a hipótese nula (i.e., ambos os algoritmos têm performances idênticas) é rejeitada ($p\text{-value} > \alpha$), pode-se concluir que um dos algoritmos tem melhor performance do que o outro, aparecendo nestas situações o nome do melhor algoritmo na tabela. As situações em que a hipótese nula é aceite, estão indicadas na tabela com H_0 . Todos os testes foram realizados com o nível de significância $\alpha = 0.05$, o que se traduz num intervalo de confiança de 95%.

Tabela 10 – Resultados do teste de hipóteses 1

	Musk1	Musk2	Fox	Tiger	Elephant
tExe	MCIL	BARTMIP	MCIL	MCIL	MCIL
Exatidão	H_0	H_0	H_0	MCIL	MCIL
AUC	MCIL	H_0	H_0	MCIL	MCIL

Desta forma pode-se concluir que o algoritmo MCIL apresenta melhor performance do que o algoritmo BARTMIP nos conjuntos de dados Tiger e Elephant, para todas as métricas de performance analisadas. O algoritmo MCIL apresentou também tempos de execução mais baixos para os conjuntos de dados Musk1 e Fox, bem como maior área abaixo da curva ROC (AUC) para o conjunto de dados Musk1.

O algoritmo BARTMIP apenas apresenta melhor performance do que o algoritmo MCIL especificamente no tempo de execução do algoritmo, para o conjunto de dados Musk2.

Para os restantes casos, segundo os resultados do teste T de Student, não há diferenças significativas na performance dos algoritmos, de forma a poder-se afirmar que um é melhor do que o outro.

Todos os resultados dos testes T de Student estão presentes com mais detalhe no Anexo A, secção Testes de Hipóteses 1.

6.3 Teste de Hipóteses 2

Pretende-se testar a hipótese de haver um algoritmo que apresente melhor performance do que os outros para a generalidade dos conjuntos de dados, para cada um dos parâmetros de performance.

A Tabela 11 apresenta os resultados do teste de hipóteses 2. Aplicou-se o teste de Wilcoxon por se tratar de análises a dois conjuntos de dados emparelhados (cada linha corresponde aos resultados para mesmo conjunto de dados) e não normais (resultados de execuções a conjuntos de dados diferentes não tendem para a normalidade). Para situações em que a hipótese nula (i.e., ambos os algoritmos têm performances idênticas) é rejeitada ($p\text{-value} > \alpha$), pode-se concluir que um dos algoritmos tem melhor performance do que o outro, aparecendo nestas situações o nome do melhor algoritmo indicado na tabela. As situações em que a hipótese nula é aceite estão indicadas na tabela com H_0 . Todos os testes foram realizados com o nível de significância $\alpha = 0.05$, o que se traduz num intervalo de confiança de 95%.

Tabela 11 – Resultados do teste de hipóteses 2

	Wilcoxon
tExe	MCIL
Exatidão	MCIL
AUC	H_0

Desta forma, pode-se concluir que o algoritmo MCIL tem melhor tempo de execução e exatidão do que o algoritmo BARTMIP, para a generalidade dos conjuntos de dados.

Em relação à área abaixo da curva ROC (AUC), segundo o teste de Wilcoxon, não há diferenças significativas de forma a poder-se afirmar que um dos algoritmos é melhor do que o outro.

Todos os resultados dos testes de Wilcoxon estão presentes com mais detalhe no Anexo A, secção Testes de Hipóteses 2.

7 Análise dos Algoritmos Multi-Grupo

Cada algoritmo foi testado utilizando o método *10-fold cross validation*, ou seja, com cada conjunto de dados dividido em 10 partes. Cada valor apresentado representa a média dos resultados obtidos para estes 10 *folds*. Por sua vez, este processo foi repetido 10 vezes, querendo isto dizer que cada algoritmo é foi executado 100 vezes. A única exceção ocorreu na execução do algoritmo LIFT para os conjuntos de dados Bibtex e Corel5k, devido ao tempo de execução verificado. Nestes dois casos, os resultados apresentados referem-se a um teste, utilizando o método *10-fold cross validation*, em vez de 10.

Os algoritmos utilizados na análise dos algoritmos multi-grupo foram o LIFT, o CLMLC, o CBMLC e o HOMER, com as configurações por omissão definidas pelos autores. O algoritmo LIFT foi executado com SVM do tipo RBF e rácio de *clusters* = 0,1. O algoritmo CLMLC foi executado com as configurações $d = 30$, $k = 100$ e $n = 5$. O algoritmo CBMLC foi executado com o número de clusters $K = 5$. O algoritmo HOMER também foi executado com o número de clusters $K = 5$, com o método de classificação base *Ridge Regression* e com o método de *threshold* SCut.

Os conjuntos de dados utilizados para a realização deste teste foram o Bibtex, Birds, Corel5k, Emotions, Genbase, Medical e Yeast. A performance dos algoritmos foi avaliada de acordo com o tempo de execução, tempo de CPU, *exact match ratio*, *hamming loss*, f1-macro e f1-micro.

As quatro tabelas seguintes apresentam as médias dos resultados dos algoritmos de agrupamento multi-grupo analisados. Os valores de tempo de execução e tempo de CPU são apresentados em segundos. Os restantes valores são apresentados no intervalo [0, 1].

Tabela 12 – Resumo com médias dos resultados gerados pelo algoritmo LIFT

	LIFT					
Dataset	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro
Bibtex	2331,04	5901,99	0,14	0,01	0,18	0,26
Birds	1,28	2,22	0,44	0,05	0	0
Corel5k	900,51	2473,21	0,01	0,01	0,03	0,10
Emotions	0,63	0,81	0,01	0,31	0,07	0,09
Genbase	2,96	7,87	0,95	0,003	0,58	0,97
Medical	12,24	34,02	0,49	0,01	0,24	0,68
Yeast	38,31	42,09	0,19	0,19	0,41	0,65

Tabela 13 – Resumo com médias dos resultados gerados pelo algoritmo CLMLC

	CLMLC					
Dataset	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro
Bibtex	6,05	20,24	0,17	0,01	0,21	0,40
Birds	0,03	0,12	0,52	0,05	0,33	0,48
Corel5k	2,25	8,08	0,03	0,012	0,04	0,17
Emotions	0,01	0,04	0,26	0,24	0,59	0,60
Genbase	0,10	0,33	0,97	0,00	0,65	0,98
Medical	0,33	1,06	0,70	0,01	0,33	0,79
Yeast	0,16	0,61	0,19	0,23	0,40	0,61

Tabela 14 – Resumo com médias dos resultados gerados pelo algoritmo CBMLC

	CBMLC					
Dataset	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro
Bibtex	118,39	387,25	0,15	0,02	0,24	0,37
Birds	0,12	0,45	0,50	0,06	0,30	0,44
Corel5k	36,35	110,95	0,01	0,01	0,02	0,17
Emotions	0,01	0,03	0,25	0,24	0,61	0,62
Genbase	2,14	6,57	0,98	0,0008	0,66	0,99
Medical	6,19	21,04	0,59	0,01	0,33	0,76
Yeast	0,06	0,21	0,19	0,22	0,38	0,62

Tabela 15 – Resumo com médias dos resultados gerados pelo algoritmo HOMER

Dataset	HOMER					
	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro
Bibtex	13,66	28,44	0,10	0,02	0,14	0,28
Birds	0,06	0,18	0,48	0,06	0,23	0,39
Corel5k	77,24	136,20	0,00	0,02	0,03	0,17
Emotions	0,02	0,05	0,28	0,21	0,66	0,68
Genbase	1,66	4,76	0,97	0,001	0,61	0,98
Medical	0,42	1,16	0,57	0,02	0,23	0,70
Yeast	0,07	0,24	0,15	0,24	0,42	0,65

O único algoritmo a utilizar 100% do CPU foi o CLMCL. O CBMLC e o HOMER utilizaram cerca de 80% e o LIFT flutuava entre 40 e 100%. Este comportamento dos algoritmos foi contante para todos os conjuntos de dados analisados.

Na Tabela 12, é possível verificar que, para o conjunto de dados Birds, os valores F1-macro e F1-micro do algoritmo LIFT são zero. Isto acontece porque o algoritmo não conseguiu atribuir classificações a nenhum objeto do conjunto de dados. Como não existem classificações positivas, os valores de F1-macro e F1-micro são zero.

É possível também verificar que o algoritmo LIFT não escala bem para conjuntos de dados maiores. Por este motivo, foi necessário, já na fase de testes, redefinir a lista do conjunto de dados a analisar, de forma se poder apresentar resultados do algoritmo LIFT para todos os conjuntos de dados.

7.1 Testes de Normalidade

Tal como esperado, todas as métricas analisadas de todos os algoritmos tendem para a normalidade, confirmadas pelo teste Shapiro-Wilk (com nível de significância $\alpha = 0.05$).

Todos os testes de Normalidade podem ser consultados no Anexo B, secção Resultados e Testes de Normalidade, incluindo o *p-value* de todos os testes Shapiro-Wilk.

7.2 Teste de Hipóteses 1

Pretende-se testar a hipótese de algum dos algoritmos ter melhor performance num determinado conjunto de dados, para cada métrica de avaliação.

A Tabela 16 apresenta os resultados do teste de hipóteses 1. Aplicou-se o teste ANOVA por se tratar de análises a quatro conjuntos de dados independentes (os conjuntos criados pela *cross validation* eram diferentes a cada execução) e normais. Como se pode verificar na tabela, a hipótese nula (i.e., todos os algoritmos têm performances idênticas) foi rejeitada para todas as variáveis e conjuntos de dados ($p\text{-value} > \alpha$). Pode-se então concluir que, para todas as variáveis e conjuntos analisados, existe pelo menos um algoritmo que é melhor ou pior do que outros algoritmos. Todos os testes foram realizados com o nível de significância $\alpha = 0.05$, o que se traduz num intervalo de confiança de 95%.

Tabela 16 – Resultados do teste de hipóteses 1 – Teste ANOVA

	Bibtex	Birds	Corel5k	Emotions	Genbase	Medical	Yeast
tExe (s)	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada
tCPU (s)	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada
EM	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada
HL	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada
F1-macro	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada
F1-micro	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada	Rejeitada

Sabendo os resultados do teste ANOVA, o passo seguinte é determinar quais os algoritmos que são piores ou melhores do que os outros, para cada métrica de performance e para cada conjunto de dados. Para esse efeito, aplicou-se o teste Tukey HSD com o nível de significância $\alpha = 0.05$, o que se traduz num intervalo de confiança de 95%. As tabelas seis tabelas seguintes apresentam as conclusões que se podem retirar dos testes Tukey HSD realizados para cada métrica e conjunto de dados.

Tabela 17 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – Tempo de Execução

	Tukey HSD - Tempo de Execução
Bibtex	LIFT tem pior performance do que todos os outros algoritmos. CBMLC é pior do que CLMLC e HOMER.
Birds	LIFT tem pior performance do que todos os outros algoritmos. CLMLC tem melhor performance do que todos os outros algoritmos. HOMER é melhor do que CBMLC.
Corel5k	LIFT tem pior performance do que todos os outros algoritmos. CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que HOMER.
Emotions	LIFT tem pior performance do que todos os outros algoritmos.
Genbase	LIFT tem pior performance do que todos os outros algoritmos. CLMLC tem melhor performance do que todos os outros algoritmos. HOMER é melhor do que CBMLC.
Medical	LIFT tem pior performance do que todos os outros algoritmos. CBMLC é pior do que CLMLC e HOMER.
Yeast	LIFT tem pior performance do que todos os outros algoritmos.

Tabela 18 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – Tempo de CPU

	Tukey HSD - Tempo de CPU
Bibtex	LIFT tem pior performance do que todos os outros algoritmos. CBMLC é pior do que CLMLC e HOMER.
Birds	LIFT tem pior performance do que todos os outros algoritmos. CLMLC tem melhor performance do que todos os outros algoritmos. HOMER é melhor do que CBMLC.
Corel5k	LIFT tem pior performance do que todos os outros algoritmos. CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que HOMER.
Emotions	LIFT tem pior performance do que todos os outros algoritmos.
Genbase	LIFT tem pior performance do que todos os outros algoritmos. CLMLC tem melhor performance do que todos os outros algoritmos. HOMER é melhor do que CBMLC.
Medical	LIFT tem pior performance do que todos os outros algoritmos. CBMLC é pior do que CLMLC e HOMER.
Yeast	LIFT tem pior performance do que todos os outros algoritmos.

Tabela 19 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – Exact Match Ratio

	Tukey HSD – Exact Match Ratio
Bibtex	HOMER tem pior performance do que todos os outros algoritmos. CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT.
Birds	LIFT tem pior performance do que todos os outros algoritmos. HOMER é pior do que CLMLC e CBMLC.
Corel5k	HOMER tem pior performance do que todos os outros algoritmos. CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT.
Emotions	LIFT tem pior performance do que todos os outros algoritmos. HOMER tem melhor performance do que todos os outros algoritmos.
Genbase	LIFT tem pior performance do que todos os outros algoritmos. CBMLC tem melhor performance do que todos os outros algoritmos.
Medical	LIFT tem pior performance do que todos os outros algoritmos. CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que HOMER.
Yeast	HOMER tem pior performance do que todos os outros algoritmos.

Tabela 20 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – Hamming Loss

	Tukey HSD – Hamming Loss
Bibtex	HOMER tem pior performance do que todos os outros algoritmos. CBMLC é pior do que CLMLC e LIFT.
Birds	CLMLC tem melhor performance do que todos os outros algoritmos. LIFT é melhor do que CBMLC e HOMER.
Corel5k	LIFT tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT e CLMLC. CLMLC é melhor do que HOMER.
Emotions	HOMER tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT e CLMLC. CLMLC é melhor do que LIFT.
Genbase	CBMLC tem melhor performance do que todos os outros algoritmos. HOMER é melhor do que LIFT e CLMLC. CLMLC é melhor do que LIFT.
Medical	CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT e HOMER. LIFT é melhor do que HOMER.
Yeast	LIFT tem melhor performance do que todos os outros algoritmos.

	CBMLC é melhor do que CLMLC e HOMER.
--	--------------------------------------

Tabela 21 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – F1-macro

	Tukey HSD – F1-macro
Bibtex	CBMLC tem melhor performance do que todos os outros algoritmos. CLMLC é melhor do que LIFT e HOMER. LIFT é melhor do que HOMER.
Birds	CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT e HOMER. HOMER é melhor do que LIFT.
Corel5k	CBMLC tem pior performance do que todos os outros algoritmos. CLMLC é melhor do que LIFT.
Emotions	LIFT tem melhor performance do que todos os outros algoritmos. HOMER é melhor do que CLMLC e CBMLC. CBMLC é melhor do que CLMLC.
Genbase	CBMLC tem melhor performance do que todos os outros algoritmos. CLMLC é melhor do que LIFT e HOMER. HOMER é melhor do que LIFT.
Medical	HOMER tem pior performance do que todos os outros algoritmos. LIFT é pior do que CLMLC e CBMLC.
Yeast	CBMLC tem pior performance do que todos os outros algoritmos. HOMER tem melhor performance do que todos os outros algoritmos.

Tabela 22 – Resultados do teste de hipóteses 1 – Teste Tukey HSD – F1-micro

	Tukey HSD – F1-micro
Bibtex	CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT e HOMER.
Birds	CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT e HOMER. HOMER é melhor do que LIFT.
Corel5k	LIFT tem pior performance do que todos os outros algoritmos.
Emotions	HOMER tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT e CLMLC. CLMLC é melhor do que LIFT.
Genbase	CBMLC tem melhor performance do que todos os outros algoritmos. HOMER é melhor do que LIFT e CLMLC. CLMLC é melhor do que LIFT.
Medical	CLMLC tem melhor performance do que todos os outros algoritmos. CBMLC é melhor do que LIFT e HOMER. HOMER é melhor do que LIFT.
Yeast	CLMLC tem pior performance do que todos os outros algoritmos. CBMLC é pior do que HOMER e LIFT.

Analisando os resultados do teste Tukey HSD, é possível concluir que o algoritmo LIFT sempre os piores tempos de Execução e tempos de CPU, e tem sempre dos piores ou o pior resultados no que diz respeito à métrica de avaliação de performance F1-micro.

Todos os resultados dos testes ANOVA e dos testes Tukey HSD estão presentes com mais detalhe no Anexo B, secção Testes de Hipóteses 1.

7.3 Teste de Hipóteses 2

Pretende-se testar a hipótese de haver um algoritmo que apresente melhor performance do que os outros para a generalidade dos conjuntos de dados, para cada um dos parâmetros de performance.

A Tabela 23 apresenta os resultados do teste de hipóteses 2. Aplicou-se o teste de Friedman pro se tratar de análises a quatro conjuntos de dados emparelhados (cada linha corresponde aos resultados do mesmo conjunto de dados) e não normais (resultados de execuções a conjuntos de dados diferentes não tendem para a normalidade). Para situações em que a hipótese nula (i.e., todos os algoritmos têm performances idênticas) é rejeitada ($p\text{-value} < \alpha$), pode-se concluir que pelo menos um dos algoritmos tem melhor ou pior performance do que os outros. As situações em que a hipótese nula é aceite estão indicadas na tabela com H_0 . Todos os testes foram realizados com o nível de significância $\alpha = 0.05$, o que se traduz num intervalo de confiança de 95%

Tabela 23 – Resultados do teste de hipóteses 2 – teste de Friedman

	Friedman
Tempo de Execução	Rejeitada
Tempo de CPU	Rejeitada
Exact Match	Rejeitada
Hamming Loss	H_0
F1-macro	H_0
F1-micro	H_0

Sabendo os resultados do teste de Friedman, o passo seguinte é determinar quais os algoritmos que são piores ou melhores do que os outros, para cada métrica de performance. Para esse efeito, aplicou-se o teste Conover com o nível de significância $\alpha = 0.05$, o que se traduz num intervalo de confiança de 95% (inicialmente estava previsto utilizar o teste de Nemenyi, mas os resultados que este apresentava não condiziam com os resultados do teste de Friedman). A Tabela 24 apresentam as conclusões que se podem retirar dos testes Tukey HSD realizados para cada métrica e conjunto de dados.

Tabela 24 – Resultados do teste de hipóteses 2 – teste de Conover

	Conover
tExe	LIFT tem pior performance do que todos os outros algoritmos.
tCPU	LIFT tem pior performance do que todos os outros algoritmos.

Exact Match	CLMLC é melhor do que LIFT e HOMER.
	CBMLC é melhor do que LIFT.

Tal como já se tinha verificado no teste de Hipóteses 1, pode-se afirmar que o algoritmo LIFT tem pior tempo de execução e tempo de CPU do que os restantes algoritmos analisados.

No que diz respeito a indicadores do acerto, o destaque vai para os algoritmos CLMLC e CBMLC, que, provavelmente, deverão atingir melhores classificações de grupos completos do que os restantes algoritmos.

Em relação às métricas *Hamming Loss*, F1-macro e F1-micro, segundo o teste de Friedman, não há diferenças significativas de forma a poder-se afirmar que algum dos algoritmos é melhor do que um ou mais dos outros.

Todos os resultados dos testes de Friedman e dos testes de Conover estão presentes com mais detalhe no Anexo B, secção Testes de Hipóteses 2.

8 Conclusão

Esta dissertação focou-se no estudo e comparação de técnicas de agrupamento de dados relacionadas com os conceitos de aprendizagem multi-grupo e multi-padrão, de forma a permitir ao leitor fazer uma escolha sustentada sobre qual dos algoritmos conhecidos será o melhor para o seu problema, dentro destes dois paradigmas.

Estas técnicas de agrupamento multi-grupo e multi-padrão podem ser aplicadas em várias áreas de estudo, incluindo biologia, classificação de imagens, áudio e vídeo, classificação de texto, entre outros. O agrupamento de dados consiste na divisão de um conjunto de dados em grupos sem uso de informação *a priori*, em que dados semelhantes são colocados no mesmo grupo e dados dissemelhantes são colocados em grupos diferentes.

Em qualquer processo de aprendizagem automática, é importante saber quais os algoritmos que produzem melhores resultados na análise de um determinado conjunto de dados. Os algoritmos de agrupamento de dados são avaliados, principalmente, pela qualidade do acerto do agrupamento de dados produzido quando comparados com a partição real dos dados, pelo seu tempo de execução e pelos recursos utilizados (p.e., tempo de CPU). O conhecimento de quais os algoritmos de agrupamento de dados obtêm melhores resultados de performance, permite a um utilizador a melhor escolha do algoritmo quando necessitar de realizar um agrupamento de dados.

Nesta dissertação foram apresentados alguns algoritmos de agrupamento multi-grupo e multi-padrão e foi analisada a performance desses algoritmos, com o uso das métricas referidas no parágrafo anterior.

Relativamente aos algoritmos multi-padrão, apenas foram analisados os algoritmos MCIL e BARTMIP porque, durante a fase de pesquisa, não foi encontrada a implementação de nenhum outro algoritmo, mesmo após tentar entrar em contacto com os autores dos algoritmos descritos no capítulo 3.1. Grande parte dos algoritmos descritos foram apresentados há mais

de 10 anos, uma altura em que ainda não existia a cultura e facilidade de partilha de informação e recursos que existe hoje. Por outro lado, os autores podem já não estar contactáveis pelos emails institucionais presentes nos seus artigos.

Os resultados da avaliação da performance dos algoritmos multi-padrão MCIL e BARTMIP permitem afirmar que o algoritmo MCIL apresenta melhores resultados do que o algoritmo BARTMIP, relativamente ao tempo de execução e exatidão, para a generalidade dos conjuntos de dados.

Relativamente à métrica AUC (área abaixo da curva ROC), os resultados da avaliação realizada, não permitem afirmar que um dos algoritmos é o melhor para a generalidade dos conjuntos de dados (teste de hipóteses 2). Isto apesar de, tendo em conta os resultados do teste de hipóteses 1, o algoritmo MCIL apresentar melhores resultados de AUC do que o algoritmo BARTMIP para três dos cinco conjuntos de dados analisados.

Os resultados da avaliação de performance aos algoritmos multi-grupo LIFT, CLMCL, CBMLC e HOMER permitem afirmar que o algoritmo LIFT tem significativamente piores tempos de execução e piores tempos de CPU do que os restantes algoritmos, para a generalidade dos conjuntos de dados.

Relativamente à métrica *exact match ratio*, pode-se concluir que o algoritmo CLMCL apresenta vantagens significativas quando comparado com os algoritmos LIFT e HOMER, para a generalidade dos conjuntos de dados. Da mesma forma, também é possível afirmar que o algoritmo CBMLC apresenta melhor *exact match ratio* do que o algoritmo LIFT.

Para as métricas *hamming loss*, f1-macro e f1-micro, os resultados da avaliação da performance não permitem afirmar que um dos algoritmos é melhor do que os outros para a generalidade dos conjuntos de dados. No entanto, analisando os resultados de cada conjunto de dados com métricas de avaliação individualmente, pode-se concluir que o algoritmo LIFT tem sempre o pior ou dos piores resultados para a métrica de avaliação f1-micro.

Durante o desenvolvimento da dissertação, foi necessário adaptar algumas das implementações dos algoritmos disponibilizadas pelos seus autores, de forma a que os algoritmos pudessem tratar os mesmos conjuntos de dados e para que pudessem ser executados com as ferramentas de cálculo e bibliotecas mais recentes. Esta adaptação foi longa e iterativa, tendo ocupado grande parte do tempo de implementação do projeto.

Existe espaço para a melhoria da qualidade dos resultados gerados pelo estudo levado a cabo nesta dissertação. Por exemplo, aumentando o número de conjuntos de dados analisados, poder-se-á permitir tirar conclusões relativas a qual o melhor algoritmo para tratar dados de uma determinada área (p.e. biologia, texto, imagens, etc.). Por outro lado, quanto mais forem os conjuntos de dados analisados, mais diversos os seus tamanhos (tanto em número de instâncias como no número de atributos) e mais áreas de estudo, mais força ganham as conclusões que indicam uma melhor ou pior qualidade de um determinado algoritmo para a generalidade dos conjuntos de dados.

Do mesmo modo, aumentando o número de algoritmos analisados, poder-se-á também aumentar a qualidade dos resultados produzidos pelo estudo desenvolvido nesta dissertação, especialmente na análise de algoritmos multi-padrão.

Em jeito de conclusão, pode-se dizer que este trabalho apresenta um estudo relativo a um conjunto de diferentes algoritmos dos paradigmas multi-padrão e multi-grupo que serão certamente úteis a quem necessitar de realizar agrupamentos de dados deste tipo.

Considero que este trabalho foi bastante interessante e que produziu bons resultados.

Referências

- Andrews, S., Tsochantaridis, I. and Hofmann, T. (2003) 'Support Vector Machines for Multiple-Instance Learning', *Advances in Neural Information Processing Systems*, (15), pp. 561–568.
- Auer, P. and Ortner, R. (2004) 'A Boosting Approach to Multiple Instance Learning', in *Machine Learning: ECML*, pp. 63–74. doi: 10.1007/978-3-540-30115-8_9.
- Babenko, B. (2008) *Multiple instance learning: algorithms and applications*, View Article PubMed/NCBI Google Scholar. University of California, San Diego. Available at: http://vision.ucsd.edu/~bbabenko/data/bbabenko_re.pdf%5Cnpapers3://publication/uuid/2CDB4FD4-9E25-4F12-826C-E67049137B7C.
- Boutell, M. R. *et al.* (2004) 'Learning multi-label scene classification', *Pattern Recognition*, 37(9), pp. 1757–1771. doi: 10.1016/j.patcog.2004.03.009.
- Clare, A. and King, R. D. (2001) 'Knowledge Discovery in Multi-label Phenotype Data', in *Pkdd*, pp. 42–53. doi: 10.1007/3-540-44794-6_4.
- Cleuziou, G. (2007) *A generalization of k-means for overlapping clustering*, Université d'Orléans, LIFORapport No RR-2007-15. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.106.1534&rep=rep1&type=pdf>.
- Dempster, A. P. A., Laird, N. M. N. and Rubin, D. D. B. (1977) 'Maximum likelihood from incomplete data via the EM algorithm', *Journal of the Royal Statistical Society Series B Methodological*, 39(1), pp. 1–38. doi: <http://dx.doi.org/10.2307/2984875>.
- Dietterich, T. G., Lathrop, R. H. and Lozano-Pérez, T. (1997) 'Solving the multiple instance problem with axis-parallel rectangles', *Artificial Intelligence*, 89(1–2), pp. 31–71. doi: 10.1016/S0004-3702(96)00034-3.
- Diplaris, S. *et al.* (2005) 'Protein Classification with Multiple Algorithms', in. Springer, Berlin, Heidelberg, pp. 448–456. doi: 10.1007/11573036_42.
- Duarte, F. J. F. (2008) *Optimização da Combinação de Agrupamentos baseado na Acumulação de Provas pesadas por Índices de Validação e com uso de Amostragem*. Universidade de Trás-os-Montes e Alto Douro.
- Duygulu, P. *et al.* (2002) 'Object Recognition as Machine Translation: Learning a Lexicon for a Fixed Image Vocabulary', in. Springer, Berlin, Heidelberg, pp. 97–112. doi: 10.1007/3-540-47979-1_7.
- Elisseeff, a. and Weston, J. (2001) 'A kernel method for multi-labelled classification', in *Advances in neural information processing systems*, pp. 681–687.
- Ester, M. *et al.* (1996) 'A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise', in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*. AAAI Press, pp. 226–231.

- Feng, S. and Xu, D. (2010) 'Transductive Multi-Instance Multi-Label learning algorithm with application to automatic image annotation', *Expert Systems with Applications*. Elsevier Ltd, 37(1), pp. 661–670. doi: 10.1016/j.eswa.2009.06.111.
- Fürnkranz, J. *et al.* (2008) 'Multilabel classification via calibrated label ranking', *Machine Learning*, 73(2), pp. 133–153. doi: 10.1007/s10994-008-5064-8.
- Ghamrawi, N. and McCallum, A. (2005) 'Collective multi-label classification', in *Proceedings of the 14th ACM international conference on Information and knowledge management - CIKM '05*. New York, New York, USA: ACM Press, p. 195. doi: 10.1145/1099554.1099591.
- Guha, S., Rastogi, R. and Shim, K. (1998) 'CURE', in *Proceedings of the 1998 ACM SIGMOD international conference on Management of data - SIGMOD '98*. New York, New York, USA: ACM Press, pp. 73–84. doi: 10.1145/276304.276312.
- Hand, D. J. (2008) 'Data Clustering: Theory, Algorithms, and Applications by Guojun Gan, Chaoqun Ma, Jianhong Wu', *International Statistical Review*, 76(1), pp. 141–141. doi: 10.1111/j.1751-5823.2007.00039_2.x.
- Hanley, J. A. and McNeil, B. J. (1982) 'The meaning and use of the area under a receiver operating characteristic (ROC) curve', *Radiology*, 143(1), pp. 29–36. doi: 10.1148/radiology.143.1.7063747.
- Henegar, C., Clément, K. and Zucker, J. (2006) 'Unsupervised Multiple-Instance Learning for Functional Profiling of Genomic Data', in *Learning*, pp. 186–197. doi: 10.1007/11871842_21.
- Herrera, F. *et al.* (2016) *Multiple Instance Learning*. Cham: Springer International Publishing. doi: 10.1007/978-3-319-47759-6.
- Irani, J., Pise, N. and Phatak, M. (2016) 'Clustering Techniques and the Similarity Measures used in Clustering: A Survey', *International Journal of Computer Applications*, 134(7), pp. 975–8887. Available at: <http://www.ijcaonline.org/research/volume134/number7/irani-2016-ijca-907841.pdf> (Accessed: 16 February 2018).
- Jain, a. K., Murty, M. N. and Flynn, P. J. (1999) 'Data clustering: a review', *ACM Computing Surveys*, 31(3), pp. 264–323. doi: 10.1145/331499.331504.
- Jain, A. K. (2010) 'Data clustering: 50 years beyond K-means', *Pattern Recognition Letters*. Elsevier B.V., 31(8), pp. 651–666. doi: 10.1016/j.patrec.2009.09.011.
- Katakis, I., Tsoumakas, G. and Vlahavas, I. P. (2008) 'Multilabel Text Classification for Automated Tag Suggestion'. Available at: <https://www.semanticscholar.org/paper/Multilabel-Text-Classification-for-Automated-Tag-Katakis-Tsoumakas/157099d6ffd3ffca8cfa7955aff7c5f1a979ac9> (Accessed: 11 October 2019).
- Kaufman, L. and Rousseeuw, P. J. (1987) 'Clustering by means of medoids', *Statistical Data Analysis Based on the L1-Norm and Related Methods*, (1), p. North-Holland.
- Kohavi, R. and Kohavi, R. (1995) 'A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection', pp. 1137--1143. Available at: <https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.48.529> (Accessed: 11 October 2019).

- Kriegel, H. *et al.* (2006) 'COSMIC: Conceptually Specified Multi-Instance Clusters', in *Sixth International Conference on Data Mining (ICDM'06)*. IEEE, pp. 917–921. doi: 10.1109/ICDM.2006.46.
- Kriegel, H. H. H. and Schubert, M. (2004) 'Classification of Websites as Sets of Feature Vectors.', in *Databases and applications*, pp. 127–132. Available at: <http://www.dbs.ifi.lmu.de/~schubert/papers/dba2004.pdf>.
- Kriegel, H. P., Pryakhin, A. and Schubert, M. (2006) 'An EM-approach for clustering multi-instance objects', in *Advances in Knowledge Discovery and Data Mining, Proceedings*, pp. 139–148.
- Liao, B. *et al.* (2014) 'Using Multi-Instance Hierarchical Clustering Learning System to Predict Yeast Gene Function', *PLoS ONE*. Edited by H. Fröhlich, 9(3), p. e90962. doi: 10.1371/journal.pone.0090962.
- MacQueen, J. B. (1967) 'Some Methods for classification and Analysis of Multivariate Observations', *5th Berkeley Symposium on Mathematical Statistics and Probability*, 1(233), pp. 281–297. doi: citeulike-article-id:6083430.
- Maron, O. and Lozano-Pérez, T. (1997) 'A Framework for Multiple-Instance Learning', *NIPS '97 Proceedings of the 1997 conference on Advances in neural information processing systems 10*, pp. 570–576. doi: ISBN:0-262-10076-2.
- McCallum, A. (1999) 'Multi-label text classification with a mixture model trained by EM', in *AAAI'99 Workshop on Text Learning*, pp. 1–7. doi: 10.1.1.35.888.
- Megano, T. *et al.* (2015) 'Evolutionary multi-objective distance metric learning for multi-label clustering', in *2015 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, pp. 2945–2952. doi: 10.1109/CEC.2015.7257255.
- Nasierding, G., Tsoumakas, G. and Kouzani, A. Z. (2009) 'Clustering based multi-label classification for image annotation and retrieval', in *2009 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, pp. 4514–4519. doi: 10.1109/ICSMC.2009.5346902.
- Read, J. *et al.* (2009) 'Classifier Chains for Multi-label Classification', in *Machine Learning and Knowledge Discovery in Databases*, pp. 254–269. doi: 10.1007/978-3-642-04174-7_17.
- Rong Jin, Shijun Wang and Zhi-Hua Zhou (2009) 'Learning a distance metric from multi-instance multi-label data', in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 896–902. doi: 10.1109/CVPRW.2009.5206684.
- Scott, S., Zhang, J. U. N. and Brown, J. (2005) 'On generalized multiple-instance learning', *International Journal of Computational Intelligence and Applications*, 5(1), pp. 21–35.
- Sorower, M. (2010) *A literature survey on algorithms for multi-label learning.*, Oregon State University, Corvallis. Available at: <http://people.oregonstate.edu/~sorowerm/pdf/Qual-Multilabel-Shahed-CompleteVersion.pdf>.
- Sun, L., Kudo, M. and Kimura, K. (2016) 'A Scalable Clustering-Based Local Multi-Label Classification Method', 0. doi: 10.3233/978-1-61499-672-9-261.
- Tang, L., Rajan, S. and Narayanan, V. K. (2009) 'Large scale multi-label classification via

metalabeler', in *Proceedings of the 18th international conference on World wide web - WWW '09*. New York, New York, USA: ACM Press, p. 211. doi: 10.1145/1526709.1526738.

Trohidis, K. *et al.* (2011) 'Multi-label classification of music by emotion', *EURASIP Journal on Audio, Speech, and Music Processing*. Springer International Publishing, 2011(1), p. 4. doi: 10.1186/1687-4722-2011-426793.

Tsoumakas, G. and Katakis, I. (2009) 'Multi-Label Classification', in *Database Technologies*. IGI Global, pp. 309–319. doi: 10.4018/978-1-60566-058-5.ch021.

Tsoumakas, G., Katakis, I. and Vlahavas, I. (2008) 'Effective and Efficient Multilabel Classification in Domains with Large Number of Labels', in *Proc. ECML/PKDD 2008 Workshop on Mining Multidimensional Data*. Available at: <http://lps.csd.auth.gr/publications/tsoumakas-mmd08.pdf> (Accessed: 3 October 2019).

Tsoumakas, G., Katakis, I. and Vlahavas, I. (2009) 'Mining Multi-label Data', in *Data Mining and Knowledge Discovery Handbook*. Boston, MA: Springer US, pp. 667–685. doi: 10.1007/978-0-387-09823-4_34.

Tsoumakas, G. and Vlahavas, I. (2007) 'Random k-Labelsets: An Ensemble Method for Multilabel Classification', in *Machine Learning: ECML 2007*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 406–417. doi: 10.1007/978-3-540-74958-5_38.

Wang, J. and Zucker, J.-D. (2000) 'Solving Multiple-Instance Problem: A Lazy Learning Approach', in *Proc. 17th International Conf. on Machine Learning*, (1994), pp. 1119--1125. Available at: <http://cogprints.org/2124/>.

Wang, W., Yang, J. and Muntz, R. (1997) 'STING : A Statistical Information Grid Approach to Spatial Data Mining', in *Proceedings of the 23rd VLDB conference*, pp. 1–10.

Weidmann, N., Frank, E. and Pfahringer, B. (2003) 'A Two-Level Learning Method for Generalized Multi-instance Problems', in *Machine Learning: ECML 2003*, pp. 468–479. doi: 10.1007/978-3-540-39857-8_42.

Xu, X.-S., Xue, X. and Zhou, Z.-H. (2011) 'Ensemble multi-instance multi-label learning approach for video annotation task', in *Proceedings of the 19th ACM international conference on Multimedia - MM '11*. New York, New York, USA: ACM Press, p. 1153. doi: 10.1145/2072298.2071962.

Xu, Y. *et al.* (2012) 'Multiple clustered instance learning for histopathology cancer image classification, segmentation and clustering', in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 964–971. doi: 10.1109/CVPR.2012.6247772.

Ying-Xin Li *et al.* (2012) 'Drosophila Gene Expression Pattern Annotation through Multi-Instance Multi-Label Learning', *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 9(1), pp. 98–112. doi: 10.1109/TCBB.2011.73.

Zhang, D. *et al.* (2009) 'M3IC: Maximum Margin Multiple Instance Clustering', in *Proceedings of the 21st International Joint Conference on Artificial Intelligence*. Morgan Kaufmann Publishers Inc., pp. 1339–1344. Available at: <http://dl.acm.org/citation.cfm?id=1661445.166166>.

- Zhang, M.-L. and Wu, L. (2015) 'LIFT: Multi-Label Learning with Label-Specific Features', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(1), pp. 107–120. doi: 10.1109/TPAMI.2014.2339815.
- Zhang, M.-L. and Zhou, Z.-H. (2007) 'ML-KNN: A lazy learning approach to multi-label learning', *Pattern Recognition*, 40(7), pp. 2038–2048. doi: 10.1016/j.patcog.2006.12.019.
- Zhang, M.-L. and Zhou, Z.-H. (2008) 'M3MIML: A Maximum Margin Method for Multi-instance Multi-label Learning', in *2008 Eighth IEEE International Conference on Data Mining*. IEEE, pp. 688–697. doi: 10.1109/ICDM.2008.27.
- Zhang, M.-L. and Zhou, Z.-H. (2009) 'Multi-instance clustering with applications to multi-instance prediction', *Applied Intelligence*, 31(1), pp. 47–68. doi: 10.1007/s10489-007-0111-x.
- Zhang, M.-L. and Zhou, Z.-H. (2014) 'A Review on Multi-Label Learning Algorithms', *IEEE Transactions on Knowledge and Data Engineering*, 26(8), pp. 1819–1837. doi: 10.1109/TKDE.2013.39.
- Zhou, Z.-H. *et al.* (2012) 'Multi-instance multi-label learning', *Artificial Intelligence*, 176(1), pp. 2291–2320. doi: 10.1016/j.artint.2011.10.002.
- Zhou, Z.-H., Jiang, K. and Li, M. (2005) 'Multi-Instance Learning Based Web Mining', *Applied Intelligence*, 22(2), pp. 135–147. doi: 10.1007/s10489-005-5602-z.
- Zhou, Z.-H. and Zhang, M. (2006) 'Multi-instance multi-label learning with application to scene classification', *Advances in Neural Information Processing Systems*, pp. 1609–1617.

A Anexo A - Processamento dos Algoritmos Multi-Padrão

A.1 Resultados e Testes de Normalidade

A.1.1 MCIL

Tabela 25 – Resultados e teste Shapiro-Wilk para o conjunto de dados Musk1

Dataset	Musk1				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	2,681933	0,2234375	0,8369565	0,9219858		
2	2,4796665	0,2140625	0,7826087	0,8683215		
3	2,5496223	0,4796875	0,8369565	0,8853428		
4	2,5270574	0,390625	0,7934783	0,8955083		
5	2,6923694	0,3984375	0,7826087	0,8836879		
6	2,7770785	0,25	0,75	0,8399527		
7	2,4897877	0,2171875	0,7173913	0,8104019		
8	2,7940601	0,39375	0,7282609	0,8212766		
9	2,4249617	0,2359375	0,7608696	0,8825059		
10	2,5075887	0,2515625	0,826087	0,9146572		
Average	2,5924125	0,3054688	0,7815217	0,8723641		
p-value	0,1943876	0,022726	0,4843947	0,4639481		
Shapiro-Wilk	yes	no	yes	yes	alpha	0,05

Tabela 26 – Resultados e teste Shapiro-Wilk para o conjunto de dados Musk2

Dataset	Musk2				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	33,495837	2,89375	0,7777778	0,8617521		
2	29,695175	2,9796875	0,8383838	0,9431624		
3	32,210023	2,8984375	0,7575758	0,8489316		
4	31,332039	2,965625	0,8181818	0,891453		
5	29,769948	3,0015625	0,7878788	0,8762821		
6	30,061774	2,8359375	0,7979798	0,908547		
7	30,574505	2,9	0,7777778	0,8683761		
8	29,397389	2,834375	0,7777778	0,8358974		
9	29,20379	2,8671875	0,7171717	0,8008547		
10	29,346987	2,8828125	0,8181818	0,8788462		
Average	30,508747	2,9059375	0,7868687	0,8714103		
p-value	0,0623947	0,2934018	0,7107743	0,995816		
Shapiro-Wilk	yes	yes	yes	yes	alpha	0,05

Tabela 27 – Resultados e teste Shapiro-Wilk para o conjunto de dados Fox

Dataset	Fox				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	9,0878302	0,7359375	0,6	0,59265		
2	9,0556387	0,7203125	0,59	0,5664		
3	9,4087444	0,6953125	0,57	0,58735		
4	9,1777687	0,6984375	0,585	0,6137		
5	9,2790801	0,7	0,585	0,57055		
6	9,2434111	0,696875	0,555	0,5671		
7	9,855967	0,7015625	0,6	0,6003		
8	9,552287	0,703125	0,61	0,6391		
9	9,1979738	0,70625	0,58	0,5584		
10	9,2100213	0,7	0,63	0,66455		
Average	9,3068722	0,7057813	0,5905	0,59601		
p-value	0,0775719	0,0040573	0,9824099	0,2538494		
Shapiro-Wilk	yes	no	yes	yes	alpha	0,05

Tabela 28 – Resultados e teste Shapiro-Wilk para o conjunto de dados Tiger

Dataset	Tiger				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	7,7777325	0,75	0,725	0,81275		
2	6,0790881	0,6484375	0,745	0,82615		
3	6,3430894	0,640625	0,745	0,80475		
4	6,346369	0,65	0,77	0,85435		
5	5,9854305	0,85	0,755	0,8197		
6	7,8160854	0,640625	0,745	0,834		
7	6,5269343	0,6453125	0,755	0,83665		
8	8,2632968	0,6640625	0,77	0,84385		
9	8,0446846	0,65	0,775	0,8391		
10	6,6900376	0,6390625	0,785	0,87085		
Average	6,9872748	0,6778125	0,757	0,834215		
p-value	0,0703467	9,184E-05	0,72932	0,9836623		
Shapiro-Wilk	yes	no	yes	yes	alpha	0,05

Tabela 29 – Resultados e teste Shapiro-Wilk para o conjunto de dados Elephant

Dataset	Elephant				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	9,2579395	0,825	0,81	0,8933		
2	8,5925198	1,0203125	0,75	0,8576		
3	8,0728057	1,1546875	0,775	0,8706		
4	9,1320888	1,125	0,795	0,91325		
5	9,3461074	1,3734375	0,765	0,9035		
6	9,3066301	1,2921875	0,785	0,8946		
7	9,4524916	0,771875	0,75	0,8649		
8	8,5038843	0,7875	0,78	0,89435		
9	8,8416842	1,0796875	0,77	0,8663		
10	8,180238	0,746875	0,74	0,85445		
Average	8,8686389	1,0176563	0,772	0,881285		
p-value	0,2340115	0,3138455	0,9459955	0,2836222		
Shapiro-Wilk	yes	yes	yes	yes	alpha	0,05

A.1.2 BARTMIP

Tabela 30 – Resultados e teste Shapiro-Wilk para o conjunto de dados Musk1

Dataset	Musk1				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	8,3568569	9,25	0,86	0,863		
2	8,3159105	9,1984375	0,7475	0,813		
3	8,235757	9,0078125	0,7875	0,817		
4	8,2503875	9,0546875	0,885	0,8915		
5	8,2783517	9,05625	0,745	0,7915		
6	8,37527	9,2609375	0,9	0,956		
7	8,3251696	9,221875	0,81	0,825		
8	8,3328577	9,1640625	0,6775	0,706		
9	8,3090551	9,0140625	0,8375	0,846		
10	8,2997966	9,1609375	0,55	0,62425		
Average	8,3079413	9,1389063	0,78	0,813325		
p-value	0,9318481	0,1826462	0,340692	0,5887022		
Shapiro-Wilk	yes	yes	yes	yes	alpha	0,05

Tabela 31 – Resultados e teste Shapiro-Wilk para o conjunto de dados Musk2

Dataset	Musk2				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	11,810326	35,546875	0,8244048	0,8095238		
2	11,774712	35,93125	0,8017857	0,881746		
3	11,76372	35,871875	0,852381	0,8666667		
4	11,760015	35,740625	0,7940476	0,777381		
5	11,784024	35,910938	0,9452381	0,9464286		
6	11,76342	35,934375	0,7446429	0,7529762		
7	11,773738	35,992188	0,9232143	0,9535714		
8	11,756291	35,842188	0,8553571	0,8714286		
9	11,781236	35,923438	0,7202381	0,7666667		
10	11,770379	35,696875	0,7107143	0,7470238		
Average	11,773786	35,839063	0,8172024	0,8373413		
p-value	0,1416388	0,1242802	0,6782925	0,2239384		
Shapiro-Wilk	yes	yes	yes	yes	alpha	0,05

Tabela 32 – Resultados e teste Shapiro-Wilk para o conjunto de dados Fox

Dataset	Fox				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	40,351532	40,539063	0,4154924	0,395		
2	41,232403	41,298438	0,37	0,39		
3	40,868741	41,060938	0,5903939	0,589		
4	40,831994	40,9625	0,4449015	0,433		
5	41,131208	41,157813	0,6894091	0,684		
6	40,868145	40,989063	0,689803	0,675		
7	41,109973	41,026563	0,5350985	0,513		
8	41,111126	41,1375	0,6647045	0,638		
9	40,57797	40,79375	0,7547045	0,735		
10	41,032947	40,914063	0,4745076	0,471		
Average	40,911604	40,987969	0,5629015	0,5523		
p-value	0,2338981	0,6925192	0,518347	0,3367454		
Shapiro-Wilk	yes	yes	yes	yes	alpha	0,05

Tabela 33 – Resultados e teste Shapiro-Wilk para o conjunto de dados Tiger

Dataset	Tiger				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	40,15024	40,239063	0,5267541	0,49		
2	42,596171	42,38125	0,4167541	0,413		
3	46,069849	43,271875	0,79	0,801		
4	44,066069	42,660938	0,7356721	0,774		
5	44,268003	43,70625	0,791082	0,813		
6	41,755437	41,66875	0,6287869	0,638		
7	43,060413	42,732813	0,5712131	0,579		
8	42,196108	42,225	0,804459	0,802		
9	40,875184	40,954688	0,794459	0,835		
10	40,714803	40,753125	0,5345902	0,531		
Average	42,575228	42,059375	0,659377	0,6676		
p-value	0,8128082	0,7940412	0,1074657	0,1182819		
Shapiro-Wilk	yes	yes	yes	yes	alpha	0,05

Tabela 34 – Resultados e teste Shapiro-Wilk para o conjunto de dados Elephant

Dataset	Elephant				nFold	10
nClass	5					
Run	tExe (s)	tCPU (s)	Accuracy	AUC		
1	38,870712	40,167188	0,6080877	0,622		
2	38,686849	39,84375	0,6245219	0,625		
3	38,804915	39,959375	0,57	0,581		
4	38,945879	40,15625	0,6154781	0,609		
5	38,910251	40,11875	0,4916535	0,504		
6	38,877561	40,010938	0,4161754	0,419		
7	38,917908	40,010938	0,5226096	0,505		
8	38,985673	40,128125	0,5116535	0,525		
9	38,920111	40,1375	0,4078289	0,404		
10	38,837089	40,04375	0,4209561	0,43		
Average	38,875695	40,057656	0,5188965	0,5224		
p-value	0,341851	0,2106043	0,1962049	0,2116276		
Shapiro-Wilk	yes	yes	yes	yes	alpha	0,05

A.2 Teste de Hipóteses 1

Tabela 35 – Teste T de Student para o tempo de execução - conjunto de dados Musk1

Parameter	tExe	
Dataset	Musk1	
Run	MCIL	BARTMIP
1	2,68193296	8,35685694
2	2,47966649	8,31591051
3	2,54962227	8,23575704
4	2,52705739	8,25038747
5	2,69236936	8,27835174
6	2,77707849	8,37526999
7	2,4897877	8,32516961
8	2,79406006	8,33285774
9	2,42496166	8,3090551
10	2,50758868	8,29979662
Average	2,59241251	8,30794128

T Test: Two Independent Samples

SUMMARY				Hyp Mean Diff	0
Groups	Count	Mean	Variance	Cohen d	
MCIL	10	2,5924125	0,0174832		
BARTMIP	10	8,30794128	0,00193035		
Pooled			0,0097068	58,0120721	

T TEST: Equal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,0440609	129,71894	18	1,685E-28	1,7340636			yes	0,9994656
Two Tail	0,04406087	129,718937	18	3,3705E-28	2,10092204	-5,8080972	-5,6229603	yes	0,99946557

T TEST: Unequal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,04406087	129,718937	10,963472	0	1,79643252			yes	0,99967439
Two Tail	0,04406087	129,718937	10,963472	0	2,20188019	-5,8125455	-5,618512	yes	0,99967439

Tabela 36 – Teste T de Student para o tempo de execução - conjunto de dados Musk2

Parameter	tExe	
Dataset	Musk2	
Run	MCIL	BARTMIP
1	33,495837	11,8103264
2	29,6951752	11,7747116
3	32,2100227	11,7637199
4	31,3320394	11,7600154
5	29,7699475	11,7840245
6	30,061774	11,7634202
7	30,5745052	11,7737382
8	29,3973895	11,7562906
9	29,2037904	11,7812365
10	29,3469874	11,7703793
Average	30,5087468	11,7737862

T Test: Two Independent Samples

SUMMARY		Hyp Mean Diff	0
<i>Groups</i>	<i>Count</i>	<i>Mean</i>	<i>Variance</i>
MCIL	10	30,508747	2,0216001
BARTMIP	10	11,7737862	0,00024583
Pooled		1,01092299	18,6334704

T TEST: Equal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,4496494	41,665706	18	1,176E-19	1,7340636			yes	0,9948557
Two Tail	0,44964942	41,6657064	18	2,3518E-19	2,10092204	17,7902822	19,679639	yes	0,99485573

T TEST: Unequal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,44964942	41,6657064	9,00218886	6,5549E-12	1,83306219			yes	0,99741729
Two Tail	0,44964942	41,6657064	9,00218886	1,311E-11	2,26207332	17,7178206	19,7521005	yes	0,99741729

Tabela 37 – Teste T de Student para o tempo de execução - conjunto de dados Fox

Parameter	tExe	
Dataset	Fox	
Run	MCIL	BARTMIP
1	9,08783018	40,351532
2	9,05563871	41,2324035
3	9,4087444	40,8687408
4	9,1777687	40,8319941
5	9,2790801	41,1312082
6	9,24341114	40,8681448
7	9,85596698	41,1099727
8	9,55228695	41,1111263
9	9,19797377	40,5779697
10	9,21002126	41,0329471
Average	9,30687222	40,9116039

T Test: Two Independent Samples

SUMMARY	Hyp Mean Diff	0
<i>Groups</i>	<i>Count</i>	<i>Mean</i>
<i>Variance</i>	<i>Cohen d</i>	
MCIL	10	9,3068722
BARTMIP	10	40,9116039
Pooled		0,06699963
		122,100068

T TEST: Equal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,115758	273,02405	18	2,583E-34	1,7340636			yes	0,9998793
Two Tail	0,11575805	273,024052	18	5,1667E-34	2,10092204	-31,84793	-31,361533	yes	0,99987928

T TEST: Unequal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,11575805	273,024052	17,7105989	0	1,73560001			yes	0,99988123
Two Tail	0,11575805	273,024052	17,7105989	0	2,10338552	-31,848216	-31,361248	yes	0,99988123

Tabela 38 – Teste T de Student para o tempo de execução - conjunto de dados Tiger

Parameter	tExe	
Dataset	Tiger	
Run	MCIL	BARTMIP
1	7,77773246	40,1502401
2	6,07908813	42,596171
3	6,34308944	46,0698489
4	6,34636901	44,0660693
5	5,98543053	44,2680026
6	7,81608536	41,7554369
7	6,52693426	43,0604132
8	8,26329679	42,196108
9	8,04468458	40,8751842
10	6,69003762	40,714803
Average	6,98727482	42,5752277

T Test: Two Independent Samples

SUMMARY	Hyp Mean Diff	0
<i>Groups</i>	<i>Count</i>	<i>Mean</i>
<i>MCIL</i>	10	6,9872748
<i>BARTMIP</i>	10	42,5752277
<i>Pooled</i>		2,08869941
		24,6243647

T TEST: Equal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,646328	55,061753	18	8,08E-22	1,7340636			yes	0,9970446
Two Tail	0,646328	55,0617534	18	1,616E-21	2,10092204	-36,945838	-34,230068	yes	0,99704462

T TEST: Unequal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,646328	55,0617534	12,9210669	5,5511E-17	1,77176106			yes	0,99787586
Two Tail	0,646328	55,0617534	12,9210669	1,1102E-16	2,16171098	-36,985127	-34,190779	yes	0,99787586

Tabela 39 – Teste T de Student para o tempo de execução - conjunto de dados Elephant

Parameter	tExe	
Dataset	Elephant	
Run	MCIL	BARTMIP
1	9,2579395	38,8707121
2	8,59251977	38,6868487
3	8,07280566	38,804915
4	9,13208877	38,945879
5	9,34610736	38,9102508
6	9,30663008	38,8775615
7	9,45249162	38,9179079
8	8,50388434	38,9856732
9	8,84168422	38,9201106
10	8,18023795	38,8370891
Average	8,86863893	38,8756948

T Test: Two Independent Samples

SUMMARY		Hyp Mean Diff	0
<i>Groups</i>	<i>Count</i>	<i>Mean</i>	<i>Variance</i>
MCIL	10	8,8686389	0,2554353
BARTMIP	10	38,8756948	0,00712903
Pooled		0,13128217	82,81719402

T TEST: Equal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,1620384	185,18488	18	2,79158E-31	1,7340636			yes	0,9997377
Two Tail	0,16203837	185,184876	18	5,58316E-31	2,10092204	-30,347486	-29,666626	yes	0,99973766

T TEST: Unequal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,16203837	185,184876	9,50197689	0	1,82214862			yes	0,99986149
Two Tail	0,16203837	185,184876	9,50197689	0	2,24407046	-30,370681	-29,64343	yes	0,99986149

Tabela 40 – Teste T de Student para a exatidão - conjunto de dados Musk1

Parameter	Accuracy	
Dataset	Musk1	
Run	MCIL	BARTMIP
1	0,83695652	0,86
2	0,7826087	0,7475
3	0,83695652	0,7875
4	0,79347826	0,885
5	0,7826087	0,745
6	0,75	0,9
7	0,7173913	0,81
8	0,72826087	0,6775
9	0,76086957	0,8375
10	0,82608696	0,55
Average	0,78152174	0,78

T Test: Two Independent Samples

SUMMARY			Hyp Mean Diff	0					
Groups	Count	Mean	Variance	Cohen d					
MCIL	10	0,7815217	0,0018497						
BARTMIP	10	0,78	0,01130833						
Pooled			0,006579	0,01876119					
T TEST: Equal Variances				Alpha	0,05				
	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,036274	0,0419513	18	0,4834997	1,7340636			no	0,0098875
Two Tail	0,03627395	0,04195129	18	0,96699938	2,10092204	-0,074687	0,07773049	no	0,00988753

T TEST: Unequal Variances				Alpha	0,05				
	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,03627395	0,04195129	11,8674796	0,48361754	1,78394727			no	0,01217682
Two Tail	0,03627395	0,04195129	11,8674796	0,96723508	2,18151433	-0,0776104	0,08065389	no	0,01217682

Tabela 41 – Teste T de Student para a exatidão - conjunto de dados Musk2

Parameter	Accuracy	
Dataset	Musk2	
Run	MCIL	BARTMIP
1	0,77777778	0,82440476
2	0,83838384	0,80178571
3	0,75757576	0,85238095
4	0,81818182	0,79404762
5	0,78787879	0,9452381
6	0,7979798	0,74464286
7	0,77777778	0,92321429
8	0,77777778	0,85535714
9	0,71717172	0,7202381
10	0,81818182	0,71071429
Average	0,78686869	0,81720238

T Test: Two Independent Samples

SUMMARY				Hyp Mean Diff	0
Groups	Count	Mean	Variance	Cohen d	
MCIL	10	0,7868687	0,0011892		
BARTMIP	10	0,81720238	0,00635082		
Pooled			0,00377002	0,49403025	

T TEST: Equal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,0274591	1,1046852	18	0,1419222	1,7340636			no	0,2519754
Two Tail	0,02745913	1,10468523	18	0,28384441	2,10092204	-0,0880232	0,02735579	no	0,25197539

T TEST: Unequal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,02745913	1,10468523	12,2564032	0,14525546	1,77918605			no	0,30091667
Two Tail	0,02745913	1,10468523	12,2564032	0,29051093	2,17376824	-0,0900235	0,02935609	no	0,30091667

Tabela 42 – Teste T de Student para a exatidão - conjunto de dados Fox

Parameter	Accuracy	
Dataset	Fox	
Run	MCIL	BARTMIP
1	0,6	0,41549242
2	0,59	0,37
3	0,57	0,59039394
4	0,585	0,44490152
5	0,585	0,68940909
6	0,555	0,68980303
7	0,6	0,53509848
8	0,61	0,66470455
9	0,58	0,75470455
10	0,63	0,47450758
Average	0,5905	0,56290152

T Test: Two Independent Samples

SUMMARY				Hyp Mean Diff	0
Groups	Count	Mean	Variance	Cohen d	
MCIL	10	0,5905	0,0004414		
BARTMIP	10	0,56290152	0,0179408		
Pooled			0,00919109	0,28787347	

T TEST: Equal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,0428745	0,6437047	18	0,2639421	1,7340636			no	0,1500059
Two Tail	0,04287445	0,64370465	18	0,52788417	2,10092204	-0,0624774	0,11767437	no	0,15000591

T TEST: Unequal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,04287445	0,64370465	9,44257751	0,26754359	1,82337913			no	0,20502921
Two Tail	0,04287445	0,64370465	9,44257751	0,53508717	2,24609738	-0,0687017	0,12389868	no	0,20502921

Tabela 43 – Teste T de Student para a exatidão - conjunto de dados Tiger

Parameter	Accuracy	
Dataset	Tiger	
Run	MCIL	BARTMIP
1	0,725	0,5267541
2	0,745	0,4167541
3	0,745	0,79
4	0,77	0,73567213
5	0,755	0,79108197
6	0,745	0,62878689
7	0,755	0,57121311
8	0,77	0,80445902
9	0,775	0,79445902
10	0,785	0,53459016
Average	0,757	0,65937705

T Test: Two Independent Samples

SUMMARY				Hyp Mean Diff	0
Groups	Count	Mean	Variance	Cohen d	
MCIL	10	0,757	0,0003233		
BARTMIP	10	0,65937705	0,02002895		
Pooled			0,01017614	0,96774373	

T TEST: Equal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,0451135	2,1639408	18	0,0220809	1,7340636			yes	0,4543584
Two Tail	0,0451135	2,16394077	18	0,04416187	2,10092204	0,002843	0,19240291	yes	0,45435841

T TEST: Unequal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,0451135	2,16394077	9,29050369	0,02888227	1,82660849			yes	0,57889302
Two Tail	0,0451135	2,16394077	9,29050369	0,05776453	2,25142039	-0,0039465	0,19919241	no	0,57889302

Tabela 44 – Teste T de Student para a exatidão - conjunto de dados Elephant

Parameter	Accuracy	
Dataset	Elephant	
Run	MCIL	BARTMIP
1	0,81	0,60808771
2	0,75	0,62452193
3	0,775	0,57
4	0,795	0,61547807
5	0,765	0,49165349
6	0,785	0,41617541
7	0,75	0,52260963
8	0,78	0,51165349
9	0,77	0,4078289
10	0,74	0,42095615
Average	0,772	0,51889648

T Test: Two Independent Samples

SUMMARY			Hyp Mean Diff	0
Groups	Count	Mean	Variance	Cohen d
MCIL	10	0,772	0,0004733	
BARTMIP	10	0,51889648	0,00714886	
Pooled			0,0038111	4,099899264

T TEST: Equal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,0276083	9,1676535	18	1,67161E-08	1,7340636			yes	0,9075289
Two Tail	0,02760832	9,16765346	18	3,34321E-08	2,10092204	0,19510059	0,31110645	yes	0,90752892

T TEST: Unequal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,02760832	9,16765346	10,1865964	1,53414E-06	1,80909805			yes	0,9444042
Two Tail	0,02760832	9,16765346	10,1865964	3,06827E-06	2,22261886	0,19174075	0,3144663	yes	0,9444042

Tabela 45 – Teste T de Student para AUC - conjunto de dados Musk1

Parameter	AUC	
Dataset	Musk1	
Run	MCIL	BARTMIP
1	0,92198582	0,863
2	0,86832151	0,813
3	0,88534279	0,817
4	0,89550827	0,8915
5	0,88368794	0,7915
6	0,83995272	0,956
7	0,81040189	0,825
8	0,8212766	0,706
9	0,88250591	0,846
10	0,91465721	0,62425
Average	0,87236407	0,813325

T Test: Two Independent Samples

SUMMARY				Hyp Mean Diff	0
Groups	Count	Mean	Variance	Cohen d	
MCIL	10	0,8723641	0,0014114		
BARTMIP	10	0,813325	0,00865522		
Pooled			0,00503329	0,8321724	

T TEST: Equal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,0317279	1,8607941	18	0,0395948	1,7340636			yes	0,4016591
Two Tail	0,03172789	1,86079405	18	0,07918959	2,10092204	-0,0076188	0,12569689	no	0,40165909

T TEST: Unequal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,03172789	1,86079405	11,8591491	0,0438686	1,78405295			yes	0,47538424
Two Tail	0,03172789	1,86079405	11,8591491	0,0877372	2,18168637	-0,0101812	0,12825937	no	0,47538424

Tabela 46 – Teste T de Student para AUC - conjunto de dados Musk2

Parameter	AUC	
Dataset	Musk2	
Run	MCIL	BARTMIP
1	0,86175214	0,80952381
2	0,94316239	0,88174603
3	0,84893162	0,86666667
4	0,89145299	0,77738095
5	0,87628205	0,94642857
6	0,90854701	0,75297619
7	0,86837607	0,95357143
8	0,83589744	0,87142857
9	0,8008547	0,76666667
10	0,87884615	0,74702381
Average	0,87141026	0,83734127

T Test: Two Independent Samples

SUMMARY		Hyp Mean	0
		Diff	
<i>Groups</i>	<i>Count</i>	<i>Mean</i>	<i>Variance</i>
MCIL	10	0,8714103	0,0015395
BARTMIP	10	0,83734127	0,00600465
Pooled		0,00377206	0,55471484

T TEST: Equal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,0274666	1,2403801	18	0,1153803	1,7340636			no	0,2806136
Two Tail	0,02746657	1,24038008	18	0,23076054	2,10092204	-0,0236361	0,09177411	no	0,28061359

T TEST: Unequal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,02746657	1,24038008	13,3302165	0,11811204	1,76758465			no	0,32167508
Two Tail	0,02746657	1,24038008	13,3302165	0,23622407	2,15494111	-0,0251199	0,09325783	no	0,32167508

Tabela 47 – Teste T de Student para AUC - conjunto de dados Fox

Parameter	AUC	
Dataset	Fox	
Run	MCIL	BARTMIP
1	0,59265	0,395
2	0,5664	0,39
3	0,58735	0,589
4	0,6137	0,433
5	0,57055	0,684
6	0,5671	0,675
7	0,6003	0,513
8	0,6391	0,638
9	0,5584	0,735
10	0,66455	0,471
Average	0,59601	0,5523

T Test: Two Independent Samples

SUMMARY				Hyp Mean Diff	0
Groups	Count	Mean	Variance	Cohen d	
MCIL	10	0,59601	0,0011942		
BARTMIP	10	0,5523	0,01643801		
Pooled			0,00881608	0,465525	

T TEST: Equal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,0419907	1,0409455	18	0,1558408	1,7340636			no	0,2382858
Two Tail	0,04199067	1,04094554	18	0,31168156	2,10092204	-0,0445091	0,13192912	no	0,23828583

T TEST: Unequal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,04199067	1,04094554	10,3007594	0,16085835	1,80710606			no	0,30851361
Two Tail	0,04199067	1,04094554	10,3007594	0,3217167	2,21935191	-0,0494821	0,13690207	no	0,30851361

Tabela 48 – Teste T de Student para AUC - conjunto de dados Tiger

Parameter	AUC	
Dataset	Tiger	
Run	MCIL	BARTMIP
1	0,81275	0,49
2	0,82615	0,413
3	0,80475	0,801
4	0,85435	0,774
5	0,8197	0,813
6	0,834	0,638
7	0,83665	0,579
8	0,84385	0,802
9	0,8391	0,835
10	0,87085	0,531
Average	0,834215	0,6676

T Test: Two Independent Samples

SUMMARY				Hyp Mean Diff	0
Groups	Count	Mean	Variance	Cohen d	
MCIL	10	0,834215	0,0003861		
BARTMIP	10	0,6676	0,02445471		
Pooled			0,0124204	1,49501735	

T TEST: Equal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,0498406	3,3429604	18	0,0018104	1,7340636			yes	0,6189035
Two Tail	0,04984055	3,34296042	18	0,00362085	2,10092204	0,06190388	0,27132612	yes	0,6189035

T TEST: Unequal Variances Alpha 0,05

	<i>std err</i>	<i>t-stat</i>	<i>df</i>	<i>p-value</i>	<i>t-crit</i>	<i>lower</i>	<i>upper</i>	<i>sig</i>	<i>effect r</i>
One Tail	0,04984055	3,34296042	9,28411779	0,00412709	1,82674665			yes	0,73906682
Two Tail	0,04984055	3,34296042	9,28411779	0,00825419	2,25164823	0,0543916	0,2788384	yes	0,73906682

Tabela 49 – Teste T de Student para AUC - conjunto de dados Elephant

Parameter	AUC	
Dataset	Elephant	
Run	MCIL	BARTMIP
1	0,8933	0,622
2	0,8576	0,625
3	0,8706	0,581
4	0,91325	0,609
5	0,9035	0,504
6	0,8946	0,419
7	0,8649	0,505
8	0,89435	0,525
9	0,8663	0,404
10	0,85445	0,43
Average	0,881285	0,5224

T Test: Two Independent Samples

SUMMARY				Hyp Mean Diff	0
Groups	Count	Mean	Variance	Cohen d	
MCIL	10	0,881285	0,0004329		
BARTMIP	10	0,5224	0,00725293		
Pooled			0,00384291	5,789292148	

T TEST: Equal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,0277233	12,945251	18	7,37525E-11	1,7340636			yes	0,9502666
Two Tail	0,02772329	12,9452508	18	1,47505E-10	2,10092204	0,30064052	0,41712948	yes	0,95026656

T TEST: Unequal Variances Alpha 0,05

	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,02772329	12,9452508	10,0704819	6,64737E-08	1,81117478			yes	0,97124291
Two Tail	0,02772329	12,9452508	10,0704819	1,32947E-07	2,22602686	0,2971722	0,4205978	yes	0,97124291

A.3 Teste de Hipóteses 2

Parameter	tExe	
	Algorithm	
Dataset	MCIL (s)	BARTMIP (s)
Musk1	2,592412506	8,307941276
Musk2	30,50874683	11,77378624
Fox	9,306872219	40,91160392
Tiger	6,987274818	42,57522772
Elephant	8,868638927	38,87569478

Tabela 50 – Teste de Wilcoxon para tempo de execução

Wilcoxon Signed-Rank Test for Paired Samples

	MCIL (s)	BARTMIP (s)
median	8,8686389	38,875695

count	5
# unequal	5
T+	2
T-	13
T	2

	one tail	two tail
mean	7,5	
std dev	3,7080992	
z-score	1,4832397	
effect r	0,4690416	
p-norm	0,0690054	0,1380107
p-exact	0,09375	0,1875
p-simul	N/A	N/A
alpha	0,05	
sig (norm)	yes	

Tabela 51 – Teste de Wilcoxon para a exatidão

Parameter	Accuracy	
	Algorithm	
Dataset	MCIL	BARTMIP
Musk1	0,7815217	0,78
Musk2	0,7868687	0,8172024
Fox	0,5905	0,5629015
Tiger	0,757	0,659377
Elephant	0,772	0,5188965

Wilcoxon Signed-Rank Test for Paired Samples

	MCIL	BARTMIP
median	0,772	0,659377

count	5
# unequal	5
T+	3
T-	12
T	3

	one tail	two tail
mean	7,5	
std dev	3,7080992	
z-score	1,2135598	
effect r	0,3837613	
p-norm	0,1124579	0,2249159
p-exact	0,15625	0,3125
p-simul	N/A	N/A
alpha	0,05	
sig (norm)	yes	

Tabela 52 – Teste de Wilcoxon para AUC

Parameter	AUC	
	Algorithm	
Dataset	MCIL	BARTMIP
Musk1	0,8723641	0,813325
Musk2	0,8714103	0,8373413
Fox	0,59601	0,5523
Tiger	0,834215	0,6676
Elephant	0,881285	0,5224

Wilcoxon Signed-Rank Test for Paired Samples

	MCIL	BARTMIP
median	0,8714103	0,6676

count	5
# unequal	5
T+	0
T-	15
T	0

	one tail	two tail
mean	7,5	
std dev	3,7080992	
z-score	2,0225996	
effect r	0,6396021	
p-norm	0,0215572	0,0431144
p-exact	0,03125	0,0625
p-simul	N/A	N/A
alpha	0,05	
sig (norm)	no	

B Anexo B - Processamento dos Algoritmos Multi-Grupo

B.1 Resultados e Testes de Normalidade

B.1.1 LIFT

Tabela 53 – Resultados e teste Shapiro-Wilk para o conjunto de dados Bibtex

Dataset	Bibtex						nFold
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro	1
1	2332,7178	5881,4063	0,1378378	0,0137855	0,179502	0,2765388	
2	2298,6894	5848,2656	0,1407307	0,0142212	0,1779358	0,2596367	
3	2375,3737	5959,3125	0,1299053	0,0149446	0,2036997	0,2683333	
4	2353,3421	5963,6719	0,1163735	0,0140254	0,1909025	0,2688554	
5	2253,8177	5844,9063	0,1123139	0,014885	0,1634171	0,2385721	
6	2280,1517	5868,2188	0,1675676	0,0131056	0,1920145	0,3128342	
7	2343,7255	5903,0625	0,1378378	0,013658	0,1623825	0,2515137	
8	2354,3687	5913,1094	0,1459459	0,0144994	0,1824167	0,2569686	
9	2360,1624	5921,4219	0,1271989	0,0139829	0,1724568	0,2595764	
10	2358,0978	5916,4844	0,1351351	0,0149329	0,1799568	0,2429987	
Average	2331,0447	5901,9859	0,1350847	0,014204	0,1804684	0,2635828	
p-value	0,104301	0,5581512	0,6322295	0,5058315	0,7861374	0,1372999	
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha 0,05

Tabela 54 – Resultados e teste Shapiro-Wilk para o conjunto de dados Birds

Dataset	Birds						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	1,2694755	2,3296875	0,4438942	0,054294	0	0		
2	1,2681411	2,178125	0,4423317	0,0543763	0	0		
3	1,2519492	2,2703125	0,4438942	0,054294	0	0		
4	1,2742995	2,2390625	0,4469952	0,0542131	0	0		
5	1,2845955	2,21875	0,4454567	0,054294	0	0		
6	1,2847182	2,2375	0,4454327	0,0541321	0	0		
7	1,271922	2,1828125	0,4454327	0,0541321	0	0		
8	1,2857565	2,196875	0,4469952	0,0540486	0	0		
9	1,2644842	2,140625	0,4438942	0,0543763	0	0		
10	1,3041345	2,159375	0,4454567	0,0541308	0	0		
Average	1,2759476	2,2153125	0,4449784	0,0542291	0	0		
p-value	0,8142419	0,7513159	0,2992412	0,276579	#VALOR!	#VALOR!		
Shapiro-Wilk	yes	yes	yes	yes	#VALOR!	#VALOR!	alpha	0,05

Tabela 55 – Resultados e teste Shapiro-Wilk para o conjunto de dados Corel5k

Dataset	Corel5k						nFold	1
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	918,93275	2492,2969	0,006	0,0096096	0,0370601	0,0826953		
2	898,93423	2482,8906	0,004	0,0098021	0,0236878	0,1019108		
3	899,4289	2483,7656	0,008	0,0095668	0,0302733	0,1121588		
4	900,5736	2489,4063	0,004	0,0097487	0,0299298	0,0925834		
5	888,68577	2468,0313	0,01	0,0094599	0,034343	0,1123934		
6	895,48303	2460,9063	0,004	0,0098984	0,0229547	0,0983926		
7	901,32805	2471,7656	0,01	0,0096257	0,0229675	0,0918264		
8	900,29185	2453,75	0	0,0096898	0,0275746	0,1020813		
9	905,39254	2460,9375	0,004	0,0094332	0,0361207	0,1197605		
10	896,03573	2468,3438	0,012	0,0095027	0,038309	0,1216016		
Average	900,50865	2473,2094	0,0062	0,0096337	0,0303221	0,1035404		
p-value	0,1123352	0,5476461	0,4384916	0,8626601	0,2244463	0,7573663		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 56 – Resultados e teste Shapiro-Wilk para o conjunto de dados Emotions

Dataset	Emotions						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,6117128	0,8	0,0067232	0,3097128	0,0750957	0,0941918		
2	0,653124	0,8359375	0,0050565	0,3102872	0,0709892	0,0889622		
3	0,6425763	0,771875	0,0067514	0,3113983	0,0663388	0,0849121		
4	0,6052423	0,7859375	0,0118079	0,3102542	0,0680019	0,0855252		
5	0,6185348	0,834375	0,0084181	0,3116855	0,0659634	0,0836931		
6	0,6084592	0,81875	0,0084463	0,3108192	0,0695947	0,088737		
7	0,6256068	0,784375	0,0067514	0,308291	0,0724137	0,0938768		
8	0,646332	0,8125	0,0084181	0,3094162	0,0745506	0,0948416		
9	0,6450493	0,7921875	0,010113	0,3088418	0,083301	0,1054054		
10	0,6477157	0,834375	0,0134746	0,3099765	0,0813438	0,104169		
Average	0,6304353	0,8070313	0,008596	0,3100683	0,0727593	0,0924314		
p-value	0,1137806	0,3099951	0,4527267	0,9681549	0,3545063	0,1980634		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 57 – Resultados e teste Shapiro-Wilk para o conjunto de dados Genbase

Dataset	Genbase						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	3,08325	7,846875	0,9470828	0,0025772	0,5830031	0,9718476		
2	2,9715671	7,8421875	0,9470828	0,0026333	0,5809043	0,971175		
3	2,8997045	7,778125	0,9470828	0,0026333	0,5821487	0,971289		
4	3,0706008	7,9578125	0,9485979	0,0025772	0,5837889	0,9719271		
5	2,9423648	7,934375	0,9440751	0,0026886	0,5831278	0,9706196		
6	2,9736569	7,909375	0,9425599	0,0027447	0,5802055	0,970009		
7	2,9452262	7,9234375	0,9470828	0,0026333	0,5825563	0,9711758		
8	2,9215498	7,84375	0,9485979	0,0026333	0,5825544	0,9713061		
9	2,8827053	7,8265625	0,9440525	0,0026333	0,583254	0,971274		
10	2,8707665	7,8296875	0,9455676	0,0026894	0,5837686	0,9707896		
Average	2,9561392	7,8692188	0,9461782	0,0026444	0,5825311	0,9711413		
p-value	0,2083788	0,3880673	0,2012733	0,1527327	0,167293	0,4622136		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 58 – Resultados e teste Shapiro-Wilk para o conjunto de dados Medical

Dataset	Medical						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	11,967848	33,56875	0,4989901	0,0148839	0,2413419	0,6836779		
2	11,571633	33,351563	0,4836524	0,0152701	0,2371455	0,6728288		
3	12,107066	33,901563	0,493846	0,0147479	0,2365563	0,6852801		
4	12,168675	34,045313	0,4867031	0,014976	0,2430218	0,6804052		
5	12,407558	34,3875	0,4908058	0,0149749	0,2397783	0,6795131		
6	12,08858	33,982813	0,4887965	0,0149513	0,2390197	0,6803988		
7	12,081702	33,851563	0,4887124	0,014839	0,2416109	0,6822565		
8	12,334434	34,36875	0,4979592	0,0147488	0,2439771	0,6870088		
9	12,92792	34,379688	0,4785399	0,0154298	0,2306994	0,6655062		
10	12,794241	34,3875	0,4999369	0,0150897	0,2399184	0,6798377		
Average	12,244966	34,0225	0,4907942	0,0149911	0,2393069	0,6796713		
p-value	0,5854625	0,1449679	0,7289745	0,2779332	0,3167969	0,1021485		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 59 – Resultados e teste Shapiro-Wilk para o conjunto de dados Yeast

Dataset	Yeast						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	40,113931	43,445313	0,193625	0,1906684	0,4051801	0,652798		
2	38,239437	42,178125	0,1948579	0,1913495	0,4034771	0,651865		
3	39,184046	42,720313	0,1931998	0,1912019	0,4059954	0,6514507		
4	39,889728	43,345313	0,1948544	0,1911122	0,407023	0,6521155		
5	39,321522	42,814063	0,1961027	0,1906397	0,4062798	0,653244		
6	37,955473	42,071875	0,1915538	0,1923823	0,4024007	0,6499681		
7	37,16388	41,0625	0,1932033	0,1919692	0,4017729	0,6506753		
8	37,104423	41,201563	0,1940348	0,1913194	0,4052188	0,6521748		
9	37,055188	40,995313	0,1923751	0,1912599	0,4068365	0,652313		
10	37,048592	41,034375	0,1923734	0,191052	0,4062951	0,6524421		
Average	38,307622	42,086875	0,193618	0,1912954	0,4050479	0,6519047		
p-value	0,0826183	0,1241293	0,9056332	0,2185767	0,1212495	0,5438845		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

B.1.2 CLMLC

Tabela 60 – Resultados e teste Shapiro-Wilk para o conjunto de dados Bibtex

Dataset	Bibtex						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	6,4007583	20,44375	0,1720056	0,0137378	0,2120616	0,4056491		
2	6,1022684	20,514063	0,1722744	0,0136103	0,2071454	0,4051464		
3	6,0591693	20,323438	0,1755285	0,0137855	0,2091536	0,4049979		
4	6,0341325	20,032813	0,1729523	0,0137889	0,2106679	0,4036823		
5	6,0011479	20,2125	0,1706519	0,0138281	0,2090456	0,4037512		
6	5,952845	20,134375	0,1736293	0,0137752	0,2062986	0,4045874		
7	5,9702738	20,392188	0,1741764	0,0137242	0,2112263	0,4054038		
8	5,8540398	20,221875	0,1757922	0,0137191	0,2061993	0,4031458		
9	5,9127478	20,084375	0,1763316	0,0137013	0,2096902	0,4054057		
10	6,2261898	20,039063	0,1737639	0,013663	0,2124576	0,4046091		
Average	6,0513572	20,239844	0,1737106	0,0137333	0,2093946	0,4046379		
p-value	0,3100169	0,4837628	0,877017	0,8871424	0,4454974	0,3162712		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 61 – Resultados e teste Shapiro-Wilk para o conjunto de dados Birds

Dataset	Birds						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,0345844	0,121875	0,5178365	0,046503	0,3318635	0,4928724		
2	0,0329596	0,1203125	0,5116346	0,0473128	0,3173483	0,4754131		
3	0,0319925	0,1265625	0,5208654	0,0478112	0,3182516	0,4712872		
4	0,0326499	0,1109375	0,5147837	0,0456048	0,3349604	0,4896076		
5	0,0348679	0,1203125	0,5286779	0,0454365	0,3472433	0,493712		
6	0,032875	0,1296875	0,525601	0,0441346	0,3153153	0,4924271		
7	0,0326299	0,128125	0,5270913	0,045124	0,330037	0,4932205		
8	0,0341136	0,1328125	0,5178606	0,0475746	0,3234527	0,4733296		
9	0,0331242	0,1296875	0,4976683	0,0489626	0,3161541	0,4536143		
10	0,0320949	0,1203125	0,525601	0,0465056	0,3233116	0,4769064		
Average	0,0331892	0,1240625	0,518762	0,046497	0,3257938	0,481239		
p-value	0,2225296	0,3981125	0,152252	0,975518	0,2053415	0,0616596		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 62 – Resultados e teste Shapiro-Wilk para o conjunto de dados Corel5k

Dataset	Corel5k						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	2,1975285	8,2734375	0,03	0,0135626	0,0370791	0,1748934		
2	2,2428207	8,246875	0,0288	0,0136524	0,0371236	0,1737701		
3	2,221231	8,36875	0,0278	0,0133834	0,0369838	0,1768728		
4	2,1866071	8,2265625	0,031	0,0136604	0,036347	0,1757331		
5	2,2482809	8,1640625	0,0296	0,0137524	0,0362133	0,171864		
6	2,3109329	7,975	0,0282	0,0136781	0,0354861	0,169064		
7	2,3115464	7,74375	0,0288	0,0138738	0,0348176	0,1658793		
8	2,3326688	7,74375	0,0286	0,0134139	0,0358521	0,1764226		
9	2,2439371	8,0015625	0,0282	0,0135364	0,0372455	0,1746484		
10	2,1840629	8,0640625	0,03	0,013538	0,0363691	0,1749876		
Average	2,2479616	8,0807813	0,0291	0,0136051	0,0363517	0,1734135		
p-value	0,2249436	0,3713872	0,5025191	0,8854186	0,3644693	0,0614338		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 63 – Resultados e teste Shapiro-Wilk para o conjunto de dados Emotions

Dataset	Emotions						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,0185303	0,0484375	0,2629379	0,2454237	0,5865555	0,6010473		
2	0,0163251	0,046875	0,2681356	0,2420339	0,5908351	0,5993481		
3	0,0147568	0,05625	0,2290678	0,2472128	0,5806784	0,599407		
4	0,0133425	0,040625	0,2578531	0,2426554	0,5871122	0,6025059		
5	0,012666	0,0296875	0,2829944	0,2368173	0,5985555	0,6125821		
6	0,0142937	0,04375	0,2561299	0,2390819	0,5926981	0,6068501		
7	0,013642	0,04375	0,2831356	0,2387288	0,5933748	0,6090468		
8	0,0126993	0,040625	0,2614124	0,2445433	0,5836177	0,5929094		
9	0,0129937	0,0390625	0,2527684	0,2480132	0,5863042	0,596628		
10	0,0145454	0,0375	0,2376554	0,2530744	0,5741953	0,5832063		
Average	0,0143795	0,0426563	0,259209	0,2437585	0,5873927	0,6003531		
p-value	0,0685328	0,9135914	0,6303527	0,909956	0,9701186	0,8414946		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 64 – Resultados e teste Shapiro-Wilk para o conjunto de dados Genbase

Dataset	Genbase						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,1015787	0,353125	0,9637268	0,0021282	0,6538669	0,9768851		
2	0,0975367	0,33125	0,9692497	0,0018477	0,648472	0,9804043		
3	0,1090338	0,321875	0,9667571	0,0020143	0,6710344	0,9783652		
4	0,0980453	0,334375	0,9652646	0,0021249	0,6603423	0,9771229		
5	0,0953422	0,328125	0,9682723	0,0019004	0,6430819	0,9796692		
6	0,0949396	0,321875	0,9667571	0,0019565	0,6418473	0,9790165		
7	0,1032169	0,3203125	0,9667797	0,0020688	0,6603423	0,9777145		
8	0,1027608	0,35	0,9683175	0,0018988	0,637975	0,9796097		
9	0,0977657	0,33125	0,9668024	0,002011	0,6377751	0,9781928		
10	0,1065128	0,3421875	0,9682949	0,0018452	0,6386733	0,9800815		
Average	0,1006733	0,3334375	0,9670222	0,0019796	0,6493411	0,9787062		
p-value	0,47853	0,2532614	0,3215964	0,339384	0,1730804	0,6095432		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 65 – Resultados e teste Shapiro-Wilk para o conjunto de dados Medical

Dataset	Medical						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,3209496	1,0578125	0,7167263	0,0107041	0,3380337	0,8020853		
2	0,3443601	1,1015625	0,6942457	0,0112493	0,3263581	0,7909406		
3	0,3414961	1,0578125	0,6901746	0,0114985	0,3308191	0,7873797		
4	0,3245554	1,0546875	0,6953293	0,0112923	0,3361008	0,7906626		
5	0,3329435	1,059375	0,7086051	0,010907	0,3341534	0,7971158		
6	0,3084471	1,0328125	0,6952556	0,011318	0,3280274	0,7908377		
7	0,303058	1,0625	0,7045024	0,0111118	0,3330257	0,7937121		
8	0,3389611	1,090625	0,6861246	0,0118379	0,3312717	0,7820447		
9	0,3332302	1,0203125	0,6963076	0,0112944	0,3323749	0,7908369		
10	0,3322039	1,0609375	0,7004313	0,0109973	0,3311388	0,796551		
Average	0,3280205	1,0598438	0,6987703	0,0112211	0,3321304	0,7922166		
p-value	0,3245646	0,3067311	0,6613912	0,9001794	0,9671518	0,7978765		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 66 – Resultados e teste Shapiro-Wilk para o conjunto de dados Yeast

Dataset	Yeast						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,1593983	0,596875	0,2031292	0,2281805	0,4061026	0,6171935		
2	0,1597529	0,6	0,203148	0,2291592	0,4042236	0,6146407		
3	0,160358	0,596875	0,1915589	0,232847	0,3963689	0,6090576		
4	0,1627685	0,6125	0,1886664	0,234329	0,3995298	0,6061745		
5	0,1585	0,5890625	0,1948647	0,2361334	0,3985261	0,6041513		
6	0,1629421	0,596875	0,1828641	0,2313969	0,4092173	0,6145854		
7	0,1652197	0,6203125	0,1857635	0,2366006	0,3868303	0,6026416		
8	0,1631672	0,6140625	0,2023233	0,2264324	0,4081767	0,6193562		
9	0,1651362	0,6203125	0,1890744	0,2322002	0,4009197	0,6091815		
10	0,1633889	0,6109375	0,1940263	0,2297783	0,4030727	0,6167978		
Average	0,1620632	0,6057813	0,1935419	0,2317057	0,4012968	0,611378		
p-value	0,3099588	0,2264184	0,3244742	0,8811936	0,4337224	0,4279762		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

B.1.3 CBMLC

Tabela 67 – Resultados e teste Shapiro-Wilk para o conjunto de dados Bibtex

Dataset	Bibtex						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	118,76626	387,55469	0,1544242	0,0162857	0,248722	0,3744065		
2	119,89788	390,83125	0,1533407	0,0161638	0,2512767	0,3780432		
3	121,17495	396,33594	0,1438776	0,0167008	0,2469711	0,3691593		
4	120,35499	391,87188	0,1557741	0,0165181	0,2473561	0,3723822		
5	122,14936	400,64063	0,1471256	0,0170033	0,2375737	0,3596386		
6	118,8498	387,53281	0,1555056	0,0168279	0,2437393	0,3653615		
7	114,87873	375,16719	0,1625418	0,0153367	0,2363288	0,3767832		
8	117,97804	387,45781	0,1560529	0,0164602	0,247883	0,3733035		
9	115,86212	380,87344	0,1598398	0,0156681	0,2421831	0,3764648		
10	114,01203	374,24375	0,1545593	0,0161374	0,2420255	0,3730285		
Average	118,39242	387,25094	0,1543042	0,0163102	0,2444059	0,3718571		
p-value	0,6619789	0,6607619	0,376767	0,7543454	0,5654319	0,1695496		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 68 – Resultados e teste Shapiro-Wilk para o conjunto de dados Birds

Dataset	Birds						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,1224564	0,459375	0,5057212	0,0592801	0,3025914	0,4382703		
2	0,1180245	0,446875	0,508726	0,0586285	0,3037657	0,4445684		
3	0,1187923	0,45625	0,5087981	0,0586247	0,305105	0,4411732		
4	0,1178382	0,453125	0,508774	0,0583831	0,3106078	0,4520128		
5	0,1252063	0,446875	0,5026202	0,059356	0,2979995	0,4351934		
6	0,1352528	0,4296875	0,5071635	0,0583009	0,3131082	0,4475654		
7	0,12721	0,425	0,5010096	0,0600114	0,2999736	0,4317109		
8	0,1267752	0,4640625	0,4978846	0,0595218	0,3041021	0,4355327		
9	0,11852	0,446875	0,5026202	0,0600089	0,3053095	0,4337468		
10	0,1158925	0,434375	0,502476	0,0590385	0,3008206	0,4331419		
Average	0,1225968	0,44625	0,5045793	0,0591154	0,3043383	0,4392916		
p-value	0,1779374	0,5654981	0,2294753	0,3938905	0,5358484	0,2899376		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 69 – Resultados e teste Shapiro-Wilk para o conjunto de dados Corel5k

Dataset	Corel5k						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	35,264717	110,83594	0,0096	0,0118316	0,0256257	0,1753521		
2	36,642758	108,60156	0,0128	0,0115749	0,0254845	0,1701501		
3	37,75137	108,19219	0,01	0,0119364	0,0242279	0,164923		
4	36,725363	112,66094	0,012	0,0119444	0,0251392	0,1609476		
5	34,876733	111,925	0,0102	0,011838	0,0238281	0,1654304		
6	35,787415	112,14688	0,0116	0,0116091	0,0248863	0,1666635		
7	35,602017	111,50156	0,0112	0,0117947	0,0244516	0,1674988		
8	35,761189	110,97813	0,012	0,0120465	0,0254633	0,1753352		
9	35,967246	112,02344	0,0098	0,0121679	0,0243743	0,1785702		
10	39,146801	110,64219	0,0112	0,011938	0,0237021	0,165872		
Average	36,352561	110,95078	0,01104	0,0118681	0,0247183	0,1690743		
p-value	0,1769609	0,1117131	0,4782291	0,7261604	0,4114386	0,3855594		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 70 – Resultados e teste Shapiro-Wilk para o conjunto de dados Emotions

Dataset	Emotions						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,0091985	0,0375	0,2598588	0,2318315	0,6173797	0,6269191		
2	0,0069416	0,021875	0,2378249	0,2380461	0,6062172	0,6160806		
3	0,0078731	0,0203125	0,254774	0,2340819	0,6138257	0,6252126		
4	0,0077174	0,03125	0,244661	0,2472175	0,593019	0,6044882		
5	0,007396	0,0265625	0,2463559	0,2371657	0,6066545	0,6192393		
6	0,007143	0,0203125	0,2377684	0,2430932	0,5985633	0,6097245		
7	0,0068245	0,025	0,2481073	0,2354002	0,6118817	0,6232482		
8	0,0068291	0,0203125	0,2513559	0,2329379	0,6116429	0,6273261		
9	0,0075584	0,0296875	0,2563559	0,238032	0,6070572	0,618603		
10	0,00822	0,025	0,2632768	0,2348588	0,6114787	0,623603		
Average	0,0075701	0,0257813	0,2500339	0,2372665	0,607772	0,6194445		
p-value	0,2016647	0,1684388	0,7616147	0,2183452	0,3290486	0,2183862		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 71 – Resultados e teste Shapiro-Wilk para o conjunto de dados Genbase

Dataset	Genbase						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	2,237083	6,609375	0,9818634	0,0008401	0,6582734	0,9908638		
2	1,9632534	6,6015625	0,9788557	0,0008954	0,6485164	0,990308		
3	2,0335263	6,3859375	0,9803709	0,0008392	0,659508	0,9909981		
4	2,1634342	6,6328125	0,9818634	0,0007278	0,6633755	0,9922861		
5	2,4056102	6,6234375	0,981886	0,0006709	0,6616028	0,9930027		
6	2,2392868	6,496875	0,9833786	0,0006717	0,6645658	0,9928289		
7	2,0664115	6,5296875	0,9818634	0,0007278	0,6571583	0,9922372		
8	2,2261342	6,4171875	0,9803483	0,0008962	0,6521006	0,9903032		
9	2,032353	6,6671875	0,981886	0,000727	0,6620613	0,992334		
10	2,0489985	6,759375	0,9848937	0,0006156	0,6645658	0,9934205		
Average	2,1416091	6,5723438	0,9817209	0,0007612	0,6591728	0,9918582		
p-value	0,3882655	0,7686664	0,3815875	0,2656437	0,1792927	0,2312921		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 72 – Resultados e teste Shapiro-Wilk para o conjunto de dados Medical

Dataset	Medical						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	6,3025295	21,359375	0,6021671	0,0130467	0,3341595	0,7703028		
2	6,3358909	21,5875	0,5847675	0,0134995	0,3284796	0,7612895		
3	6,32893	21,49375	0,5889017	0,0137034	0,3361733	0,7596039		
4	6,0056306	20,453125	0,584778	0,0140031	0,3311212	0,7547784		
5	6,1692838	20,94375	0,5766148	0,0140931	0,3350544	0,7521559		
6	6,4751597	22,065625	0,5776036	0,0132969	0,3263727	0,7636631		
7	6,1527825	20,970313	0,5919419	0,0136169	0,3255734	0,7591127		
8	5,7801988	19,698438	0,5899222	0,0132501	0,3365282	0,7691363		
9	6,2939292	21,25625	0,5909741	0,0132279	0,3331884	0,7668245		
10	6,0725961	20,6	0,5878603	0,0132524	0,332628	0,7647825		
Average	6,1916931	21,042813	0,5875531	0,013499	0,3319279	0,762165		
p-value	0,613677	0,902337	0,5560333	0,3189514	0,3217916	0,8725581		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 73 – Resultados e teste Shapiro-Wilk para o conjunto de dados Yeast

Dataset	Yeast						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,060588	0,225	0,1960821	0,2185176	0,3822071	0,620685		
2	0,059499	0,21875	0,194016	0,217656	0,3837695	0,6182734		
3	0,057099	0,203125	0,1981688	0,2168014	0,3892786	0,6243455		
4	0,0571378	0,203125	0,1894877	0,2200784	0,382154	0,6187149		
5	0,0577724	0,209375	0,1927763	0,2187271	0,379052	0,6191854		
6	0,0621338	0,209375	0,1919413	0,2201978	0,3784584	0,615677		
7	0,0593782	0,1859375	0,196936	0,2206996	0,3730207	0,6153211		
8	0,0564419	0,209375	0,1923751	0,2194023	0,3845436	0,6192716		
9	0,057809	0,2015625	0,1911166	0,2193988	0,3718595	0,6170428		
10	0,0584857	0,196875	0,2035373	0,215528	0,3885584	0,6245873		
Average	0,0586345	0,20625	0,1946437	0,2187007	0,3812902	0,6193104		
p-value	0,5283916	0,8595571	0,4189992	0,5440459	0,6162444	0,2814691		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

B.1.4 HOMER

Tabela 74 – Resultados e teste Shapiro-Wilk para o conjunto de dados Bibtex

Dataset	Bibtex						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	13,31345	28,175	0,0969528	0,0223211	0,141043	0,2723002		
2	14,587563	30,78125	0,0969528	0,0216092	0,1380363	0,2750879		
3	13,443727	28,235938	0,0928989	0,0216618	0,1405498	0,2776778		
4	13,285983	27,7875	0,094385	0,022492	0,1363961	0,2753221		
5	13,437977	28,582813	0,0972212	0,0212664	0,1414128	0,2823415		
6	13,378678	27,709375	0,093709	0,0217868	0,1382592	0,2752266		
7	14,191533	29,554688	0,0979026	0,0222064	0,1394917	0,2737887		
8	13,937126	28,175	0,1001966	0,021989	0,1396608	0,2786635		
9	12,952443	26,392188	0,0958726	0,0219596	0,1395106	0,2749507		
10	14,076119	28,978125	0,0969557	0,0215734	0,1432945	0,2801886		
Average	13,66046	28,437188	0,0963047	0,0218866	0,1397655	0,2765548		
p-value	0,4075719	0,70092	0,6319479	0,9474991	0,985206	0,5175024		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 75 – Resultados e teste Shapiro-Wilk para o conjunto de dados Birds

Dataset	Birds						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,0628761	0,1859375	0,4683654	0,0604466	0,2287405	0,3890612		
2	0,0630199	0,190625	0,4604567	0,0623317	0,2427702	0,3735969		
3	0,0587685	0,1796875	0,4883894	0,0579314	0,2409572	0,4034842		
4	0,0589255	0,1734375	0,4682692	0,0601506	0,2235861	0,3790824		
5	0,0601378	0,1953125	0,483726	0,0596736	0,2366501	0,3943994		
6	0,0577009	0,1671875	0,4760096	0,0584337	0,2401483	0,4106729		
7	0,0599682	0,1828125	0,480649	0,0577771	0,2345176	0,4068832		
8	0,0579874	0,165625	0,4790865	0,0570369	0,23074	0,4025037		
9	0,05942	0,1640625	0,4603846	0,0612082	0,2297269	0,3808898		
10	0,0608264	0,1890625	0,4914663	0,0574608	0,2248906	0,3934766		
Average	0,0599631	0,179375	0,4756803	0,0592451	0,2332728	0,393405		
p-value	0,305505	0,4412094	0,5630639	0,5353997	0,5882324	0,6587089		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 76 – Resultados e teste Shapiro-Wilk para o conjunto de dados Corel5k

Dataset	Corel5k						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	69,746622	125,99531	0,0014	0,0171759	0,0335599	0,1699389		
2	72,347737	129,67031	0,0006	0,0164824	0,0339201	0,1710059		
3	75,15166	133,00156	0,001	0,0165503	0,0347116	0,1657211		
4	82,753285	143,21406	0,0014	0,0166545	0,0350414	0,1768889		
5	79,673939	138,46406	0,0012	0,0162941	0,0326507	0,160015		
6	75,631269	134,25781	0,0012	0,0159428	0,0336343	0,1754316		
7	74,46616	130,15	0,0008	0,0165535	0,03307	0,1692785		
8	81,504968	143,45469	0,001	0,0168497	0,0343217	0,1470935		
9	78,989771	138,38281	0,0012	0,016584	0,0332812	0,1698895		
10	82,153863	145,41406	0,001	0,0167	0,033859	0,1665707		
Average	77,241927	136,20047	0,00108	0,0165787	0,033805	0,1671834		
p-value	0,566217	0,6282618	0,4453433	0,7294129	0,9788942	0,0804837		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 77 – Resultados e teste Shapiro-Wilk para o conjunto de dados Emotions

Dataset	Emotions						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,0169331	0,046875	0,3037571	0,2068362	0,6619583	0,6781408		
2	0,0151408	0,0578125	0,2800565	0,2099765	0,6597767	0,6765824		
3	0,0148853	0,0546875	0,2814972	0,2131591	0,656011	0,6725701		
4	0,0149133	0,059375	0,2934463	0,2100094	0,6564243	0,6763578		
5	0,0165459	0,0546875	0,284887	0,2066573	0,6704246	0,6832693		
6	0,0163248	0,05	0,2664124	0,2167467	0,6508514	0,6646896		
7	0,0159996	0,0578125	0,2783616	0,2102872	0,6628002	0,6754865		
8	0,0158715	0,05	0,2883051	0,2080556	0,6624381	0,677962		
9	0,0148965	0,034375	0,2648305	0,2108145	0,6626814	0,6730962		
10	0,0162907	0,0375	0,2949435	0,2066714	0,6613563	0,6809681		
Average	0,0157801	0,0503125	0,2836497	0,2099214	0,6604722	0,6759123		
p-value	0,1709056	0,1248187	0,8610823	0,1618957	0,5030181	0,5011364		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 78 – Resultados e teste Shapiro-Wilk para o conjunto de dados Genbase

Dataset	Genbase						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	1,6933532	5,09375	0,9652872	0,0016224	0,6119803	0,9824346		
2	1,5946389	4,609375	0,9667797	0,001511	0,611381	0,9834578		
3	1,6729627	4,515625	0,9683175	0,001454	0,6150237	0,9841884		
4	1,6445829	4,796875	0,9698327	0,0013418	0,616172	0,9854423		
5	1,6483034	4,796875	0,9668024	0,001454	0,6148731	0,9842211		
6	1,583485	4,421875	0,9698327	0,0013979	0,6134618	0,9849706		
7	1,6875254	4,90625	0,96981	0,0013426	0,6130027	0,9853386		
8	1,7292176	5,109375	0,9713252	0,0012865	0,6158427	0,9860071		
9	1,6888881	4,890625	0,9683175	0,001454	0,614344	0,9842065		
10	1,697507	4,4375	0,9682949	0,001511	0,6131447	0,9835875		
Average	1,6640464	4,7578125	0,96846	0,0014375	0,6139226	0,9843854		
p-value	0,3397325	0,379783	0,7093488	0,7961999	0,8210605	0,8821579		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 79 – Resultados e teste Shapiro-Wilk para o conjunto de dados Medical

Dataset	Medical						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,4122062	1,178125	0,5644435	0,0171118	0,2241661	0,6922703		
2	0,4012394	1,15625	0,5756364	0,0162714	0,2220586	0,7028662		
3	0,4207621	1,1328125	0,5776036	0,0163658	0,23999	0,703493		
4	0,412301	1,1546875	0,5551862	0,0167929	0,2345804	0,6983101		
5	0,4088029	1,1265625	0,5797076	0,0161117	0,2259162	0,7086982		
6	0,4255765	1,125	0,5643909	0,0170666	0,2181739	0,6920847		
7	0,4414593	1,10625	0,5674732	0,0169984	0,2225269	0,6926891		
8	0,437355	1,2421875	0,574595	0,0163188	0,2225226	0,7028835		
9	0,4299917	1,240625	0,578603	0,016708	0,21526	0,6974441		
10	0,4325801	1,1609375	0,5602882	0,0164759	0,2293676	0,7055134		
Average	0,4222274	1,1623438	0,5697928	0,0166221	0,2254562	0,6996252		
p-value	0,7957662	0,1074199	0,3127513	0,3828036	0,564341	0,3010986		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

Tabela 80 – Resultados e teste Shapiro-Wilk para o conjunto de dados Yeast

Dataset	Yeast						nFold	10
Run	tExe (s)	tCPU (s)	EM	HL	F1-macro	F1-micro		
1	0,0756316	0,246875	0,1464713	0,2327571	0,421455	0,6486819		
2	0,0704473	0,225	0,151454	0,2318595	0,4248668	0,6520733		
3	0,070845	0,2265625	0,140285	0,2395986	0,4233469	0,643081		
4	0,0738246	0,2375	0,1440005	0,2385177	0,4232775	0,6455042		
5	0,0743141	0,246875	0,1443881	0,2340286	0,421711	0,6492297		
6	0,0727437	0,253125	0,1439937	0,2364452	0,4299202	0,6475917		
7	0,0729271	0,2640625	0,1481465	0,2385632	0,4189004	0,644109		
8	0,07222	0,234375	0,1419379	0,233153	0,410096	0,6485117		
9	0,0719356	0,2390625	0,1423528	0,2352003	0,4219567	0,6489452		
10	0,0735846	0,259375	0,1477093	0,2334445	0,415849	0,6478336		
Average	0,0728473	0,2432813	0,1450739	0,2353568	0,4211379	0,6475561		
p-value	0,9815764	0,7949825	0,8341675	0,2910421	0,5520136	0,5954989		
Shapiro-Wilk	yes	yes	yes	yes	yes	yes	alpha	0,05

B.2 Teste de Hipóteses 1

B.2.1 Tempo de Execução

Tabela 81 – Teste ANOVA para o tempo de execução - conjunto de dados Bibtex

Parameter	tExe			
Dataset	Bibtex			
Run	LIFT	CLMLC	CBMLC	HOMER
1	2332,7178	6,4007583	118,76626	13,31345
2	2298,6894	6,1022684	119,89788	14,587563
3	2375,3737	6,0591693	121,17495	13,443727
4	2353,3421	6,0341325	120,35499	13,285983
5	2253,8177	6,0011479	122,14936	13,437977
6	2280,1517	5,952845	118,8498	13,378678
7	2343,7255	5,9702738	114,87873	14,191533
8	2354,3687	5,8540398	117,97804	13,937126
9	2360,1624	5,9127478	115,86212	12,952443
10	2358,0978	6,2261898	114,01203	14,076119
Average	2331,0447	6,0513572	118,39242	13,66046

ANOVA: Single Factor

DESCRIPTION	Alpha	0,05						
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	23310,447	2331,0447	1594,503	14350,527	6,3288407	2318,2092	2343,8802
CLMLC	10	60,513572	6,0513572	0,0258382	0,2325441	6,3288407	6,7841265	18,886841
CBMLC	10	1183,9242	118,39242	7,3813436	66,432093	6,3288407	105,55693	131,2279
HOMER	10	136,6046	13,66046	0,2587428	2,3286856	6,3288407	0,824976	26,495944

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	39238352	3	13079451	32654,36	7,25E-62	2,8662656	57,143994	0,9995918
Within Groups	14419,521	36	400,54224					
Total	39252771	39	1006481,3					

Tabela 82 – Teste Tukey HSD para o tempo de execução - conjunto de dados Bibtex

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	2331,0447	10	14350,527			
CLMLC	6,0513572	10	0,2325441			
CBMLC	118,39242	10	66,432093			
HOMER	13,66046	10	2,3286856			
		40	14419,521	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	2324,9933	6,3288407	367,36481	2300,8868	2349,0999	5,773E-15	24,106554	116,17095
LIFT	CBMLC	2212,6523	6,3288407	349,61415	2188,5457	2236,7588	5,995E-15	24,106554	110,5577
LIFT	HOMER	2317,3842	6,3288407	366,16252	2293,2777	2341,4908	5,773E-15	24,106554	115,79075
CLMLC	CBMLC	112,34106	6,3288407	17,750654	88,234505	136,44761	6,95E-14	24,106554	5,6132496
CLMLC	HOMER	7,6091026	6,3288407	1,2022901	16,497451	31,715657	0,8300927	24,106554	0,3801975
CBMLC	HOMER	104,73196	6,3288407	16,548364	80,625403	128,83851	4,821E-13	24,106554	5,2330521

Tabela 83 – Teste ANOVA para o tempo de execução - conjunto de dados Birds

Parameter	tExe			
Dataset	Birds			
Run	LIFT	CLMLC	CBMLC	HOMER
1	1,2694755	0,0345844	0,1224564	0,0628761
2	1,2681411	0,0329596	0,1180245	0,0630199
3	1,2519492	0,0319925	0,1187923	0,0587685
4	1,2742995	0,0326499	0,1178382	0,0589255
5	1,2845955	0,0348679	0,1252063	0,0601378
6	1,2847182	0,032875	0,1352528	0,0577009
7	1,271922	0,0326299	0,12721	0,0599682
8	1,2857565	0,0341136	0,1267752	0,0579874
9	1,2644842	0,0331242	0,11852	0,05942
10	1,3041345	0,0320949	0,1158925	0,0608264
Average	1,2759476	0,0331892	0,1225968	0,0599631

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
<i>Group</i>	<i>Count</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>	<i>SS</i>	<i>Std Err</i>	<i>Lower</i>	<i>Upper</i>	
LIFT	10	12,759476	1,2759476	0,0002079	0,0018715	0,0024924	1,2708929	1,2810024	
CLMLC	10	0,3318918	0,0331892	1,001E-06	9,008E-06	0,0024924	0,0281344	0,0382439	
CBMLC	10	1,2259681	0,1225968	3,615E-05	0,0003253	0,0024924	0,1175421	0,1276516	
HOMER	10	0,5996307	0,0599631	3,381E-06	3,043E-05	0,0024924	0,0549083	0,0650178	

ANOVA

<i>Sources</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P value</i>	<i>F crit</i>	<i>RMSSE</i>	<i>Omega Sq</i>
Between Groups	10,914797	3	3,6382655	58569,382	1,97E-66	2,8662656	76,530636	0,9997724
Within Groups	0,0022363	36	6,212E-05					
Total	10,917033	39	0,2799239					

Tabela 84 – Teste Tukey HSD para o tempo de execução - conjunto de dados Birds

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	1,2759476	10	0,0018715			
CLMLC	0,0331892	10	9,008E-06			
CBMLC	0,1225968	10	0,0003253			
HOMER	0,0599631	10	3,043E-05			
		40	0,0022363	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	1,2427584	0,0024924	498,62592	1,233265	1,2522519	4,441E-15	0,0094934	0,0620958
LIFT	CBMLC	1,1533508	0,0024924	462,75333	1,1438574	1,1628442	4,885E-15	0,0094934	0,0576285
LIFT	HOMER	1,2159846	0,0024924	487,88356	1,2064911	1,225478	4,663E-15	0,0094934	0,0607581
CLMLC	CBMLC	0,0894076	0,0024924	35,872588	0,0799142	0,0989011	7,994E-15	0,0094934	0,0044674
CLMLC	HOMER	0,0267739	0,0024924	10,742356	0,0172805	0,0362673	3,256E-08	0,0094934	0,0013378
CBMLC	HOMER	0,0626337	0,0024924	25,130232	0,0531403	0,0721272	7,994E-15	0,0094934	0,0031296

Tabela 85 – Teste ANOVA para o tempo de execução - conjunto de dados Corel5k

Parameter	tExe			
Dataset	Corel5k			
Run	LIFT	CLMLC	CBMLC	HOMER
1	918,93275	2,1975285	35,264717	69,746622
2	898,93423	2,2428207	36,642758	72,347737
3	899,4289	2,221231	37,75137	75,15166
4	900,5736	2,1866071	36,725363	82,753285
5	888,68577	2,2482809	34,876733	79,673939
6	895,48303	2,3109329	35,787415	75,631269
7	901,32805	2,3115464	35,602017	74,46616
8	900,29185	2,3326688	35,761189	81,504968
9	905,39254	2,2439371	35,967246	78,989771
10	896,03573	2,1840629	39,146801	82,153863
Average	900,50865	2,2479616	36,352561	77,241927

ANOVA: Single Factor

DESCRIPTION	Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT		10	9005,0865	900,50865	61,412381	552,71143	1,4375076	897,59324	903,42405
CLMLC		10	22,479616	2,2479616	0,0029258	0,0263319	1,4375076	0,6674389	5,1633622
CBMLC		10	363,52561	36,352561	1,6366673	14,730006	1,4375076	33,43716	39,267961
HOMER		10	772,41927	77,241927	19,60515	176,44635	1,4375076	74,326527	80,157328

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	5599663,1	3	1866554,4	90327,574	8,099E-70	2,8662656	95,04082	0,9998524
Within Groups	743,91411	36	20,664281					
Total	5600407	39	143600,18					

Tabela 86 – Teste Tukey HSD para o tempo de execução - conjunto de dados Corel5k

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	900,50865	10	552,71143			
CLMLC	2,2479616	10	0,0263319			
CBMLC	36,352561	10	14,730006			
HOMER	77,241927	10	176,44635			
		40	743,91411	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	898,26068	1,4375076	624,87369	892,78522	903,73615	3,775E-15	5,4754664	44,882623
LIFT	CBMLC	864,15608	1,4375076	601,14888	858,68062	869,63155	3,997E-15	5,4754664	43,178548
LIFT	HOMER	823,26672	1,4375076	572,70426	817,79125	828,74218	4,219E-15	5,4754664	41,135464
CLMLC	CBMLC	34,104599	1,4375076	23,724813	28,629133	39,580066	7,994E-15	5,4754664	1,7040753
CLMLC	HOMER	74,993966	1,4375076	52,16944	69,518499	80,469432	7,994E-15	5,4754664	3,7471593
CBMLC	HOMER	40,889366	1,4375076	28,444626	35,4139	46,364833	7,994E-15	5,4754664	2,043084

Tabela 87 – Teste ANOVA para o tempo de execução - conjunto de dados Emotions

Parameter	tExe			
Dataset	Emotions			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,6117128	0,0185303	0,0091985	0,0169331
2	0,653124	0,0163251	0,0069416	0,0151408
3	0,6425763	0,0147568	0,0078731	0,0148853
4	0,6052423	0,0133425	0,0077174	0,0149133
5	0,6185348	0,012666	0,007396	0,0165459
6	0,6084592	0,0142937	0,007143	0,0163248
7	0,6256068	0,013642	0,0068245	0,0159996
8	0,646332	0,0126993	0,0068291	0,0158715
9	0,6450493	0,0129937	0,0075584	0,0148965
10	0,6477157	0,0145454	0,00822	0,0162907
Average	0,6304353	0,0143795	0,0075701	0,0157801

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	6,3043531	0,6304353	0,00034	0,0030602	0,0029349	0,6244831	0,6363875	
CLMLC	10	0,1437946	0,0143795	3,388E-06	3,049E-05	0,0029349	0,0084273	0,0203316	
CBMLC	10	0,0757015	0,0075701	5,445E-07	4,9E-06	0,0029349	0,001618	0,0135223	
HOMER	10	0,1578014	0,0157801	5,857E-07	5,271E-06	0,0029349	0,009828	0,0217323	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	2,8635064	3	0,9545021	11081,538	2,008E-53	2,8662656	33,288944	0,9987981
Within Groups	0,0031008	36	8,613E-05					
Total	2,8666072	39	0,0735027					

Tabela 88 – Teste Tukey HSD para o tempo de execução - conjunto de dados Emotions

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,6304353	10	0,0030602			
CLMLC	0,0143795	10	3,049E-05			
CBMLC	0,0075701	10	4,9E-06			
HOMER	0,0157801	10	5,271E-06			
		40	0,0031008	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,6160559	0,0029349	209,90929	0,6048769	0,6272348	7,438E-15	0,0111789	0,0307819
LIFT	CBMLC	0,6228652	0,0029349	212,22943	0,6116863	0,6340441	7,438E-15	0,0111789	0,0311222
LIFT	HOMER	0,6146552	0,0029349	209,43203	0,6034763	0,6258341	7,438E-15	0,0111789	0,0307119
CLMLC	CBMLC	0,0068093	0,0029349	2,320144	0,0043696	0,0179882	0,3694715	0,0111789	0,0003402
CLMLC	HOMER	0,0014007	0,0029349	0,4772564	0,0097782	0,0125796	0,9865434	0,0111789	6,999E-05
CBMLC	HOMER	0,00821	0,0029349	2,7974003	0,0029689	0,0193889	0,2150865	0,0111789	0,0004102

Tabela 89 – Teste ANOVA para o tempo de execução - conjunto de dados Genbase

Parameter	tExe			
Dataset	Genbase			
Run	LIFT	CLMLC	CBMLC	HOMER
1	3,08325	0,1015787	2,237083	1,6933532
2	2,9715671	0,0975367	1,9632534	1,5946389
3	2,8997045	0,1090338	2,0335263	1,6729627
4	3,0706008	0,0980453	2,1634342	1,6445829
5	2,9423648	0,0953422	2,4056102	1,64883034
6	2,9736569	0,0949396	2,2392868	1,583485
7	2,9452262	0,1032169	2,0664115	1,6875254
8	2,9215498	0,1027608	2,2261342	1,7292176
9	2,8827053	0,0977657	2,032353	1,6888881
10	2,8707665	0,1065128	2,0489985	1,697507
Average	2,9561392	0,1006733	2,1416091	1,6640464

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05							
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	29,561392	2,9561392	0,0052416	0,0471747	0,0254088	2,9046077	3,0076707
CLMLC	10	1,0067326	0,1006733	2,246E-05	0,0002021	0,0254088	0,0491417	0,1522048
CBMLC	10	21,416091	2,1416091	0,0184038	0,1656345	0,0254088	2,0900776	2,1931406
HOMER	10	16,640464	1,6640464	0,0021564	0,0194079	0,0254088	1,6125149	1,7155779

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	43,310674	3	14,436891	2236,1665	6,029E-41	2,8662656	14,953817	0,9940701
Within Groups	0,2324192	36	0,0064561					
Total	43,543093	39	1,1164896					

Tabela 90 – Teste Tukey HSD para o tempo de execução - conjunto de dados Genbase

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	2,9561392	10	0,0471747			
CLMLC	0,1006733	10	0,0002021			
CBMLC	2,1416091	10	0,1656345			
HOMER	1,6640464	10	0,0194079			
		40	0,2324192	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	2,8554659	0,0254088	112,38082	2,7586837	2,9522482	7,994E-15	0,0967823	0,1426766
LIFT	CBMLC	0,8145301	0,0254088	32,056961	0,7177478	0,9113123	7,994E-15	0,0967823	0,0406989
LIFT	HOMER	1,2920928	0,0254088	50,852103	1,1953105	1,388875	7,994E-15	0,0967823	0,0645609
CLMLC	CBMLC	2,0409358	0,0254088	80,323861	1,9441536	2,1377181	7,994E-15	0,0967823	0,1019777
CLMLC	HOMER	1,5633732	0,0254088	61,528719	1,4665909	1,6601554	7,994E-15	0,0967823	0,0781157
CBMLC	HOMER	0,4775627	0,0254088	18,795142	0,3807804	0,5743449	1,932E-14	0,0967823	0,023862

Tabela 91 – Teste ANOVA para o tempo de execução - conjunto de dados Medical

Parameter	tExe			
Dataset	Medical			
Run	LIFT	CLMLC	CBMLC	HOMER
1	11,967848	0,3209496	6,3025295	0,4122062
2	11,571633	0,3443601	6,3358909	0,4012394
3	12,107066	0,3414961	6,32893	0,4207621
4	12,168675	0,3245554	6,0056306	0,412301
5	12,407558	0,3329435	6,1692838	0,4088029
6	12,08858	0,3084471	6,4751597	0,4255765
7	12,081702	0,303058	6,1527825	0,4414593
8	12,334434	0,3389611	5,7801988	0,437355
9	12,92792	0,3332302	6,2939292	0,4299917
10	12,794241	0,3322039	6,0725961	0,4325801
Average	12,244966	0,3280205	6,1916931	0,4222274

ANOVA: Single Factor

DESCRIPTION	Alpha		0,05					
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	122,44966	12,244966	0,1565212	1,4086906	0,0702151	12,102563	12,387369
CLMLC	10	3,2802048	0,3280205	0,0001895	0,0017054	0,0702151	0,1856176	0,4704234
CBMLC	10	61,916931	6,1916931	0,0403183	0,3628643	0,0702151	6,0492902	6,334096
HOMER	10	4,2222741	0,4222274	0,0001777	0,0015992	0,0702151	0,2798245	0,5646303

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	965,27777	3	321,75926	6526,338	2,726E-49	2,8662656	25,546698	0,9979608
Within Groups	1,7748595	36	0,0493017					
Total	967,05263	39	24,796221					

Tabela 92 – Teste Tukey HSD para o tempo de execução - conjunto de dados Medical

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	12,244966	10	1,4086906			
CLMLC	0,3280205	10	0,0017054			
CBMLC	6,1916931	10	0,3628643			
HOMER	0,4222274	10	0,0015992			
		40	1,7748595	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	11,916945	0,0702151	169,72046	11,649496	12,184395	7,772E-15	0,2674495	0,5954438
LIFT	CBMLC	6,0532728	0,0702151	86,210368	5,7858233	6,3207222	7,994E-15	0,2674495	0,3024587
LIFT	HOMER	11,822738	0,0702151	168,37877	11,555289	12,090188	7,772E-15	0,2674495	0,5907367
CLMLC	CBMLC	5,8636726	0,0702151	83,510094	5,5962232	6,1311221	7,994E-15	0,2674495	0,2929851
CLMLC	HOMER	0,0942069	0,0702151	1,3416899	0,1732425	0,3616564	0,7788979	0,2674495	0,0047072
CBMLC	HOMER	5,7694657	0,0702151	82,168404	5,5020162	6,0369151	7,994E-15	0,2674495	0,288278

Tabela 93 – Teste ANOVA para o tempo de execução - conjunto de dados Yeast

Parameter	tExe			
Dataset	Yeast			
Run	LIFT	CLMLC	CBMLC	HOMER
1	40,113931	0,1593983	0,060588	0,0756316
2	38,239437	0,1597529	0,059499	0,0704473
3	39,184046	0,160358	0,057099	0,070845
4	39,889728	0,1627685	0,0571378	0,0738246
5	39,321522	0,1585	0,0577724	0,0743141
6	37,955473	0,1629421	0,0621338	0,0727437
7	37,16388	0,1652197	0,0593782	0,0729271
8	37,104423	0,1631672	0,0564419	0,07222
9	37,055188	0,1651362	0,057809	0,0719356
10	37,048592	0,1633889	0,0584857	0,0735846
Average	38,307622	0,1620632	0,0586345	0,0728473

ANOVA: Single Factor

DESCRIPTION	Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT		10	383,07622	38,307622	1,5111293	13,600164	0,194367	37,913428	38,701817
CLMLC		10	1,6206317	0,1620632	5,744E-06	5,17E-05	0,194367	0,2321313	0,5562576
CBMLC		10	0,5863447	0,0586345	3,135E-06	2,821E-05	0,194367	-0,33556	0,4528289
HOMER		10	0,7284735	0,0728473	2,49E-06	2,241E-05	0,194367	0,3213471	0,4670418

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	10949,964	3	3649,988	9661,5437	2,363E-52	2,8662656	31,083024	0,9986217
Within Groups	13,600266	36	0,3777852					
Total	10963,564	39	281,11703					

Tabela 94 – Teste Tukey HSD para o tempo de execução - conjunto de dados Yeast

TUKEY HSD/KRAMER			alpha		0,05	
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	38,307622	10	13,600164			
CLMLC	0,1620632	10	5,17E-05			
CBMLC	0,0586345	10	2,821E-05			
HOMER	0,0728473	10	2,241E-05			
		40	13,600266	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	38,145559	0,194367	196,25536	37,405215	38,885903	7,772E-15	0,7403438	1,9059865
LIFT	CBMLC	38,248988	0,194367	196,78749	37,508644	38,989331	7,772E-15	0,7403438	1,9111544
LIFT	HOMER	38,234775	0,194367	196,71437	37,494431	38,975118	7,772E-15	0,7403438	1,9104443
CLMLC	CBMLC	0,1034287	0,194367	0,5321311	-	0,6369151	0,8437725	0,9815533	0,7403438
CLMLC	HOMER	0,0892158	0,194367	0,4590071	-0,651128	0,8295596	0,9879877	0,7403438	0,0044578
CBMLC	HOMER	0,0142129	0,194367	0,073124	-	0,7261309	0,7545567	0,9999494	0,7403438

B.2.2 Tempo de CPU

Tabela 95 – Teste ANOVA para o tempo de CPU - conjunto de dados Bibtex

Parameter	tCPU			
Dataset	Bibtex			
Run	LIFT	CLMLC	CBMLC	HOMER
1	5881,4063	20,44375	387,55469	28,175
2	5848,2656	20,514063	390,83125	30,78125
3	5959,3125	20,323438	396,33594	28,235938
4	5963,6719	20,032813	391,87188	27,7875
5	5844,9063	20,2125	400,64063	28,582813
6	5868,2188	20,134375	387,53281	27,709375
7	5903,0625	20,392188	375,16719	29,554688
8	5913,1094	20,221875	387,45781	28,175
9	5921,4219	20,084375	380,87344	26,392188
10	5916,4844	20,039063	374,24375	28,978125
Average	5901,9859	20,239844	387,25094	28,437188

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05							
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	59019,859	5901,9859	1723,561	15512,049	6,7033922	5888,3908	5915,581
CLMLC	10	202,39844	20,239844	0,0296122	0,26651	6,7033922	6,6447342	33,834953
CBMLC	10	3872,5094	387,25094	72,453735	652,08362	6,7033922	373,65583	400,84605
HOMER	10	284,37188	28,437188	1,3743413	12,369071	6,7033922	14,842078	42,032297

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	249423314	3	83141105	185023,35	2,012E-75	2,8662656	136,02329	0,9999279
Within Groups	16176,768	36	449,35467					
Total	249439491	39	6395884,4					

Tabela 96 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Bibtex

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	5901,9859	10	15512,049			
CLMLC	20,239844	10	0,26651			
CBMLC	387,25094	10	652,08362			
HOMER	28,437188	10	12,369071			
		40	16176,768	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	5881,7461	6,7033922	877,42831	5856,2129	5907,2793	3,331E-15	25,533221	277,46719
LIFT	CBMLC	5514,735	6,7033922	822,67825	5489,2018	5540,2682	3,331E-15	25,533221	260,15371
LIFT	HOMER	5873,5488	6,7033922	876,20544	5848,0155	5899,082	3,331E-15	25,533221	277,08049
CLMLC	CBMLC	367,01109	6,7033922	54,750055	341,47787	392,54431	7,994E-15	25,533221	17,313488
CLMLC	HOMER	8,1973438	6,7033922	1,222865	17,335877	33,730565	0,8228797	25,533221	0,3867039
CBMLC	HOMER	358,81375	6,7033922	53,52719	333,28053	384,34697	7,994E-15	25,533221	16,926784

Tabela 97 – Teste ANOVA para o tempo de CPU - conjunto de dados Birds

Parameter	tCPU			
Dataset	Birds			
Run	LIFT	CLMLC	CBMLC	HOMER
1	2,3296875	0,121875	0,459375	0,1859375
2	2,178125	0,1203125	0,446875	0,190625
3	2,2703125	0,1265625	0,45625	0,1796875
4	2,2390625	0,1109375	0,453125	0,1734375
5	2,21875	0,1203125	0,446875	0,1953125
6	2,2375	0,1296875	0,4296875	0,1671875
7	2,1828125	0,128125	0,425	0,1828125
8	2,196875	0,1328125	0,4640625	0,165625
9	2,140625	0,1296875	0,446875	0,1640625
10	2,159375	0,1203125	0,434375	0,1890625
Average	2,2153125	0,1240625	0,44625	0,179375

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
<i>Group</i>	<i>Count</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>	<i>SS</i>	<i>Std Err</i>	<i>Lower</i>	<i>Upper</i>	
LIFT	10	22,153125	2,2153125	0,0031846	0,0286611	0,0093815	2,1962859	2,2343391	
CLMLC	10	1,240625	0,1240625	4,243E-05	0,0003818	0,0093815	0,1050359	0,1430891	
CBMLC	10	4,4625	0,44625	0,0001672	0,0015049	0,0093815	0,4272234	0,4652766	
HOMER	10	1,79375	0,179375	0,0001263	0,0011367	0,0093815	0,1603484	0,1984016	

ANOVA

<i>Sources</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P value</i>	<i>F crit</i>	<i>RMSSE</i>	<i>Omega Sq</i>	
Between Groups	29,565092	3	9,8550306	11197,283	1,666E-53	2,8662656	33,462342	0,9988105	
Within Groups	0,0316846	36	0,0008801						
Total	29,596776	39	0,7588917						

Tabela 98 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Birds

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	2,2153125	10	0,0286611			
CLMLC	0,1240625	10	0,0003818			
CBMLC	0,44625	10	0,0015049			
HOMER	0,179375	10	0,0011367			
		40	0,0316846	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	2,09125	0,0093815	222,91192	2,0555158	2,1269842	7,438E-15	0,0357342	0,0986532
LIFT	CBMLC	1,7690625	0,0093815	188,56909	1,7333283	1,8047967	7,772E-15	0,0357342	0,0834543
LIFT	HOMER	2,0359375	0,0093815	217,01601	2,0002033	2,0716717	7,438E-15	0,0357342	0,0960439
CLMLC	CBMLC	0,3221875	0,0093815	34,342826	0,2864533	0,3579217	7,994E-15	0,0357342	0,015199
CLMLC	HOMER	0,0553125	0,0093815	5,895907	0,0195783	0,0910467	0,0010122	0,0357342	0,0026093
CBMLC	HOMER	0,266875	0,0093815	28,446919	0,2311408	0,3026092	7,994E-15	0,0357342	0,0125896

Tabela 99 – Teste ANOVA para o tempo de CPU - conjunto de dados Corel5k

Parameter	tCPU			
Dataset	Corel5k			
Run	LIFT	CLMLC	CBMLC	HOMER
1	2492,2969	8,2734375	110,83594	125,99531
2	2482,8906	8,246875	108,60156	129,67031
3	2483,7656	8,36875	108,19219	133,00156
4	2489,4063	8,2265625	112,66094	143,21406
5	2468,0313	8,1640625	111,925	138,46406
6	2460,9063	7,975	112,14688	134,25781
7	2471,7656	7,74375	111,50156	130,15
8	2453,75	7,74375	110,97813	143,45469
9	2460,9375	8,0015625	112,02344	138,38281
10	2468,3438	8,0640625	110,64219	145,41406
Average	2473,2094	8,0807813	110,95078	136,20047

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	24732,094	2473,2094	173,89302	1565,0372	2,3448811	2468,4537	2477,965	
CLMLC	10	80,807813	8,0807813	0,0467492	0,4207424	2,3448811	3,3251419	12,836421	
CBMLC	10	1109,5078	110,95078	2,2217959	19,996163	2,3448811	106,19514	115,70642	
HOMER	10	1362,0047	136,20047	43,777135	393,99422	2,3448811	131,44483	140,95611	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	42865924	3	14288641	259865,88	4,45E-78	2,8662656	161,20356	0,9999487
Within Groups	1979,4483	36	54,984676					
Total	42867904	39	1099177					

Tabela 100 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Corel5k

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	2473,2094	10	1565,0372			
CLMLC	8,0807813	10	0,4207424			
CBMLC	110,95078	10	19,996163			
HOMER	136,20047	10	393,99422			
		40	1979,4483	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	2465,1286	2,3448811	1051,2808	2456,1969	2474,0602	3,109E-15	8,9316523	116,29069
LIFT	CBMLC	2362,2586	2,3448811	1007,4108	2353,3269	2371,1902	3,22E-15	8,9316523	111,43787
LIFT	HOMER	2337,0089	2,3448811	996,6428	2328,0773	2345,9406	3,331E-15	8,9316523	110,24673
CLMLC	CBMLC	102,87	2,3448811	43,870027	93,938348	111,80165	7,994E-15	8,9316523	4,8528192
CLMLC	HOMER	128,11969	2,3448811	54,63803	119,18804	137,05134	7,994E-15	8,9316523	6,0439552
CBMLC	HOMER	25,249688	2,3448811	10,768003	16,318035	34,18134	3,085E-08	8,9316523	1,1911361

Tabela 101 – Teste ANOVA para o tempo de CPU - conjunto de dados Emotions

Parameter	tCPU			
Dataset	Emotions			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,8	0,0484375	0,0375	0,046875
2	0,8359375	0,046875	0,021875	0,0578125
3	0,771875	0,05625	0,0203125	0,0546875
4	0,7859375	0,040625	0,03125	0,059375
5	0,834375	0,0296875	0,0265625	0,0546875
6	0,81875	0,04375	0,0203125	0,05
7	0,784375	0,04375	0,025	0,0578125
8	0,8125	0,040625	0,0203125	0,05
9	0,7921875	0,0390625	0,0296875	0,034375
10	0,834375	0,0375	0,025	0,0375
Average	0,8070313	0,0426563	0,0257813	0,0503125

ANOVA: Single Factor

DESCRIPTION	Alpha		0,05					
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	8,0703125	0,8070313	0,0005514	0,0049622	0,0042061	0,7985009	0,8155616
CLMLC	10	0,4265625	0,0426563	5,048E-05	0,0004543	0,0042061	0,0341259	0,0511866
CBMLC	10	0,2578125	0,0257813	3,215E-05	0,0002893	0,0042061	0,0172509	0,0343116
HOMER	10	0,503125	0,0503125	7,368E-05	0,0006631	0,0042061	0,0417821	0,0588429

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	4,4204728	3	1,4734909	8328,8642	3,405E-51	2,8662656	28,859772	0,9984015
Within Groups	0,0063689	36	0,0001769					
Total	4,4268417	39	0,1135088					

Tabela 102 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Emotions

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,8070313	10	0,0049622			
CLMLC	0,0426563	10	0,0004543			
CBMLC	0,0257813	10	0,0002893			
HOMER	0,0503125	10	0,0006631			
		40	0,0063689	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,764375	0,0042061	181,72958	0,7483539	0,7803961	7,772E-15	0,0160211	0,0360588
LIFT	CBMLC	0,78125	0,0042061	185,7416	0,7652289	0,7972711	7,772E-15	0,0160211	0,0368549
LIFT	HOMER	0,7567188	0,0042061	179,90932	0,7406977	0,7727398	7,772E-15	0,0160211	0,0356977
CLMLC	CBMLC	0,016875	0,0042061	4,0120186	0,0008539	0,0328961	0,0357436	0,0160211	0,0007961
CLMLC	HOMER	0,0076563	0,0042061	1,8202677	0,0083648	0,0236773	0,5768825	0,0160211	0,0003612
CBMLC	HOMER	0,0245313	0,0042061	5,8322863	0,0085102	0,0405523	0,0011528	0,0160211	0,0011572

Tabela 103 – Teste ANOVA para o tempo de CPU - conjunto de dados Genbase

Parameter	tCPU			
Dataset	Genbase			
Run	LIFT	CLMLC	CBMLC	HOMER
1	7,846875	0,353125	6,609375	5,09375
2	7,8421875	0,33125	6,6015625	4,609375
3	7,778125	0,321875	6,3859375	4,515625
4	7,9578125	0,334375	6,6328125	4,796875
5	7,934375	0,328125	6,6234375	4,796875
6	7,909375	0,321875	6,496875	4,421875
7	7,9234375	0,3203125	6,5296875	4,90625
8	7,84375	0,35	6,4171875	5,109375
9	7,8265625	0,33125	6,6671875	4,890625
10	7,8296875	0,3421875	6,759375	4,4375
Average	7,8692188	0,3334375	6,5723438	4,7578125

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	78,692188	7,8692188	0,0033561	0,0302053	0,0448932	7,7781711	7,9602664	
CLMLC	10	3,334375	0,3334375	0,0001352	0,0012168	0,0448932	0,2423899	0,4244851	
CBMLC	10	65,723438	6,5723438	0,0132004	0,118804	0,0448932	6,4812961	6,6633914	
HOMER	10	47,578125	4,7578125	0,0639242	0,5753174	0,0448932	4,6667649	4,8488601	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	324,85575	3	108,28525	5372,8953	8,966E-48	2,8662656	23,179507	0,9975241
Within Groups	0,7255435	36	0,020154					
Total	325,5813	39	8,3482384					

Tabela 104 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Genbase

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	7,8692188	10	0,0302053			
CLMLC	0,3334375	10	0,0012168			
CBMLC	6,5723438	10	0,118804			
HOMER	4,7578125	10	0,5753174			
		40	0,7255435	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	7,5357813	0,0448932	167,86023	7,3647831	7,7067794	7,772E-15	0,1709982	0,3554951
LIFT	CBMLC	1,296875	0,0448932	28,888012	1,1258768	1,4678732	7,994E-15	0,1709982	0,0611792
LIFT	HOMER	3,1114063	0,0448932	69,306865	2,9404081	3,2824044	7,994E-15	0,1709982	0,1467784
CLMLC	CBMLC	6,2389063	0,0448932	138,97222	6,0679081	6,4099044	7,994E-15	0,1709982	0,294316
CLMLC	HOMER	4,424375	0,0448932	98,553367	4,2533768	4,5953732	7,994E-15	0,1709982	0,2087167
CBMLC	HOMER	1,8145313	0,0448932	40,418853	1,6435331	1,9855294	7,994E-15	0,1709982	0,0855992

Tabela 105 – Teste ANOVA para o tempo de CPU - conjunto de dados Medical

Parameter	tCPU			
Dataset	Medical			
Run	LIFT	CLMLC	CBMLC	HOMER
1	33,56875	1,0578125	21,359375	1,178125
2	33,351563	1,1015625	21,5875	1,15625
3	33,901563	1,0578125	21,49375	1,1328125
4	34,045313	1,0546875	20,453125	1,1546875
5	34,3875	1,059375	20,94375	1,1265625
6	33,982813	1,0328125	22,065625	1,125
7	33,851563	1,0625	20,970313	1,10625
8	34,36875	1,090625	19,698438	1,2421875
9	34,379688	1,0203125	21,25625	1,240625
10	34,3875	1,0609375	20,6	1,1609375
Average	34,0225	1,0598438	21,042813	1,1623438

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05							
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	340,225	34,0225	0,135101	1,2159092	0,1213197	33,776452	34,268548
CLMLC	10	10,598438	1,0598438	0,0005583	0,0050247	0,1213197	0,813796	1,3058915
CBMLC	10	210,42813	21,042813	0,4509078	4,0581699	0,1213197	20,796765	21,28886
HOMER	10	11,623438	1,1623438	0,0021715	0,0195432	0,1213197	0,916296	1,4083915

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	7823,4036	3	2607,8012	17717,89	4,339E-57	2,8662656	42,092624	0,999248
Within Groups	5,298647	36	0,1471846					
Total	7828,7023	39	200,73596					

Tabela 106 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Medical

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	34,0225	10	1,2159092			
CLMLC	1,0598438	10	0,0050247			
CBMLC	21,042813	10	4,0581699			
HOMER	1,1623438	10	0,0195432			
		40	5,298647	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	32,962656	0,1213197	271,70083	32,50055	33,424763	6,883E-15	0,4621066	1,5549899
LIFT	CBMLC	12,979688	0,1213197	106,98749	12,517581	13,441794	7,994E-15	0,4621066	0,6123075
LIFT	HOMER	32,860156	0,1213197	270,85595	32,39805	33,322263	6,883E-15	0,4621066	1,5501545
CLMLC	CBMLC	19,982969	0,1213197	164,71334	19,520862	20,445075	7,772E-15	0,4621066	0,9426824
CLMLC	HOMER	0,1025	0,1213197	0,8448753	0,3596066	0,5646066	0,9321889	0,4621066	0,0048354
CBMLC	HOMER	19,880469	0,1213197	163,86846	19,418362	20,342575	7,772E-15	0,4621066	0,937847

Tabela 107 – Teste ANOVA para o tempo de CPU - conjunto de dados Yeast

Parameter	tCPU			
Dataset	Yeast			
Run	LIFT	CLMLC	CBMLC	HOMER
1	43,445313	0,596875	0,225	0,246875
2	42,178125	0,6	0,21875	0,225
3	42,720313	0,596875	0,203125	0,2265625
4	43,345313	0,6125	0,203125	0,2375
5	42,814063	0,5890625	0,209375	0,246875
6	42,071875	0,596875	0,209375	0,253125
7	41,0625	0,6203125	0,1859375	0,2640625
8	41,201563	0,6140625	0,209375	0,234375
9	40,995313	0,6203125	0,2015625	0,2390625
10	41,034375	0,6109375	0,196875	0,259375
Average	42,086875	0,6057813	0,20625	0,2432813

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	420,86875	42,086875	0,9444446	8,500001	0,153693	41,775171	42,398579	
CLMLC	10	6,0578125	0,6057813	0,0001237	0,0011135	0,153693	0,2940774	0,9174851	
CBMLC	10	2,0625	0,20625	0,0001199	0,0010791	0,153693	0,1054538	0,5179538	
HOMER	10	2,4328125	0,2432813	0,0001731	0,0015579	0,153693	0,0684226	0,5549851	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	13064,617	3	4354,8722	18436,028	2,123E-57	2,8662656	42,937196	0,9992773
Within Groups	8,5037515	36	0,2362153					
Total	13073,12	39	335,20821					

Tabela 108 – Teste Tukey HSD para o tempo de CPU - conjunto de dados Yeast

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	42,086875	10	8,500001			
CLMLC	0,6057813	10	0,0011135			
CBMLC	0,20625	10	0,0010791			
HOMER	0,2432813	10	0,0015579			
		40	8,5037515	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	41,481094	0,153693	269,89583	40,895677	42,06651	6,883E-15	0,5854166	1,9568411
LIFT	CBMLC	41,880625	0,153693	272,49537	41,295208	42,466042	6,883E-15	0,5854166	1,9756887
LIFT	HOMER	41,843594	0,153693	272,25443	41,258177	42,42901	6,883E-15	0,5854166	1,9739418
CLMLC	CBMLC	0,3995313	0,153693	2,5995413	0,1858853	0,9849478	0,2726586	0,5854166	0,0188476
CLMLC	HOMER	0,3625	0,153693	2,3585983	0,2229166	0,9479166	0,3551338	0,5854166	0,0171007
CBMLC	HOMER	0,0370313	0,153693	0,240943	0,5483853	0,6224478	0,9982083	0,5854166	0,0017469

B.2.3 Exact Match Ratio

Tabela 109 – Teste ANOVA para o exact match ratio - conjunto de dados Bibtex

Parameter	EM			
Dataset	Bibtex			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,1378378	0,1720056	0,1544242	0,0969528
2	0,1407307	0,1722744	0,1533407	0,0969528
3	0,1299053	0,1755285	0,1438776	0,0928989
4	0,1163735	0,1729523	0,1557741	0,094385
5	0,1123139	0,1706519	0,1471256	0,0972212
6	0,1675676	0,1736293	0,1555056	0,093709
7	0,1378378	0,1741764	0,1625418	0,0979026
8	0,1459459	0,1757922	0,1560529	0,1001966
9	0,1271989	0,1763316	0,1598398	0,0958726
10	0,1351351	0,1737639	0,1545593	0,0969557
Average	0,1350847	0,1737106	0,1543042	0,0963047

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	1,3508467	0,1350847	0,000242	0,0021778	0,0026436	0,1297233	0,1404461	
CLMLC	10	1,737106	0,1737106	3,293E-06	2,963E-05	0,0026436	0,1683492	0,179072	
CBMLC	10	1,5430415	0,1543042	2,96E-05	0,0002664	0,0026436	0,1489428	0,1596656	
HOMER	10	0,9630472	0,0963047	4,668E-06	4,201E-05	0,0026436	0,0909433	0,1016661	

ANOVA								
Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0327436	3	0,0109145	156,18011	1,084E-20	2,8662656	3,951963	0,9208767
Within Groups	0,0025158	36	6,988E-05					
Total	0,0352595	39	0,0009041					

Tabela 110 – Teste Tukey HSD para o exact match ratio - conjunto de dados Bibtex

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,1350847	10	0,0021778			
CLMLC	0,1737106	10	2,963E-05			
CBMLC	0,1543042	10	0,0002664			
HOMER	0,0963047	10	4,201E-05			
		40	0,0025158	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0386259	0,0026436	14,611309	0,0285566	0,0486953	1,536E-11	0,0100693	4,6205017
LIFT	CBMLC	0,0192195	0,0026436	7,2702936	0,0091502	0,0292888	5,59E-05	0,0100693	2,2990687
LIFT	HOMER	0,0387799	0,0026436	14,669566	0,0287106	0,0488493	1,379E-11	0,0100693	4,6389242
CLMLC	CBMLC	0,0194064	0,0026436	7,3410157	0,0093371	0,0294758	4,802E-05	0,0100693	2,321433
CLMLC	HOMER	0,0774059	0,0026436	29,280876	0,0673365	0,0874752	7,994E-15	0,0100693	9,2594259
CBMLC	HOMER	0,0579994	0,0026436	21,93986	0,0479301	0,0680688	7,994E-15	0,0100693	6,9379929

Tabela 111 – Teste ANOVA para o exact match ratio - conjunto de dados Birds

Parameter	EM			
Dataset	Birds			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,4438942	0,5178365	0,5057212	0,4683654
2	0,4423317	0,5116346	0,508726	0,4604567
3	0,4438942	0,5208654	0,5087981	0,4883894
4	0,4469952	0,5147837	0,508774	0,4682692
5	0,4454567	0,5286779	0,5026202	0,483726
6	0,4454327	0,525601	0,5071635	0,4760096
7	0,4454327	0,5270913	0,5010096	0,480649
8	0,4469952	0,5178606	0,4978846	0,4790865
9	0,4438942	0,4976683	0,5026202	0,4603846
10	0,4454567	0,525601	0,502476	0,4914663
Average	0,4449784	0,518762	0,5045793	0,4756803

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
<i>Group</i>	<i>Count</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>	<i>SS</i>	<i>Std Err</i>	<i>Lower</i>	<i>Upper</i>	
LIFT	10	4,4497837	0,4449784	2,171E-06	1,954E-05	0,0023657	0,4401806	0,4497761	
CLMLC	10	5,1876202	0,518762	8,655E-05	0,0007789	0,0023657	0,5139642	0,5235598	
CBMLC	10	5,0457933	0,5045793	1,447E-05	0,0001302	0,0023657	0,4997816	0,5093771	
HOMER	10	4,7568029	0,4756803	0,0001207	0,001086	0,0023657	0,4708825	0,4804781	

ANOVA

<i>Sources</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P value</i>	<i>F crit</i>	<i>RMSSE</i>	<i>Omega Sq</i>	
Between Groups	0,0320781	3	0,0106927	191,06634	3,666E-22	2,8662656	4,3711136	0,9344476	
Within Groups	0,0020147	36	5,596E-05						
Total	0,0340928	39	0,0008742						

Tabela 112 – Teste Tukey HSD para o exact match ratio - conjunto de dados Birds

TUKEY HSD/KRAMER		alpha		0,05	
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>
LIFT	0,4449784	10	1,954E-05		
CLMLC	0,518762	10	0,0007789		
CBMLC	0,5045793	10	0,0001302		
HOMER	0,4756803	10	0,001086		
		40	0,0020147	36	3,809

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0737837	0,0023657	31,1895	0,0647729	0,0827944	7,994E-15	0,0090108	8,8261288
LIFT	CBMLC	0,059601	0,0023657	25,194255	0,0505902	0,0686117	7,994E-15	0,0090108	7,1295705
LIFT	HOMER	0,0307019	0,0023657	12,978181	0,0216911	0,0397127	3,476E-10	0,0090108	3,6726174
CLMLC	CBMLC	0,0141827	0,0023657	5,995245	0,0051719	0,0231935	0,0008254	0,0090108	1,6965583
CLMLC	HOMER	0,0430817	0,0023657	18,211319	0,0340709	0,0520925	3,664E-14	0,0090108	5,1535115
CBMLC	HOMER	0,028899	0,0023657	12,216074	0,0198883	0,0379098	1,58E-09	0,0090108	3,4569532

Tabela 113 – Teste ANOVA para o exact match ratio - conjunto de dados Corel5k

Parameter	EM			
Dataset	Corel5k			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,006	0,03	0,0096	0,0014
2	0,004	0,0288	0,0128	0,0006
3	0,008	0,0278	0,01	0,001
4	0,004	0,031	0,012	0,0014
5	0,01	0,0296	0,0102	0,0012
6	0,004	0,0282	0,0116	0,0012
7	0,01	0,0288	0,0112	0,0008
8	0	0,0286	0,012	0,001
9	0,004	0,0282	0,0098	0,0012
10	0,012	0,03	0,0112	0,001
Average	0,0062	0,0291	0,01104	0,00108

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05							
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	0,062	0,0062	1,373E-05	0,0001236	0,0006327	0,0049169	0,0074831
CLMLC	10	0,291	0,0291	1,024E-06	9,22E-06	0,0006327	0,0278169	0,0303831
CBMLC	10	0,1104	0,01104	1,189E-06	1,07E-05	0,0006327	0,0097569	0,0123231
HOMER	10	0,0108	0,00108	6,4E-08	5,76E-07	0,0006327	0,0002031	0,0023631

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0044613	3	0,0014871	371,52025	3,977E-27	2,8662656	6,0952461	0,9652645
Within Groups	0,0001441	36	4,003E-06					
Total	0,0046054	39	0,0001181					

Tabela 114 – Teste Tukey HSD para o exact match ratio - conjunto de dados Corel5k

TUKEY HSD/KRAMER			alpha			0,05			
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>				
LIFT	0,0062	10	0,0001236						
CLMLC	0,0291	10	9,22E-06						
CBMLC	0,01104	10	1,07E-05						
HOMER	0,00108	10	5,76E-07						
		40	0,0001441	36	3,809				

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0229	0,0006327	36,195514	0,0204901	0,0253099	7,994E-15	0,0024099	2,7393378
LIFT	CBMLC	0,00484	0,0006327	7,6500561	0,0024301	0,0072499	2,467E-05	0,0024099	0,5789692
LIFT	HOMER	0,00512	0,0006327	8,0926214	0,0027101	0,0075299	9,47E-06	0,0024099	0,6124633
CLMLC	CBMLC	0,01806	0,0006327	28,545457	0,0156501	0,0204699	7,994E-15	0,0024099	2,1603686
CLMLC	HOMER	0,02802	0,0006327	44,288135	0,0256101	0,0304299	7,994E-15	0,0024099	3,3518011
CBMLC	HOMER	0,00996	0,0006327	15,742677	0,0075501	0,0123699	1,964E-12	0,0024099	1,1914325

Tabela 115 – Teste ANOVA para o exact match ratio - conjunto de dados Emotions

Parameter	EM			
Dataset	Emotions			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,0067232	0,2629379	0,2598588	0,3037571
2	0,0050565	0,2681356	0,2378249	0,2800565
3	0,0067514	0,2290678	0,254774	0,2814972
4	0,0118079	0,2578531	0,244661	0,2934463
5	0,0084181	0,2829944	0,2463559	0,284887
6	0,0084463	0,2561299	0,2377684	0,2664124
7	0,0067514	0,2831356	0,2481073	0,2783616
8	0,0084181	0,2614124	0,2513559	0,2883051
9	0,010113	0,2527684	0,2563559	0,2648305
10	0,0134746	0,2376554	0,2632768	0,2949435
Average	0,008596	0,259209	0,2500339	0,2836497

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	0,0859605	0,008596	6,594E-06	5,934E-05	0,0036329	0,0012281	0,015964	
CLMLC	10	2,5920904	0,259209	0,0002958	0,0026625	0,0036329	0,2518411	0,266577	
CBMLC	10	2,500339	0,2500339	7,574E-05	0,0006817	0,0036329	0,2426659	0,2574019	
HOMER	10	2,8364972	0,2836497	0,0001498	0,0013478	0,0036329	0,2762818	0,2910177	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,496413	3	0,165471	1253,7287	1,863E-36	2,8662656	11,197003	0,9894687
Within Groups	0,0047514	36	0,000132					
Total	0,5011644	39	0,0128504					

Tabela 116 – Teste Tukey HSD para o exact match ratio - conjunto de dados Emotions

TUKEY HSD/KRAMER		alpha		0,05	
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>
LIFT	0,008596	10	5,934E-05		
CLMLC	0,259209	10	0,0026625		
CBMLC	0,2500339	10	0,0006817		
HOMER	0,2836497	10	0,0013478		
		40	0,0047514	36	3,809

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,250613	0,0036329	68,983376	0,2367751	0,2644509	7,994E-15	0,0138379	29,978762
LIFT	CBMLC	0,2414379	0,0036329	66,45784	0,2276	0,2552758	7,994E-15	0,0138379	28,881215
LIFT	HOMER	0,2750537	0,0036329	75,710882	0,2612158	0,2888916	7,994E-15	0,0138379	32,902398
CLMLC	CBMLC	0,0091751	0,0036329	2,5255363	0,0046628	0,023013	0,2965837	0,0138379	1,0975463
CLMLC	HOMER	0,0244407	0,0036329	6,7275062	0,0106028	0,0382786	0,0001783	0,0138379	2,9236363
CBMLC	HOMER	0,0336158	0,0036329	9,2530425	0,0197779	0,0474537	7,711E-07	0,0138379	4,0211827

Tabela 117 – Teste ANOVA para o exact match ratio - conjunto de dados Genbase

Parameter	EM			
Dataset	Genbase			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,9470828	0,9637268	0,9818634	0,9652872
2	0,9470828	0,9692497	0,9788557	0,9667797
3	0,9470828	0,9667571	0,9803709	0,9683175
4	0,9485979	0,9652646	0,9818634	0,9698327
5	0,9440751	0,9682723	0,981886	0,9668024
6	0,9425599	0,9667571	0,9833786	0,9698327
7	0,9470828	0,9667797	0,9818634	0,96981
8	0,9485979	0,9683175	0,9803483	0,9713252
9	0,9440525	0,9668024	0,981886	0,9683175
10	0,9455676	0,9682949	0,9848937	0,9682949
Average	0,9461782	0,9670222	0,9817209	0,96846

ANOVA: Single Factor

DESCRIPTION	Alpha		0,05						
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	9,461782	0,9461782	4,154E-06	3,739E-05	0,0005671	0,9450282	0,9473282	
CLMLC	10	9,6702221	0,9670222	2,669E-06	2,402E-05	0,0005671	0,9658722	0,9681723	
CBMLC	10	9,8172094	0,9817209	2,761E-06	2,485E-05	0,0005671	0,9805709	0,982871	
HOMER	10	9,6845997	0,96846	3,278E-06	2,951E-05	0,0005671	0,9673099	0,96961	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0064705	3	0,0021568	670,75376	1,241E-31	2,8662656	8,1899558	0,9804808
Within Groups	0,0001158	36	3,216E-06					
Total	0,0065863	39	0,0001689					

Tabela 118 – Teste Tukey HSD para o exact match ratio - conjunto de dados Genbase

TUKEY HSD/KRAMER		alpha		0,05	
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>
LIFT	0,9461782	10	3,739E-05		
CLMLC	0,9670222	10	2,402E-05		
CBMLC	0,9817209	10	2,485E-05		
HOMER	0,96846	10	2,951E-05		
		40	0,0001158	36	3,809

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,020844	0,0005671	36,758152	0,0186841	0,0230039	7,994E-15	0,0021599	2,4933963
LIFT	CBMLC	0,0355427	0,0005671	62,679189	0,0333828	0,0377027	7,994E-15	0,0021599	4,2516844
LIFT	HOMER	0,0222818	0,0005671	39,293634	0,0201218	0,0244417	7,994E-15	0,0021599	2,6653844
CLMLC	CBMLC	0,0146987	0,0005671	25,921037	0,0125388	0,0168587	7,994E-15	0,0021599	1,758288
CLMLC	HOMER	0,0014378	0,0005671	2,5354823	0,0007222	0,0035977	0,2932938	0,0021599	0,171988
CBMLC	HOMER	0,013261	0,0005671	23,385555	0,011101	0,0154209	7,994E-15	0,0021599	1,5863

Tabela 119 – Teste ANOVA para o exact match ratio - conjunto de dados Medical

Parameter	EM			
Dataset	Medical			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,4989901	0,7167263	0,6021671	0,5644435
2	0,4836524	0,6942457	0,5847675	0,5756364
3	0,493846	0,6901746	0,5889017	0,5776036
4	0,4867031	0,6953293	0,584778	0,5551862
5	0,4908058	0,7086051	0,5766148	0,5797076
6	0,4887965	0,6952556	0,5776036	0,5643909
7	0,4887124	0,7045024	0,5919419	0,5674732
8	0,4979592	0,6861246	0,5899222	0,574595
9	0,4785399	0,6963076	0,5909741	0,578603
10	0,4999369	0,7004313	0,5878603	0,5602882
Average	0,4907942	0,6987703	0,5875531	0,5697928

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05							
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	4,9079424	0,4907942	4,863E-05	0,0004376	0,0025408	0,4856412	0,4959473
CLMLC	10	6,9877025	0,6987703	8,214E-05	0,0007393	0,0025408	0,6936172	0,7039233
CBMLC	10	5,8755312	0,5875531	5,402E-05	0,0004862	0,0025408	0,5824001	0,5927062
HOMER	10	5,6979276	0,5697928	7,345E-05	0,000661	0,0025408	0,5646397	0,5749458

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,2204424	3	0,0734808	1138,2088	1,043E-35	2,8662656	10,668687	0,9884113
Within Groups	0,0023241	36	6,456E-05					
Total	0,2227665	39	0,005712					

Tabela 120 – Teste Tukey HSD para o exact match ratio - conjunto de dados Medical

TUKEY HSD/KRAMER			alpha			0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>			
LIFT	0,4907942	10	0,0004376					
CLMLC	0,6987703	10	0,0007393					
CBMLC	0,5875531	10	0,0004862					
HOMER	0,5697928	10	0,000661					
		40	0,0023241	36	3,809			

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,207976	0,0025408	81,853513	0,198298	0,217654	7,994E-15	0,009678	24,878452
LIFT	CBMLC	0,0967589	0,0025408	38,081579	0,0870809	0,1064369	7,994E-15	0,009678	11,574466
LIFT	HOMER	0,0789985	0,0025408	31,091599	0,0693205	0,0886766	7,994E-15	0,009678	9,449941
CLMLC	CBMLC	0,1112171	0,0025408	43,771934	0,1015391	0,1208952	7,994E-15	0,009678	13,303986
CLMLC	HOMER	0,1289775	0,0025408	50,761913	0,1192995	0,1386555	7,994E-15	0,009678	15,428511
CBMLC	HOMER	0,0177604	0,0025408	6,9899791	0,0080823	0,0274384	0,0001019	0,009678	2,1245253

Tabela 121 – Teste ANOVA para o exact match ratio - conjunto de dados Yeast

Parameter	EM			
Dataset	Yeast			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,193625	0,2031292	0,1960821	0,1464713
2	0,1948579	0,203148	0,194016	0,151454
3	0,1931998	0,1915589	0,1981688	0,140285
4	0,1948544	0,1886664	0,1894877	0,1440005
5	0,1961027	0,1948647	0,1927763	0,1443881
6	0,1915538	0,1828641	0,1919413	0,1439937
7	0,1932033	0,1857635	0,196936	0,1481465
8	0,1940348	0,2023233	0,1923751	0,1419379
9	0,1923751	0,1890744	0,1911166	0,1423528
10	0,1923734	0,1940263	0,2035373	0,1477093
Average	0,193618	0,1935419	0,1946437	0,1450739

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	1,9361802	0,193618	1,902E-06	1,712E-05	0,0014518	0,1906736	0,1965624	
CLMLC	10	1,9354189	0,1935419	5,394E-05	0,0004855	0,0014518	0,1905975	0,1964863	
CBMLC	10	1,946437	0,1946437	1,713E-05	0,0001542	0,0014518	0,1916993	0,1975881	
HOMER	10	1,450739	0,1450739	1,133E-05	0,000102	0,0014518	0,1421295	0,1480183	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0179128	3	0,0059709	283,28287	4,379E-25	2,8662656	5,3224324	0,9548965
Within Groups	0,0007588	36	2,108E-05					
Total	0,0186716	39	0,0004788					

Tabela 122 – Teste Tukey HSD para o exact match ratio - conjunto de dados Yeast

TUKEY HSD/KRAMER		alpha		0,05	
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>
LIFT	0,193618	10	1,712E-05		
CLMLC	0,1935419	10	0,0004855		
CBMLC	0,1946437	10	0,0001542		
HOMER	0,1450739	10	0,000102		
		40	0,0007588	36	3,809

Q TEST										
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>	
LIFT	CLMLC	7,613E-05	0,0014518	0,0524372	0,0054538	-	0,0056061	0,9999813	0,00553	0,0091067
LIFT	CBMLC	0,0010257	0,0014518	0,7064854	0,0045043	-	0,0065556	0,9586442	0,00553	0,1226942
LIFT	HOMER	0,0485441	0,0014518	33,436886	0,0430142	-	0,0540741	7,994E-15	0,00553	5,8069316
CLMLC	CBMLC	0,0011018	0,0014518	0,7589227	0,0044281	-	0,0066318	0,9495057	0,00553	0,1318009
CLMLC	HOMER	0,048468	0,0014518	33,384449	0,042938	-	0,0539979	7,994E-15	0,00553	5,7978249
CBMLC	HOMER	0,0495698	0,0014518	34,143371	0,0440398	-	0,0550998	7,994E-15	0,00553	5,9296258

B.2.4 Hamming Loss

Tabela 123 – Teste ANOVA para o hamming loss - conjunto de dados Bibtex

Parameter	HL			
Dataset	Bibtex			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,0137855	0,0137378	0,0162857	0,0223211
2	0,0142212	0,0136103	0,0161638	0,0216092
3	0,0149446	0,0137855	0,0167008	0,0216618
4	0,0140254	0,0137889	0,0165181	0,022492
5	0,014885	0,0138281	0,0170033	0,0212664
6	0,0131056	0,0137752	0,0168279	0,0217868
7	0,013658	0,0137242	0,0153367	0,0222064
8	0,0144994	0,0137191	0,0164602	0,021989
9	0,0139829	0,0137013	0,0156681	0,0219596
10	0,0149329	0,013663	0,0161374	0,0215734
Average	0,014204	0,0137333	0,0163102	0,0218866

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	0,1420403	0,014204	3,768E-07	3,392E-06	0,0001405	0,0139192	0,0144889	
CLMLC	10	0,1373333	0,0137333	4,218E-09	3,796E-08	0,0001405	0,0134485	0,0140182	
CBMLC	10	0,1631021	0,0163102	2,643E-07	2,379E-06	0,0001405	0,0160254	0,0165951	
HOMER	10	0,2188657	0,0218866	1,437E-07	1,294E-06	0,0001405	0,0216017	0,0221714	

ANOVA								
Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0004197	3	0,0001399	709,22083	4,63E-32	2,8662656	8,421525	0,9815214
Within Groups	7,102E-06	36	1,973E-07					
Total	0,0004268	39	1,094E-05					

Tabela 124 – Teste Tukey HSD para o hamming loss - conjunto de dados Bibtex

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,014204	10	3,392E-06			
CLMLC	0,0137333	10	3,796E-08			
CBMLC	0,0163102	10	2,379E-06			
HOMER	0,0218866	10	1,294E-06			
		40	7,102E-06	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0004707	0,0001405	3,3512569	-6,429E-05	0,0010057	0,1014055	0,000535	1,0597605
LIFT	CBMLC	0,0021062	0,0001405	14,995649	0,0015712	0,0026412	7,562E-12	0,000535	4,7420407
LIFT	HOMER	0,0076825	0,0001405	54,698251	0,0071476	0,0082175	7,994E-15	0,000535	17,297106
CLMLC	CBMLC	0,0025769	0,0001405	18,346906	0,0020419	0,0031119	3,109E-14	0,000535	5,8018012
CLMLC	HOMER	0,0081532	0,0001405	58,049508	0,0076183	0,0086882	7,994E-15	0,000535	18,356866
CBMLC	HOMER	0,0055764	0,0001405	39,702602	0,0050414	0,0061113	7,994E-15	0,000535	12,555065

Tabela 125 – Teste ANOVA para o hamming loss - conjunto de dados Birds

Parameter	HL			
Dataset	Birds			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,054294	0,046503	0,0592801	0,0604466
2	0,0543763	0,0473128	0,0586285	0,0623317
3	0,054294	0,0478112	0,0586247	0,0579314
4	0,0542131	0,0456048	0,0583831	0,0601506
5	0,054294	0,0454365	0,059356	0,0596736
6	0,0541321	0,0441346	0,0583009	0,0584337
7	0,0541321	0,045124	0,0600114	0,0577771
8	0,0540486	0,0475746	0,0595218	0,0570369
9	0,0543763	0,0489626	0,0600089	0,0612082
10	0,0541308	0,0465056	0,0590385	0,0574608
Average	0,0542291	0,046497	0,0591154	0,0592451

ANOVA: Single Factor

DESCRIPTION		Alpha		0,05				
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	0,5422912	0,0542291	1,303E-08	1,173E-07	0,0003768	0,053465	0,0549932
CLMLC	10	0,4649696	0,046497	2,113E-06	1,902E-05	0,0003768	0,0457329	0,0472611
CBMLC	10	0,5911538	0,0591154	3,927E-07	3,535E-06	0,0003768	0,0583513	0,0598795
HOMER	10	0,5924507	0,0592451	3,159E-06	2,843E-05	0,0003768	0,058481	0,0600092

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0010764	3	0,0003588	252,78963	3,108E-24	2,8662656	5,0278189	0,9497089
Within Groups	5,11E-05	36	1,419E-06					
Total	0,0011275	39	2,891E-05					

Tabela 126 – Teste Tukey HSD para o hamming loss - conjunto de dados Birds

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,0542291	10	1,173E-07			
CLMLC	0,046497	10	1,902E-05			
CBMLC	0,0591154	10	3,535E-06			
HOMER	0,0592451	10	2,843E-05			
		40	5,11E-05	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0077322	0,0003768	20,523231	0,0062971	0,0091672	8,66E-15	0,001435	17,408807
LIFT	CBMLC	0,0048863	0,0003768	12,969446	0,0034512	0,0063213	3,536E-10	0,001435	11,001318
LIFT	HOMER	0,0050159	0,0003768	13,313655	0,0035809	0,006451	1,806E-10	0,001435	11,293292
CLMLC	CBMLC	0,0126184	0,0003768	33,492677	0,0111834	0,0140535	7,994E-15	0,001435	28,410125
CLMLC	HOMER	0,0127481	0,0003768	33,836885	0,0113131	0,0141831	7,994E-15	0,001435	28,702099
CBMLC	HOMER	0,0001297	0,0003768	0,3442086	0,0013054	0,0015647	0,9948395	0,001435	0,2919746

Tabela 127 – Teste ANOVA para o hamming loss - conjunto de dados Corel5k

Parameter	HL			
Dataset	Corel5k			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,0096096	0,0135626	0,0118316	0,0171759
2	0,0098021	0,0136524	0,0115749	0,0164824
3	0,0095668	0,0133834	0,0119364	0,0165503
4	0,0097487	0,0136604	0,0119444	0,0166545
5	0,0094599	0,0137524	0,011838	0,0162941
6	0,0098984	0,0136781	0,0116091	0,0159428
7	0,0096257	0,0138738	0,0117947	0,0165535
8	0,0096898	0,0134139	0,0120465	0,0168497
9	0,0094332	0,0135364	0,0121679	0,016584
10	0,0095027	0,013538	0,011938	0,0167
Average	0,0096337	0,0136051	0,0118681	0,0165787

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	0,0963369	0,0096337	2,306E-08	2,075E-07	6,777E-05	0,0094962	0,0097711	
CLMLC	10	0,1360513	0,0136051	2,236E-08	2,013E-07	6,777E-05	0,0134677	0,0137426	
CBMLC	10	0,1186813	0,0118681	3,309E-08	2,978E-07	6,777E-05	0,0117307	0,0120056	
HOMER	10	0,1657872	0,0165787	1,052E-07	9,468E-07	6,777E-05	0,0164413	0,0167162	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0002576	3	8,587E-05	1869,832	1,481E-39	2,8662656	13,67418	0,992916
Within Groups	1,653E-06	36	4,593E-08					
Total	0,0002593	39	6,648E-06					

Tabela 128 – Teste Tukey HSD para o hamming loss - conjunto de dados Corel5k

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,0096337	10	2,075E-07			
CLMLC	0,0136051	10	2,013E-07			
CBMLC	0,0118681	10	2,978E-07			
HOMER	0,0165787	10	9,468E-07			
		40	1,653E-06	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0039714	6,777E-05	58,603237	0,0037133	0,0042296	7,994E-15	0,0002581	8,9416271
LIFT	CBMLC	0,0022344	6,777E-05	32,971719	0,0019763	0,0024926	7,994E-15	0,0002581	5,030794
LIFT	HOMER	0,006945	6,777E-05	102,48188	0,0066869	0,0072032	7,994E-15	0,0002581	15,63659
CLMLC	CBMLC	0,001737	6,777E-05	25,631518	0,0014789	0,0019951	7,994E-15	0,0002581	3,9108331
CLMLC	HOMER	0,0029736	6,777E-05	43,878647	0,0027155	0,0032317	7,994E-15	0,0002581	6,6949629
CBMLC	HOMER	0,0047106	6,777E-05	69,510165	0,0044525	0,0049687	7,994E-15	0,0002581	10,605796

Tabela 129 – Teste ANOVA para o hamming loss - conjunto de dados Emotions

Parameter	HL			
Dataset	Emotions			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,3097128	0,2454237	0,2318315	0,2068362
2	0,3102872	0,2420339	0,2380461	0,2099765
3	0,3113983	0,2472128	0,2340819	0,2131591
4	0,3102542	0,2426554	0,2472175	0,2100094
5	0,3116855	0,2368173	0,2371657	0,2066573
6	0,3108192	0,2390819	0,2430932	0,2167467
7	0,308291	0,2387288	0,2354002	0,2102872
8	0,3094162	0,2445433	0,2329379	0,2080556
9	0,3088418	0,2480132	0,238032	0,2108145
10	0,3099765	0,2530744	0,2348588	0,2066714
Average	0,3100683	0,2437585	0,2372665	0,2099214

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	3,1006827	0,3100683	1,139E-06	1,025E-05	0,0012066	0,3076212	0,3125153	
CLMLC	10	2,4375847	0,2437585	2,442E-05	0,0002197	0,0012066	0,2413114	0,2462055	
CBMLC	10	2,3726648	0,2372665	2,243E-05	0,0002019	0,0012066	0,2348194	0,2397135	
HOMER	10	2,0992137	0,2099214	1,025E-05	9,223E-05	0,0012066	0,2074743	0,2123684	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0541533	3	0,0180511	1239,9026	2,27E-36	2,8662656	11,135091	0,9893524
Within Groups	0,0005241	36	1,456E-05					
Total	0,0546775	39	0,001402					

Tabela 130 – Teste Tukey HSD para o hamming loss - conjunto de dados Emotions

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,3100683	10	1,025E-05			
CLMLC	0,2437585	10	0,0002197			
CBMLC	0,2372665	10	0,0002019			
HOMER	0,2099214	10	9,223E-05			
		40	0,0005241	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0663098	0,0012066	54,956545	0,0617139	0,0709057	7,994E-15	0,0045959	149,29518
LIFT	CBMLC	0,0728018	0,0012066	60,337012	0,0682059	0,0773977	7,994E-15	0,0045959	163,91178
LIFT	HOMER	0,1001469	0,0012066	83,000217	0,095551	0,1047428	7,994E-15	0,0045959	225,47874
CLMLC	CBMLC	0,006492	0,0012066	5,3804675	0,0018961	0,0110879	0,0028601	0,0045959	14,616601
CLMLC	HOMER	0,0338371	0,0012066	28,043672	0,0292412	0,038433	7,994E-15	0,0045959	76,183559
CBMLC	HOMER	0,0273451	0,0012066	22,663205	0,0227492	0,031941	7,994E-15	0,0045959	61,566959

Tabela 131 – Teste ANOVA para o hamming loss - conjunto de dados Genbase

Parameter	HL			
Dataset	Genbase			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,0025772	0,0021282	0,0008401	0,0016224
2	0,0026333	0,0018477	0,0008954	0,001511
3	0,0026333	0,0020143	0,0008392	0,001454
4	0,0025772	0,0021249	0,0007278	0,0013418
5	0,0026886	0,0019004	0,0006709	0,001454
6	0,0027447	0,0019565	0,0006717	0,0013979
7	0,0026333	0,0020688	0,0007278	0,0013426
8	0,0026333	0,0018988	0,0008962	0,0012865
9	0,0026333	0,002011	0,000727	0,001454
10	0,0026894	0,0018452	0,0006156	0,001511
Average	0,0026444	0,0019796	0,0007612	0,0014375

ANOVA: Single Factor

DESCRIPTION		Alpha 0,05							
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	0,0264435	0,0026444	2,632E-09	2,369E-08	2,903E-05	0,0025855	0,0027032	
CLMLC	10	0,0197958	0,0019796	1,135E-08	1,022E-07	2,903E-05	0,0019207	0,0020385	
CBMLC	10	0,0076118	0,0007612	9,921E-09	8,929E-08	2,903E-05	0,0007023	0,0008201	
HOMER	10	0,0143751	0,0014375	9,814E-09	8,832E-08	2,903E-05	0,0013786	0,0014964	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	1,92E-05	3	6,4E-06	759,27453	1,384E-32	2,8662656	8,7136361	0,9827201
Within Groups	3,035E-07	36	8,43E-09					
Total	1,95E-05	39	5,001E-07					

Tabela 132 – Teste Tukey HSD para o hamming loss - conjunto de dados Genbase

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,0026444	10	2,369E-08			
CLMLC	0,0019796	10	1,022E-07			
CBMLC	0,0007612	10	8,929E-08			
HOMER	0,0014375	10	8,832E-08			
		40	3,035E-07	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0006648	2,903E-05	22,896468	0,0005542	0,0007754	7,994E-15	0,0001106	1,496725
LIFT	CBMLC	0,0018832	2,903E-05	64,861305	0,0017726	0,0019938	7,994E-15	0,0001106	4,239935
LIFT	HOMER	0,0012068	2,903E-05	41,566739	0,0010963	0,0013174	7,994E-15	0,0001106	2,7171866
CLMLC	CBMLC	0,0012184	2,903E-05	41,964838	0,0011078	0,001329	7,994E-15	0,0001106	2,7432101
CLMLC	HOMER	0,0005421	2,903E-05	18,670271	0,0004315	0,0006527	2,154E-14	0,0001106	1,2204616
CBMLC	HOMER	0,0006763	2,903E-05	23,294567	0,0005657	0,0007869	7,994E-15	0,0001106	1,5227484

Tabela 133 – Teste ANOVA para o hamming loss - conjunto de dados Medical

Parameter	HL			
Dataset	Medical			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,0148839	0,0107041	0,0130467	0,0171118
2	0,0152701	0,0112493	0,0134995	0,0162714
3	0,0147479	0,0114985	0,0137034	0,0163658
4	0,014976	0,0112923	0,0140031	0,0167929
5	0,0149749	0,010907	0,0140931	0,0161117
6	0,0149513	0,011318	0,0132969	0,0170666
7	0,014839	0,0111118	0,0136169	0,0169984
8	0,0147488	0,0118379	0,0132501	0,0163188
9	0,0154298	0,0112944	0,0132279	0,016708
10	0,0150897	0,0109973	0,0132524	0,0164759
Average	0,0149911	0,0112211	0,013499	0,0166221

ANOVA: Single Factor

DESCRIPTION		Alpha		0,05				
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	0,1499114	0,0149911	4,828E-08	4,346E-07	0,0001003	0,0147877	0,0151946
CLMLC	10	0,1122105	0,0112211	1,007E-07	9,06E-07	0,0001003	0,0110176	0,0114245
CBMLC	10	0,1349901	0,013499	1,227E-07	1,104E-06	0,0001003	0,0132956	0,0137025
HOMER	10	0,1662213	0,0166221	1,309E-07	1,178E-06	0,0001003	0,0164187	0,0168256

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,000158	3	5,268E-05	523,44859	9,832E-30	2,8662656	7,2349747	0,9751143
Within Groups	3,623E-06	36	1,006E-07					
Total	0,0001617	39	4,145E-06					

Tabela 134 – Teste Tukey HSD para o hamming loss - conjunto de dados Medical

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,0149911	10	4,346E-07			
CLMLC	0,0112211	10	9,06E-07			
CBMLC	0,013499	10	1,104E-06			
HOMER	0,0166221	10	1,178E-06			
		40	3,623E-06	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0037701	0,0001003	37,581105	0,003388	0,0041522	7,994E-15	0,0003821	8,4882755
LIFT	CBMLC	0,0014921	0,0001003	14,873937	0,00111	0,0018742	9,453E-12	0,0003821	3,3595093
LIFT	HOMER	0,001631	0,0001003	16,258122	0,0012489	0,0020131	7,94E-13	0,0003821	3,6721491
CLMLC	CBMLC	0,002278	0,0001003	22,707169	0,0018958	0,0026601	7,994E-15	0,0003821	5,1287663
CLMLC	HOMER	0,0054011	0,0001003	53,839228	0,005019	0,0057832	7,994E-15	0,0003821	12,160425
CBMLC	HOMER	0,0031231	0,0001003	31,132059	0,002741	0,0035052	7,994E-15	0,0003821	7,0316583

Tabela 135 – Teste ANOVA para o hamming loss - conjunto de dados Yeast

Parameter	HL			
Dataset	Yeast			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,1906684	0,2281805	0,2185176	0,2327571
2	0,1913495	0,2291592	0,217656	0,2318595
3	0,1912019	0,232847	0,2168014	0,2395986
4	0,1911122	0,234329	0,2200784	0,2385177
5	0,1906397	0,2361334	0,2187271	0,2340286
6	0,1923823	0,2313969	0,2201978	0,2364452
7	0,1919692	0,2366006	0,2206996	0,2385632
8	0,1913194	0,2264324	0,2194023	0,233153
9	0,1912599	0,2322002	0,2193988	0,2352003
10	0,191052	0,2297783	0,215528	0,2334445
Average	0,1912954	0,2317057	0,2187007	0,2353568

ANOVA: Single Factor

DESCRIPTION		Alpha		0,05				
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	1,9129544	0,1912954	2,85E-07	2,565E-06	0,0007411	0,1897923	0,1927985
CLMLC	10	2,3170574	0,2317057	1,139E-05	0,0001025	0,0007411	0,2302026	0,2332088
CBMLC	10	2,1870071	0,2187007	2,657E-06	2,391E-05	0,0007411	0,2171976	0,2202038
HOMER	10	2,3535676	0,2353568	7,638E-06	6,874E-05	0,0007411	0,2338537	0,2368599

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0119633	3	0,0039878	725,98763	3,062E-32	2,8662656	8,5204908	0,981941
Within Groups	0,0001977	36	5,493E-06					
Total	0,0121611	39	0,0003118					

Tabela 136 – Teste Tukey HSD para o hamming loss - conjunto de dados Yeast

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,1912954	10	2,565E-06			
CLMLC	0,2317057	10	0,0001025			
CBMLC	0,2187007	10	2,391E-05			
HOMER	0,2353568	10	6,874E-05			
		40	0,0001977	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0404103	0,0007411	54,524479	0,0375873	0,0432333	7,994E-15	0,002823	90,98298
LIFT	CBMLC	0,0274053	0,0007411	36,977159	0,0245823	0,0302283	7,994E-15	0,002823	61,702416
LIFT	HOMER	0,0440613	0,0007411	59,450703	0,0412383	0,0468843	7,994E-15	0,002823	99,203188
CLMLC	CBMLC	0,013005	0,0007411	17,547321	0,010182	0,015828	9,415E-14	0,002823	29,280564
CLMLC	HOMER	0,003651	0,0007411	4,9262238	0,000828	0,006474	0,0069071	0,002823	8,2202073
CBMLC	HOMER	0,0166561	0,0007411	22,473544	0,013833	0,0194791	7,994E-15	0,002823	37,500772

B.2.5 F1-macro

Tabela 137 – Teste ANOVA para o f1-macro - conjunto de dados Bibtex

Parameter	F1-macro			
Dataset	Bibtex			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,179502	0,2120616	0,248722	0,141043
2	0,1779358	0,2071454	0,2512767	0,1380363
3	0,2036997	0,2091536	0,2469711	0,1405498
4	0,1909025	0,2106679	0,2473561	0,1363961
5	0,1634171	0,2090456	0,2375737	0,1414128
6	0,1920145	0,2062986	0,2437393	0,1382592
7	0,1623825	0,2112263	0,2363288	0,1394917
8	0,1824167	0,2061993	0,247883	0,1396608
9	0,1724568	0,2096902	0,2421831	0,1395106
10	0,1799568	0,2124576	0,2420255	0,1432945
Average	0,1804684	0,2093946	0,2444059	0,1397655

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	1,8046843	0,1804684	0,0001639	0,0014753	0,0022188	0,1759685	0,1849683	
CLMLC	10	2,0939461	0,2093946	5,177E-06	4,659E-05	0,0022188	0,2048947	0,2138945	
CBMLC	10	2,4440591	0,2444059	2,402E-05	0,0002162	0,0022188	0,239906	0,2489058	
HOMER	10	1,397655	0,1397655	3,798E-06	3,418E-05	0,0022188	0,1352656	0,1442654	

ANOVA								
Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0590127	3	0,0196709	399,57342	1,117E-27	2,8662656	6,321182	0,9676302
Within Groups	0,0017723	36	4,923E-05					
Total	0,060785	39	0,0015586					

Tabela 138 – Teste Tukey HSD para o f1-macro - conjunto de dados Bibtex

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,1804684	10	0,0014753			
CLMLC	0,2093946	10	4,659E-05			
CBMLC	0,2444059	10	0,0002162			
HOMER	0,1397655	10	3,418E-05			
		40	0,0017723	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0289262	0,0022188	13,036986	0,0204748	0,0373775	3,098E-10	0,0084513	4,122657
LIFT	CBMLC	0,0639375	0,0022188	28,816534	0,0554862	0,0723888	7,994E-15	0,0084513	9,112588
LIFT	HOMER	0,0407029	0,0022188	18,344757	0,0322516	0,0491543	3,109E-14	0,0084513	5,8011216
CLMLC	CBMLC	0,0350113	0,0022188	15,779547	0,02656	0,0434626	1,84E-12	0,0084513	4,989931
CLMLC	HOMER	0,0696291	0,0022188	31,381744	0,0611778	0,0780804	7,994E-15	0,0084513	9,9237786
CBMLC	HOMER	0,1046404	0,0022188	47,161291	0,0961891	0,1130917	7,994E-15	0,0084513	14,91371

Tabela 139 – Teste ANOVA para o f1-macro - conjunto de dados Birds

Parameter	F1-macro			
Dataset	Birds			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0	0,3318635	0,3025914	0,2287405
2	0	0,3173483	0,3037657	0,2427702
3	0	0,3182516	0,305105	0,2409572
4	0	0,3349604	0,3106078	0,2235861
5	0	0,3472433	0,2979995	0,2366501
6	0	0,3153153	0,3131082	0,2401483
7	0	0,330037	0,2999736	0,2345176
8	0	0,3234527	0,3041021	0,23074
9	0	0,3161541	0,3053095	0,2297269
10	0	0,3233116	0,3008206	0,2248906
Average	0	0,3257938	0,3043383	0,2332728

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05							
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	0	0	0	0	0,0020733	0,0042048	0,0042048
CLMLC	10	3,2579378	0,3257938	0,0001046	0,0009415	0,0020733	0,321589	0,3299985
CBMLC	10	3,0433833	0,3043383	2,142E-05	0,0001928	0,0020733	0,3001336	0,3085431
HOMER	10	2,3327276	0,2332728	4,59E-05	0,0004131	0,0020733	0,229068	0,2374775

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,6681259	3	0,2227086	5181,2135	1,722E-47	2,8662656	22,762279	0,9974327
Within Groups	0,0015474	36	4,298E-05					
Total	0,6696734	39	0,0171711					

Tabela 140 – Teste Tukey HSD para o f1-macro - conjunto de dados Birds

TUKEY HSD/KRAMER			alpha		0,05	
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0	10	0			
CLMLC	0,3257938	10	0,0009415			
CBMLC	0,3043383	10	0,0001928			
HOMER	0,2332728	10	0,0004131			
		40	0,0015474	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,3257938	0,0020733	157,14117	0,3178967	0,3336908	7,772E-15	0,007897	46,433244
LIFT	CBMLC	0,3043383	0,0020733	146,7925	0,2964413	0,3122354	7,994E-15	0,007897	43,375339
LIFT	HOMER	0,2332728	0,0020733	112,51521	0,2253757	0,2411698	7,994E-15	0,007897	33,246832
CLMLC	CBMLC	0,0214555	0,0020733	10,348677	0,0135584	0,0293525	7,448E-08	0,007897	3,0579042
CLMLC	HOMER	0,092521	0,0020733	44,625964	0,084624	0,100418	7,994E-15	0,007897	13,186412
CBMLC	HOMER	0,0710656	0,0020733	34,277287	0,0631685	0,0789626	7,994E-15	0,007897	10,128508

Tabela 141 – Teste ANOVA para o f1-macro - conjunto de dados Corel5k

Parameter	F1-macro			
Dataset	Corel5k			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,0370601	0,0370791	0,0256257	0,0335599
2	0,0236878	0,0371236	0,0254845	0,0339201
3	0,0302733	0,0369838	0,0242279	0,0347116
4	0,0299298	0,036347	0,0251392	0,0350414
5	0,034343	0,0362133	0,0238281	0,0326507
6	0,0229547	0,0354861	0,0248863	0,0336343
7	0,0229675	0,0348176	0,0244516	0,03307
8	0,0275746	0,0358521	0,0254633	0,0343217
9	0,0361207	0,0372455	0,0243743	0,0332812
10	0,038309	0,0363691	0,0237021	0,033859
Average	0,0303221	0,0363517	0,0247183	0,033805

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	0,3032205	0,0303221	3,545E-05	0,0003191	0,0009633	0,0283684	0,0322757	
CLMLC	10	0,363517	0,0363517	6,328E-07	5,695E-06	0,0009633	0,034398	0,0383054	
CBMLC	10	0,2471829	0,0247183	4,918E-07	4,426E-06	0,0009633	0,0227646	0,026672	
HOMER	10	0,3380498	0,033805	5,411E-07	4,87E-06	0,0009633	0,0318513	0,0357586	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0007607	3	0,0002536	27,325411	2,168E-09	2,8662656	1,65304	0,6637984
Within Groups	0,0003341	36	9,28E-06					
Total	0,0010948	39	2,807E-05					

Tabela 142 – Teste Tukey HSD para o f1-macro - conjunto de dados Corel5k

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,0303221	10	0,0003191			
CLMLC	0,0363517	10	5,695E-06			
CBMLC	0,0247183	10	4,426E-06			
HOMER	0,033805	10	4,87E-06			
		40	0,0003341	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0060297	0,0009633	6,2593556	0,0023604	0,0096989	0,0004775	0,0036692	0,8593664
LIFT	CBMLC	0,0056038	0,0009633	5,8172456	0,0019345	0,009273	0,0011887	0,0036692	0,7986677
LIFT	HOMER	0,0034829	0,0009633	3,6156083	0,0001863	0,0071521	0,0679793	0,0036692	0,4963981
CLMLC	CBMLC	0,0116334	0,0009633	12,076601	0,0079642	0,0153026	2,093E-09	0,0036692	1,6580341
CLMLC	HOMER	0,0025467	0,0009633	2,6437473	0,0011225	0,0062159	0,2589841	0,0036692	0,3629683
CBMLC	HOMER	0,0090867	0,0009633	9,4328539	0,0054175	0,0127559	5,239E-07	0,0036692	1,2950658

Tabela 143 – Teste ANOVA para o f1-macro - conjunto de dados Emotions

Parameter	F1-macro			
Dataset	Emotions			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,0750957	0,5865555	0,6173797	0,6619583
2	0,0709892	0,5908351	0,6062172	0,6597767
3	0,0663388	0,5806784	0,6138257	0,656011
4	0,0680019	0,5871122	0,593019	0,6564243
5	0,0659634	0,5985555	0,6066545	0,6704246
6	0,0695947	0,5926981	0,5985633	0,6508514
7	0,0724137	0,5933748	0,6118817	0,6628002
8	0,0745506	0,5836177	0,6116429	0,6624381
9	0,083301	0,5863042	0,6070572	0,6626814
10	0,0813438	0,5741953	0,6114787	0,6613563
Average	0,0727593	0,5873927	0,607772	0,6604722

ANOVA: Single Factor

DESCRIPTION					Alpha	0,05			
<i>Group</i>	<i>Count</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>	<i>SS</i>	<i>Std Err</i>	<i>Lower</i>	<i>Upper</i>	
LIFT	10	0,7275929	0,0727593	3,52E-05	0,0003168	0,0020272	0,068648	0,0768706	
CLMLC	10	5,8739269	0,5873927	4,843E-05	0,0004359	0,0020272	0,5832814	0,591504	
CBMLC	10	6,0777198	0,607772	5,347E-05	0,0004812	0,0020272	0,6036607	0,6118833	
HOMER	10	6,6047222	0,6604722	2,728E-05	0,0002455	0,0020272	0,6563609	0,6645835	

ANOVA

<i>Sources</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P value</i>	<i>F crit</i>	<i>RMSSE</i>	<i>Omega Sq</i>
Between Groups	2,2625646	3	0,7541882	18352,739	2,303E-57	2,8662656	42,840097	0,999274
Within Groups	0,0014794	36	4,109E-05					
Total	2,264044	39	0,0580524					

Tabela 144 – Teste Tukey HSD para o f1-macro - conjunto de dados Emotions

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,0727593	10	0,0003168			
CLMLC	0,5873927	10	0,0004359			
CBMLC	0,607772	10	0,0004812			
HOMER	0,6604722	10	0,0002455			
		40	0,0014794	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,5146334	0,0020272	253,86834	0,5069119	0,5223549	7,105E-15	0,0077215	73,347313
LIFT	CBMLC	0,5350127	0,0020272	263,92144	0,5272912	0,5427342	7,105E-15	0,0077215	76,251839
LIFT	HOMER	0,5877129	0,0020272	289,91843	0,5799914	0,5954344	6,661E-15	0,0077215	83,762857
CLMLC	CBMLC	0,0203793	0,0020272	10,053094	0,0126578	0,0281008	1,393E-07	0,0077215	2,9045268
CLMLC	HOMER	0,0730795	0,0020272	36,050086	0,065358	0,080801	7,994E-15	0,0077215	10,415544
CBMLC	HOMER	0,0527002	0,0020272	25,996992	0,0449788	0,0604217	7,994E-15	0,0077215	7,5110172

Tabela 145 – Teste ANOVA para o f1-macro - conjunto de dados Genbase

Parameter	F1-macro			
Dataset	Genbase			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,5830031	0,6538669	0,6582734	0,6119803
2	0,5809043	0,648472	0,6485164	0,611381
3	0,5821487	0,6710344	0,659508	0,6150237
4	0,5837889	0,6603423	0,6633755	0,616172
5	0,5831278	0,6430819	0,6616028	0,6148731
6	0,5802055	0,6418473	0,6645658	0,6134618
7	0,5825563	0,6603423	0,6571583	0,6130027
8	0,5825544	0,637975	0,6521006	0,6158427
9	0,583254	0,6377751	0,6620613	0,614344
10	0,5837686	0,6386733	0,6645658	0,6131447
Average	0,5825311	0,6493411	0,6591728	0,6139226

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	5,8253114	0,5825311	1,38E-06	1,242E-05	0,0020412	0,5783914	0,5866709	
CLMLC	10	6,4934106	0,6493411	0,000134	0,001206	0,0020412	0,6452013	0,6534808	
CBMLC	10	6,5917279	0,6591728	2,874E-05	0,0002586	0,0020412	0,655033	0,6633126	
HOMER	10	6,1392259	0,6139226	2,548E-06	2,294E-05	0,0020412	0,6097828	0,6180624	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0368041	3	0,012268	294,4382	2,248E-25	2,8662656	5,426216	0,9565366
Within Groups	0,0015	36	4,167E-05					
Total	0,0383041	39	0,0009822					

Tabela 146 – Teste Tukey HSD para o f1-macro - conjunto de dados Genbase

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,5825311	10	1,242E-05			
CLMLC	0,6493411	10	0,001206			
CBMLC	0,6591728	10	0,0002586			
HOMER	0,6139226	10	2,294E-05			
		40	0,0015	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0668099	0,0020412	32,73034	0,0590349	0,0745849	7,994E-15	0,007775	9,5219772
LIFT	CBMLC	0,0766416	0,0020412	37,546929	0,0688666	0,0844167	7,994E-15	0,007775	10,923229
LIFT	HOMER	0,0313914	0,0020412	15,378748	0,0236164	0,0391665	3,769E-12	0,007775	4,4740166
CLMLC	CBMLC	0,0098317	0,0020412	4,8165889	0,0020567	0,0176067	0,0084967	0,007775	1,4012519
CLMLC	HOMER	0,0354185	0,0020412	17,351592	0,0276434	0,0431935	1,281E-13	0,007775	5,0479606
CBMLC	HOMER	0,0452502	0,0020412	22,168181	0,0374752	0,0530252	7,994E-15	0,007775	6,4492125

Tabela 147 – Teste ANOVA para o f1-macro - conjunto de dados Medical

Parameter	F1-macro			
Dataset	Medical			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,2413419	0,3380337	0,3341595	0,2241661
2	0,2371455	0,3263581	0,3284796	0,2220586
3	0,2365563	0,3308191	0,3361733	0,23999
4	0,2430218	0,3361008	0,3311212	0,2345804
5	0,2397783	0,3341534	0,3350544	0,2259162
6	0,2390197	0,3280274	0,3263727	0,2181739
7	0,2416109	0,3330257	0,3255734	0,2225269
8	0,2439771	0,3312717	0,3365282	0,2225226
9	0,2306994	0,3323749	0,3331884	0,21526
10	0,2399184	0,3311388	0,332628	0,2293676
Average	0,2393069	0,3321304	0,3319279	0,2254562

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	2,3930691	0,2393069	1,467E-05	0,0001321	0,0015621	0,2361389	0,2424749	
CLMLC	10	3,3213036	0,3321304	1,213E-05	0,0001092	0,0015621	0,3289624	0,3352984	
CBMLC	10	3,3192786	0,3319279	1,553E-05	0,0001398	0,0015621	0,3287599	0,3350959	
HOMER	10	2,2545623	0,2254562	5,526E-05	0,0004974	0,0015621	0,2222882	0,2286242	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,1002557	3	0,0334186	1369,5957	3,851E-37	2,8662656	11,702973	0,9903517
Within Groups	0,0008784	36	2,44E-05					
Total	0,1011341	39	0,0025932					

Tabela 148 – Teste Tukey HSD para o f1-macro - conjunto de dados Medical

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,2393069	10	0,0001321			
CLMLC	0,3321304	10	0,0001092			
CBMLC	0,3319279	10	0,0001398			
HOMER	0,2254562	10	0,0004974			
		40	0,0008784	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0928234	0,0015621	59,423732	0,0868736	0,0987733	7,994E-15	0,0059499	13,229515
LIFT	CBMLC	0,0926209	0,0015621	59,294094	0,0866711	0,0985708	7,994E-15	0,0059499	13,200654
LIFT	HOMER	0,0138507	0,0015621	8,8669314	0,0079008	0,0198006	1,773E-06	0,0059499	1,9740464
CLMLC	CBMLC	0,0002025	0,0015621	0,129638	0,0057474	0,0061524	0,9997186	0,0059499	0,0288613
CLMLC	HOMER	0,1066741	0,0015621	68,290663	0,1007242	0,112624	7,994E-15	0,0059499	15,203562
CBMLC	HOMER	0,1064716	0,0015621	68,161026	0,1005217	0,1124215	7,994E-15	0,0059499	15,1747

Tabela 149 – Teste ANOVA para o f1-macro - conjunto de dados Yeast

Parameter	F1-macro			
Dataset	Yeast			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,4051801	0,4061026	0,3822071	0,421455
2	0,4034771	0,4042236	0,3837695	0,4248668
3	0,4059954	0,3963689	0,3892786	0,4233469
4	0,407023	0,3995298	0,382154	0,4232775
5	0,4062798	0,3985261	0,379052	0,421711
6	0,4024007	0,4092173	0,3784584	0,4299202
7	0,4017729	0,3868303	0,3730207	0,4189004
8	0,4052188	0,4081767	0,3845436	0,410096
9	0,4068365	0,4009197	0,3718595	0,4219567
10	0,4062951	0,4030727	0,3885584	0,415849
Average	0,4050479	0,4012968	0,3812902	0,4211379

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	4,0504794	0,4050479	3,48E-06	3,132E-05	0,001652	0,4016975	0,4083984	
CLMLC	10	4,0129676	0,4012968	4,327E-05	0,0003894	0,001652	0,3979463	0,4046472	
CBMLC	10	3,8129019	0,3812902	3,392E-05	0,0003053	0,001652	0,3779397	0,3846407	
HOMER	10	4,2113795	0,4211379	2,85E-05	0,0002565	0,001652	0,4177875	0,4244884	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0080479	3	0,0026826	98,294345	2,112E-17	2,8662656	3,1351929	0,8794756
Within Groups	0,0009825	36	2,729E-05					
Total	0,0090304	39	0,0002315					

Tabela 150 – Teste Tukey HSD para o f1-macro - conjunto de dados Yeast

TUKEY HSD/KRAMER			alpha			0,05			
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>				
LIFT	0,4050479	10	3,132E-05						
CLMLC	0,4012968	10	0,0003894						
CBMLC	0,3812902	10	0,0003053						
HOMER	0,4211379	10	0,0002565						
		40	0,0009825	36	3,809				

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0037512	0,001652	2,2706532	0,0025414	0,0100437	0,3883595	0,0062926	0,534631
LIFT	CBMLC	0,0237578	0,001652	14,380974	0,0174652	0,0300503	2,36E-11	0,0062926	3,3860364
LIFT	HOMER	0,01609	0,001652	9,7395551	0,0097974	0,0223826	2,717E-07	0,0062926	2,2932026
CLMLC	CBMLC	0,0200066	0,001652	12,110321	0,013714	0,0262991	1,955E-09	0,0062926	2,8514054
CLMLC	HOMER	0,0198412	0,001652	12,010208	0,0135486	0,0261338	2,393E-09	0,0062926	2,8278336
CBMLC	HOMER	0,0398478	0,001652	24,120529	0,0335552	0,0461403	7,994E-15	0,0062926	5,679239

B.2.6 F1-micro

Tabela 151 – Teste ANOVA para o f1-micro - conjunto de dados Bibtex

Parameter	F1-micro			
Dataset	Bibtex			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,2765388	0,4056491	0,3744065	0,2723002
2	0,2596367	0,4051464	0,3780432	0,2750879
3	0,2683333	0,4049979	0,3691593	0,2776778
4	0,2688554	0,4036823	0,3723822	0,2753221
5	0,2385721	0,4037512	0,3596386	0,2823415
6	0,3128342	0,4045874	0,3653615	0,2752266
7	0,2515137	0,4054038	0,3767832	0,2737887
8	0,2569686	0,4031458	0,3733035	0,2786635
9	0,2595764	0,4054057	0,3764648	0,2749507
10	0,2429987	0,4046091	0,3730285	0,2801886
Average	0,2635828	0,4046379	0,3718571	0,2765548

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05							
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
LIFT	10	2,6358279	0,2635828	0,0004349	0,0039142	0,0034559	0,2565738	0,2705917
CLMLC	10	4,0463787	0,4046379	7,24E-07	6,516E-06	0,0034559	0,3976289	0,4116468
CBMLC	10	3,7185712	0,3718571	3,253E-05	0,0002928	0,0034559	0,3648482	0,3788661
HOMER	10	2,7655477	0,2765548	9,573E-06	8,616E-05	0,0034559	0,2695458	0,2835637

ANOVA								
Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,1458763	3	0,0486254	407,13023	8,053E-28	2,8662656	6,3806758	0,9682134
Within Groups	0,0042996	36	0,0001194					
Total	0,150176	39	0,0038507					

Tabela 152 – Teste Tukey HSD para o f1-macro - conjunto de dados Bibtex

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,2635828	10	0,0039142			
CLMLC	0,4046379	10	6,516E-06			
CBMLC	0,3718571	10	0,0002928			
HOMER	0,2765548	10	8,616E-05			
		40	0,0042996	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,1410551	0,0034559	40,815359	0,1278914	0,1542187	7,994E-15	0,0131636	12,90695
LIFT	CBMLC	0,1082743	0,0034559	31,33	0,0951107	0,121438	7,994E-15	0,0131636	9,9074158
LIFT	HOMER	0,012972	0,0034559	3,7535413	0,0001917	0,0261356	0,0546587	0,0131636	1,186974
CLMLC	CBMLC	0,0327807	0,0034559	9,485359	0,0196171	0,0459444	4,681E-07	0,0131636	2,9995339
CLMLC	HOMER	0,1280831	0,0034559	37,061817	0,1149195	0,1412467	7,994E-15	0,0131636	11,719976
CBMLC	HOMER	0,0953023	0,0034559	27,576458	0,0821387	0,108466	7,994E-15	0,0131636	8,7204418

Tabela 153 – Teste ANOVA para o f1-micro - conjunto de dados Birds

Parameter	F1-micro			
Dataset	Birds			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0	0,4928724	0,4382703	0,3890612
2	0	0,4754131	0,4445684	0,3735969
3	0	0,4712872	0,4411732	0,4034842
4	0	0,4896076	0,4520128	0,3790824
5	0	0,493712	0,4351934	0,3943994
6	0	0,4924271	0,4475654	0,4106729
7	0	0,4932205	0,4317109	0,4068832
8	0	0,4733296	0,4355327	0,4025037
9	0	0,4536143	0,4337468	0,3808898
10	0	0,4769064	0,4331419	0,3934766
Average	0	0,481239	0,4392916	0,393405

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	0	0	0	0	0,0031024	-0,006292	0,006292	
CLMLC	10	4,8123902	0,481239	0,0001786	0,0016071	0,0031024	0,474947	0,487531	
CBMLC	10	4,3929159	0,4392916	4,663E-05	0,0004197	0,0031024	0,4329996	0,4455836	
HOMER	10	3,9340502	0,393405	0,0001598	0,0014382	0,0031024	0,387113	0,399697	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	1,477289	3	0,4924297	5116,1758	2,16E-47	2,8662656	22,618965	0,9974002
Within Groups	0,003465	36	9,625E-05					
Total	1,4807539	39	0,037968					

Tabela 154 – Teste Tukey HSD para o f1-macro - conjunto de dados Birds

TUKEY HSD/KRAMER			alpha		0,05	
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0	10	0			
CLMLC	0,481239	10	0,0016071			
CBMLC	0,4392916	10	0,0004197			
HOMER	0,393405	10	0,0014382			
		40	0,003465	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,481239	0,0031024	155,11774	0,4694219	0,4930561	7,883E-15	0,0118171	44,03477
LIFT	CBMLC	0,4392916	0,0031024	141,59683	0,4274745	0,4511087	7,994E-15	0,0118171	40,196458
LIFT	HOMER	0,393405	0,0031024	126,80621	0,3815879	0,4052221	7,994E-15	0,0118171	35,997704
CLMLC	CBMLC	0,0419474	0,0031024	13,520911	0,0301303	0,0537645	1,209E-10	0,0118171	3,8383116
CLMLC	HOMER	0,087834	0,0031024	28,311525	0,0760169	0,0996511	7,994E-15	0,0118171	8,037066
CBMLC	HOMER	0,0458866	0,0031024	14,790614	0,0340695	0,0577037	1,102E-11	0,0118171	4,1987544

Tabela 155 – Teste ANOVA para o f1-micro - conjunto de dados Corel5k

Parameter	F1-micro			
Dataset	Corel5k			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,0826953	0,1748934	0,1753521	0,1699389
2	0,1019108	0,1737701	0,1701501	0,1710059
3	0,1121588	0,1768728	0,164923	0,1657211
4	0,0925834	0,1757331	0,1609476	0,1768889
5	0,1123934	0,171864	0,1654304	0,160015
6	0,0983926	0,169064	0,1666635	0,1754316
7	0,0918264	0,1658793	0,1674988	0,1692785
8	0,1020813	0,1764226	0,1753352	0,1470935
9	0,1197605	0,1746484	0,1785702	0,1698895
10	0,1216016	0,1749876	0,165872	0,1665707
Average	0,1035404	0,1734135	0,1690743	0,1671834

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05							
<i>Group</i>	<i>Count</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>	<i>SS</i>	<i>Std Err</i>	<i>Lower</i>	<i>Upper</i>
LIFT	10	1,035404	0,1035404	0,0001628	0,001465	0,0026429	0,0981804	0,1089004
CLMLC	10	1,7341351	0,1734135	1,231E-05	0,0001108	0,0026429	0,1680535	0,1787735
CBMLC	10	1,6907428	0,1690743	3,167E-05	0,000285	0,0026429	0,1637143	0,1744343
HOMER	10	1,6718336	0,1671834	7,263E-05	0,0006536	0,0026429	0,1618234	0,1725434

ANOVA

<i>Sources</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P value</i>	<i>F crit</i>	<i>RMSSE</i>	<i>Omega Sq</i>
Between Groups	0,0332215	3	0,0110738	158,54193	8,439E-21	2,8662656	3,9817324	0,9219704
Within Groups	0,0025145	36	6,985E-05					
Total	0,035736	39	0,0009163					

Tabela 156 – Teste Tukey HSD para o f1-macro - conjunto de dados Corel5k

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,1035404	10	0,001465			
CLMLC	0,1734135	10	0,0001108			
CBMLC	0,1690743	10	0,000285			
HOMER	0,1671834	10	0,0006536			
		40	0,0025145	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0698731	0,0026429	26,43829	0,0598064	0,0799398	7,994E-15	0,0100667	6,3935928
LIFT	CBMLC	0,0655339	0,0026429	24,796431	0,0554672	0,0756006	7,994E-15	0,0100667	5,9965409
LIFT	HOMER	0,063643	0,0026429	24,080954	0,0535763	0,0737097	7,994E-15	0,0100667	5,8235163
CLMLC	CBMLC	0,0043392	0,0026429	1,6418581	0,0057275	0,0144059	0,654818	0,0100667	0,3970519
CLMLC	HOMER	0,0062301	0,0026429	2,357336	0,0038366	0,0162969	0,3555996	0,0100667	0,5700764
CBMLC	HOMER	0,0018909	0,0026429	0,7154779	0,0081758	0,0119576	0,9571532	0,0100667	0,1730246

Tabela 157 – Teste ANOVA para o f1-micro - conjunto de dados Emotions

Parameter	F1-micro			
Dataset	Emotions			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,0941918	0,6010473	0,6269191	0,6781408
2	0,0889622	0,5993481	0,6160806	0,6765824
3	0,0849121	0,599407	0,6252126	0,6725701
4	0,0855252	0,6025059	0,6044882	0,6763578
5	0,0836931	0,6125821	0,6192393	0,6832693
6	0,088737	0,6068501	0,6097245	0,6646896
7	0,0938768	0,6090468	0,6232482	0,6754865
8	0,0948416	0,5929094	0,6273261	0,677962
9	0,1054054	0,596628	0,618603	0,6730962
10	0,104169	0,5832063	0,623603	0,6809681
Average	0,0924314	0,6003531	0,6194445	0,6759123

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	0,9243142	0,0924314	5,815E-05	0,0005234	0,0023028	0,0877612	0,0971017	
CLMLC	10	6,003531	0,6003531	7,085E-05	0,0006377	0,0023028	0,5956828	0,6050234	
CBMLC	10	6,1944446	0,6194445	5,703E-05	0,0005133	0,0023028	0,6147742	0,6241147	
HOMER	10	6,7591228	0,6759123	2,608E-05	0,0002347	0,0023028	0,671242	0,6805826	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	2,2135984	3	0,7378661	13914,577	3,349E-55	2,8662656	37,302249	0,9990426
Within Groups	0,001909	36	5,303E-05					
Total	2,2155075	39	0,0568079					

Tabela 158 – Teste Tukey HSD para o f1-macro - conjunto de dados Emotions

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,0924314	10	0,0005234			
CLMLC	0,6003531	10	0,0006377			
CBMLC	0,6194445	10	0,0005133			
HOMER	0,6759123	10	0,0002347			
		40	0,001909	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,5079217	0,0023028	220,56823	0,4991504	0,516693	7,438E-15	0,0087713	46,476311
LIFT	CBMLC	0,527013	0,0023028	228,85878	0,5182417	0,5357844	7,327E-15	0,0087713	48,223226
LIFT	HOMER	0,5834809	0,0023028	253,38029	0,5747095	0,5922522	7,105E-15	0,0087713	53,390196
CLMLC	CBMLC	0,0190914	0,0023028	8,2905441	0,01032	0,0278627	6,169E-06	0,0087713	1,7469148
CLMLC	HOMER	0,0755592	0,0023028	32,812059	0,0667879	0,0843305	7,994E-15	0,0087713	6,9138852
CBMLC	HOMER	0,0564678	0,0023028	24,521515	0,0476965	0,0652391	7,994E-15	0,0087713	5,1669705

Tabela 159 – Teste ANOVA para o f1-micro - conjunto de dados Genbase

Parameter	F1-micro			
Dataset	Genbase			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,9718476	0,9768851	0,9908638	0,9824346
2	0,971175	0,9804043	0,990308	0,9834578
3	0,971289	0,9783652	0,9909981	0,9841884
4	0,9719271	0,9771229	0,9922861	0,9854423
5	0,9706196	0,9796692	0,9930027	0,9842211
6	0,970009	0,9790165	0,9928289	0,9849706
7	0,9711758	0,9777145	0,9922372	0,9853386
8	0,9713061	0,9796097	0,9903032	0,9860071
9	0,971274	0,9781928	0,992334	0,9842065
10	0,9707896	0,9800815	0,9934205	0,9835875
Average	0,9711413	0,9787062	0,9918582	0,9843854

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	9,7114129	0,9711413	3,182E-07	2,863E-06	0,0003285	0,970475	0,9718076	
CLMLC	10	9,7870617	0,9787062	1,533E-06	1,38E-05	0,0003285	0,9780399	0,9793725	
CBMLC	10	9,9185825	0,9918582	1,311E-06	1,18E-05	0,0003285	0,9911919	0,9925246	
HOMER	10	9,8438545	0,9843854	1,155E-06	1,04E-05	0,0003285	0,9837191	0,9850518	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0023073	3	0,0007691	712,5029	4,267E-32	2,8662656	8,4409887	0,981605
Within Groups	3,886E-05	36	1,079E-06					
Total	0,0023461	39	6,016E-05					

Tabela 160 – Teste Tukey HSD para o f1-macro - conjunto de dados Genbase

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,9711413	10	2,863E-06			
CLMLC	0,9787062	10	1,38E-05			
CBMLC	0,9918582	10	1,18E-05			
HOMER	0,9843854	10	1,04E-05			
		40	3,886E-05	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0075649	0,0003285	23,025463	0,0063135	0,0088163	7,994E-15	0,0012514	0,6922087
LIFT	CBMLC	0,020717	0,0003285	63,056846	0,0194655	0,0219684	7,994E-15	0,0012514	1,895662
LIFT	HOMER	0,0132442	0,0003285	40,311652	0,0119927	0,0144956	7,994E-15	0,0012514	1,211879
CLMLC	CBMLC	0,0131521	0,0003285	40,031383	0,0119007	0,0144035	7,994E-15	0,0012514	1,2034534
CLMLC	HOMER	0,0056793	0,0003285	17,28619	0,0044279	0,0069307	1,421E-13	0,0012514	0,5196704
CBMLC	HOMER	0,0074728	0,0003285	22,745193	0,0062214	0,0087242	7,994E-15	0,0012514	0,683783

Tabela 161 – Teste ANOVA para o f1-micro - conjunto de dados Medical

Parameter	F1-micro			
Dataset	Medical			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,6836779	0,8020853	0,7703028	0,6922703
2	0,6728288	0,7909406	0,7612895	0,7028662
3	0,6852801	0,7873797	0,7596039	0,703493
4	0,6804052	0,7906626	0,7547784	0,6983101
5	0,6795131	0,7971158	0,7521559	0,7086982
6	0,6803988	0,7908377	0,7636631	0,6920847
7	0,6822565	0,7937121	0,7591127	0,6926891
8	0,6870088	0,7820447	0,7691363	0,7028835
9	0,6655062	0,7908369	0,7668245	0,6974441
10	0,6798377	0,796551	0,7647825	0,7055134
Average	0,6796713	0,7922166	0,762165	0,6996252

ANOVA: Single Factor

DESCRIPTION	Alpha		0,05					
<i>Group</i>	<i>Count</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>	<i>SS</i>	<i>Std Err</i>	<i>Lower</i>	<i>Upper</i>
LIFT	10	6,7967131	0,6796713	3,962E-05	0,0003566	0,0018778	0,675863	0,6834796
CLMLC	10	7,9221663	0,7922166	3,079E-05	0,0002771	0,0018778	0,7884083	0,796025
CBMLC	10	7,6216497	0,762165	3,522E-05	0,000317	0,0018778	0,7583566	0,7659733
HOMER	10	6,9962525	0,6996252	3,54E-05	0,0003186	0,0018778	0,6958169	0,7034336

ANOVA

<i>Sources</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P value</i>	<i>F crit</i>	<i>RMSSE</i>	<i>Omega Sq</i>
Between Groups	0,0831433	3	0,0277144	785,98252	7,503E-33	2,8662656	8,8655655	0,9832982
Within Groups	0,0012694	36	3,526E-05					
Total	0,0844126	39	0,0021644					

Tabela 162 – Teste Tukey HSD para o f1-macro - conjunto de dados Medical

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,6796713	10	0,0003566			
CLMLC	0,7922166	10	0,0002771			
CBMLC	0,762165	10	0,000317			
HOMER	0,6996252	10	0,0003186			
		40	0,0012694	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,1125453	0,0018778	59,935073	0,1053928	0,1196978	7,994E-15	0,0071525	10,298225
LIFT	CBMLC	0,0824937	0,0018778	43,931313	0,0753412	0,0896462	7,994E-15	0,0071525	7,5484106
LIFT	HOMER	0,0199539	0,0018778	10,626304	0,0128014	0,0271064	4,152E-08	0,0071525	1,8258436
CLMLC	CBMLC	0,0300517	0,0018778	16,003761	0,0228992	0,0372042	1,238E-12	0,0071525	2,7498144
CLMLC	HOMER	0,0925914	0,0018778	49,308769	0,0854389	0,0997439	7,994E-15	0,0071525	8,4723814
CBMLC	HOMER	0,0625397	0,0018778	33,305009	0,0553872	0,0696922	7,994E-15	0,0071525	5,722567

Tabela 163 – Teste ANOVA para o f1-micro - conjunto de dados Yeast

Parameter	F1-micro			
Dataset	Yeast			
Run	LIFT	CLMLC	CBMLC	HOMER
1	0,652798	0,6171935	0,620685	0,6486819
2	0,651865	0,6146407	0,6182734	0,6520733
3	0,6514507	0,6090576	0,6243455	0,643081
4	0,6521155	0,6061745	0,6187149	0,6455042
5	0,653244	0,6041513	0,6191854	0,6492297
6	0,6499681	0,6145854	0,615677	0,6475917
7	0,6506753	0,6026416	0,6153211	0,644109
8	0,6521748	0,6193562	0,6192716	0,6485117
9	0,652313	0,6091815	0,6170428	0,6489452
10	0,6524421	0,6167978	0,6245873	0,6478336
Average	0,6519047	0,611378	0,6193104	0,6475561

ANOVA: Single Factor

DESCRIPTION	Alpha 0,05								
Group	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper	
LIFT	10	6,5190465	0,6519047	9,594E-07	8,634E-06	0,0011506	0,6495712	0,6542381	
CLMLC	10	6,1137801	0,611378	3,485E-05	0,0003136	0,0011506	0,6090446	0,6137114	
CBMLC	10	6,1931039	0,6193104	1,009E-05	9,079E-05	0,0011506	0,616977	0,6216438	
HOMER	10	6,4755614	0,6475561	7,057E-06	6,351E-05	0,0011506	0,6452227	0,6498896	

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0,0122333	3	0,0040778	308,04166	1,03E-25	2,8662656	5,5501501	0,9583821
Within Groups	0,0004766	36	1,324E-05					
Total	0,0127098	39	0,0003259					

Tabela 164 – Teste Tukey HSD para o f1-macro - conjunto de dados Yeast

TUKEY HSD/KRAMER			alpha	0,05		
<i>group</i>	<i>mean</i>	<i>n</i>	<i>ss</i>	<i>df</i>	<i>q-crit</i>	
LIFT	0,6519047	10	8,634E-06			
CLMLC	0,611378	10	0,0003136			
CBMLC	0,6193104	10	9,079E-05			
HOMER	0,6475561	10	6,351E-05			
		40	0,0004766	36	3,809	

Q TEST									
<i>group 1</i>	<i>group 2</i>	<i>mean</i>	<i>std err</i>	<i>q-stat</i>	<i>lower</i>	<i>upper</i>	<i>p-value</i>	<i>mean-crit</i>	<i>Cohen d</i>
LIFT	CLMLC	0,0405266	0,0011506	35,223686	0,0361442	0,0449091	7,994E-15	0,0043824	3,7083054
LIFT	CBMLC	0,0325943	0,0011506	28,329263	0,0282118	0,0369767	7,994E-15	0,0043824	2,9824693
LIFT	HOMER	0,0043485	0,0011506	3,7795057	-3,393E-05	0,008731	0,0524242	0,0043824	0,3979016
CLMLC	CBMLC	0,0079324	0,0011506	6,8944226	0,0035499	0,0123148	0,000125	0,0043824	0,7258362
CLMLC	HOMER	0,0361781	0,0011506	31,44418	0,0317957	0,0405606	7,994E-15	0,0043824	3,3104038
CBMLC	HOMER	0,0282457	0,0011506	24,549757	0,0238633	0,0326282	7,994E-15	0,0043824	2,5845676

B.3 Teste de Hipóteses 2

Tabela 165 – Teste de Friedman para tempo de execução

Parameter	tExe			
	Algorithm			
Dataset	LIFT	CLMLC	CBMLC	HOMER
Bibtex	2331,0447	6,0513572	118,39242	13,66046
Birds	1,2759476	0,0331892	0,1225968	0,0599631
Corel5k	900,50865	2,2479616	36,352561	77,241927
Emotions	0,6304353	0,0143795	0,0075701	0,0157801
Genbase	2,9561392	0,1006733	2,1416091	1,6640464
Medical	12,244966	0,3280205	6,1916931	0,4222274
Yeast	38,307622	0,1620632	0,0586345	0,0728473

Friedman's Test

Alpha

Q-stat

df

p-value

Q*-stat

df1

df2

p-value

Tabela 166 – Teste de Conover para tempo de execução

Row Ranks Table

LIFT	CLMLC	CBMLC	HOMER
4	1	3	2
4	1	3	2
4	1	2	3
4	2	1	3
4	1	3	2
4	1	3	2
4	3	1	2
28	10	16	16

FRIEDMAN-CONOVER TEST			alpha	0,05	0,0083333	
group	R-sum	R-mean	t-crit	df	std err	R-crit
LIFT	28	4				
CLMLC	10	1,4285714				
CBMLC	16	2,2857143				
HOMER	16	2,2857143				
			2,100922	18	0,4096345	0,8606102
T TEST						
group 1	group 2	R-mean	t-stat	p-value		
LIFT	CLMLC	2,5714286	6,2773725	6,418E-06		
LIFT	CBMLC	1,7142857	4,184915	0,0005566		
LIFT	HOMER	1,7142857	4,184915	0,0005566		
CLMLC	CBMLC	0,8571429	2,0924575	0,0508355		
CLMLC	HOMER	0,8571429	2,0924575	0,0508355		
CBMLC	HOMER	0	0	1		

Tabela 167 – Teste de Friedman para tempo de CPU

Parameter	tCPU			
	Algorithm			
Dataset	LIFT	CLMLC	CBMLC	HOMER
Bibtex	5901,9859	20,239844	387,25094	28,437188
Birds	2,2153125	0,1240625	0,44625	0,179375
Corel5k	2473,2094	8,0807813	110,95078	136,20047
Emotions	0,8070313	0,0426563	0,0257813	0,0503125
Genbase	7,8692188	0,3334375	6,5723438	4,7578125
Medical	34,0225	1,0598438	21,042813	1,1623438
Yeast	42,086875	0,6057813	0,20625	0,2432813

Friedman's Test

Alpha

Q-stat

df

p-value

Q*-stat

df1

df2

p-value

Tabela 168 – Teste de Conover para tempo de CPU

Row Ranks Table

LIFT	CLMLC	CBMLC	HOMER
4	1	3	2
4	1	3	2
4	1	2	3
4	2	1	3
4	1	3	2
4	1	3	2
4	3	1	2
28	10	16	16

FRIEDMAN-CONOVER TEST			alpha	0,05	0,0083333	
group	R-sum	R-mean	t-crit	df	std err	R-crit
LIFT	28	4				
CLMLC	10	1,4285714				
CBMLC	16	2,2857143				
HOMER	16	2,2857143				
			2,100922	18	0,4096345	0,8606102
T TEST						
group 1	group 2	R-mean	t-stat	p-value		
LIFT	CLMLC	2,5714286	6,2773725	6,418E-06		
LIFT	CBMLC	1,7142857	4,184915	0,0005566		
LIFT	HOMER	1,7142857	4,184915	0,0005566		
CLMLC	CBMLC	0,8571429	2,0924575	0,0508355		
CLMLC	HOMER	0,8571429	2,0924575	0,0508355		
CBMLC	HOMER	0	0	1		

Tabela 169 – Teste de Friedman para exact match ratio

Parameter	EM			
	Algorithm			
Dataset	LIFT	CLMLC	CBMLC	HOMER
Bibtex	0,1350847	0,1737106	0,1543042	0,0963047
Birds	0,4449784	0,518762	0,5045793	0,4756803
Corel5k	0,0062	0,0291	0,01104	0,00108
Emotions	0,008596	0,259209	0,2500339	0,2836497
Genbase	0,9461782	0,9670222	0,9817209	0,96846
Medical	0,4907942	0,6987703	0,5875531	0,5697928
Yeast	0,193618	0,1935419	0,1946437	0,1450739

Friedman's Test

Alpha

Q-stat

df

p-value

Q*-stat

df1

df2

p-value

Tabela 170 – Teste de Conover para exact match ratio

Row Ranks Table

LIFT	CLMLC	CBMLC	HOMER
2	4	3	1
1	4	3	2
2	4	3	1
1	3	2	4
1	2	4	3
1	4	3	2
3	2	4	1
11	23	22	14

FRIEDMAN-CONOVER TEST			alpha	0,05	0,0083333	
group	R-sum	R-mean	t-crit	df	std err	R-crit
LIFT	11	1,5714286				
CLMLC	23	3,2857143				
CBMLC	22	3,1428571				
HOMER	14	2				
			2,100922	18	0,5634362	1,1837355
T TEST						
group 1	group 2	R-mean	t-stat	p-value		
LIFT	CLMLC	1,7142857	3,0425553	0,0070053		
LIFT	CBMLC	1,5714286	2,789009	0,0121187		
LIFT	HOMER	0,4285714	0,7606388	0,4567219		
CLMLC	CBMLC	0,1428571	0,2535463	0,8027195		
CLMLC	HOMER	1,2857143	2,2819165	0,03488		
CBMLC	HOMER	1,1428571	2,0283702	0,0575839		

Tabela 171 – Teste de Friedman para hamming loss

Parameter	HL			
	Algorithm			
Dataset	LIFT	CLMLC	CBMLC	HOMER
Bibtex	0,014204	0,0137333	0,0163102	0,0218866
Birds	0,0542291	0,046497	0,0591154	0,0592451
Corel5k	0,0096337	0,0136051	0,0118681	0,0165787
Emotions	0,3100683	0,2437585	0,2372665	0,2099214
Genbase	0,0026444	0,0019796	0,0007612	0,0014375
Medical	0,0149911	0,0112211	0,013499	0,0166221
Yeast	0,1912954	0,2317057	0,2187007	0,2353568

Friedman's Test

Alpha

Q-stat

df

p-value

Q*-stat

df1

df2

p-value

Tabela 172 – Teste de Friedman para f1-macro

Parameter	F1-macro			
	Algorithm			
Dataset	LIFT	CLMLC	CBMLC	HOMER
Bibtex	0,1804684	0,2093946	0,2444059	0,1397655
Birds	0	0,3257938	0,3043383	0,2332728
Corel5k	0,0303221	0,0363517	0,0247183	0,033805
Emotions	0,0727593	0,5873927	0,607772	0,6604722
Genbase	0,5825311	0,6493411	0,6591728	0,6139226
Medical	0,2393069	0,3321304	0,3319279	0,2254562
Yeast	0,4050479	0,4012968	0,3812902	0,4211379

Friedman's Test

Alpha

Q-stat

df

p-value

Q*-stat

df1

df2

p-value

Tabela 173 – Teste de Friedman para f1-micro

Parameter	F1-micro			
	Algorithm			
Dataset	LIFT	CLMLC	CBMLC	HOMER
Bibtex	0,2635828	0,4046379	0,3718571	0,2765548
Birds	0	0,481239	0,4392916	0,393405
Corel5k	0,1035404	0,1734135	0,1690743	0,1671834
Emotions	0,0924314	0,6003531	0,6194445	0,6759123
Genbase	0,9711413	0,9787062	0,9918582	0,9843854
Medical	0,6796713	0,7922166	0,762165	0,6996252
Yeast	0,6519047	0,611378	0,6193104	0,6475561

Friedman's Test

Alpha

Q-stat

df

p-value

Q*-stat

df1

df2

p-value

