

Tags de Opinião

Dulce de Fátima Neves Morim

Dissertação para obtenção do Grau de Mestre em
Engenharia Informática

Área de Especialização em
Tecnologias do Conhecimento e Decisão

Orientadora: Mestre Isabel de Fátima Silva Azevedo

Júri:

Presidente:

Vogais:

Porto, Outubro 2011

Agradecimentos

Neste ponto escrevo alguns agradecimentos às pessoas que, de alguma forma, contribuíram para a concretização deste trabalho. Agradeço a todos que me apoiaram e fizeram com que o meu esforço e dedicação permitissem concluir esta dissertação.

Agradeço à minha família o apoio e força incondicional. Não existem palavras que consigam expressar este agradecimento da forma merecida.

Aos meus amigos pela paciência, compreensão e pela força em momentos menos bons. Obrigada pelos programas de convívio que me permitiram descontraír e ter mais força para atingir os meus objectivos.

Agradeço à Eng.^a Isabel Azevedo pela disponibilidade, comentários, apoio e ajuda no desenvolvimento deste trabalho, escrita do presente documento e desenvolvimento do artigo submetido sobre o estudo realizado. Obrigada pelas palavras de incentivo que constituíram um apoio muito importante.

Agradeço a disponibilidade e ajuda da Prof. Doutora Fátima Rodrigues na identificação de análises complementares a aplicar neste estudo.

Um agradecimento à professora Ana Barata pela disponibilidade e ajuda na revisão do resumo deste trabalho.

Também um agradecimento aos meus colegas de trabalho pela compreensão nos dias em que estava mais cansada ao longo da realização deste trabalho.

Resumo

As tags podem ser utilizadas com diferentes propósitos, entre os quais organizar os recursos para fins pessoais e partilhar informação potencialmente relevante com outros utilizadores.

Este trabalho conjuga diversas áreas de conhecimento e explora a utilização de tags, debruçando-se numa categoria em particular que engloba as tags de opinião. Estas podem ser usadas para expressar sentimentos ou opiniões sobre os recursos.

Foram realizadas análises sobre a utilização de tags na loja online Amazon. Recolhida uma amostra de dados, as tags foram classificadas e analisadas segundo diversos aspectos, inclusive quanto à sua polaridade.

Além da atribuição de tags, na Amazon é possível atribuir pontuações (de 1 a 5) aos recursos. Neste trabalho compararam-se ainda as duas formas referidas de classificação de recursos, verificando a existência de alguma correspondência entre ambas, com significado estatístico.

Adicionalmente, foi desenvolvido um classificador semi-automático que tem como objectivo classificar as tags atribuídas aos recursos para proporcionar uma classificação rápida e eficiente da polaridade das tags que considera também a informação disponível sobre os recursos durante o processo.

Palavras-chave: Folksonomia, tags de opinião, Amazon.

Abstract

Tags can be used for different purposes, including organizing resources for personal use and sharing potentially relevant information with other users.

This work combines several knowledge areas and explores the use of tags, in particular the opinion tags. These can be used to express feelings or opinions about the resources.

Analyses were performed on the use of tags in the Amazon online store. After selecting a data sample, the tags were classified and analyzed according to various aspects, including their polarity.

In addition to the allocation of tags in Amazon, scores (1-5) can be assigned to resources. This work also compared the two mentioned resource classification forms, checking for any correspondence between them, with a statistical significance.

Additionally, we developed a semi-automatic classifier which aims to classify the tags assigned to resources to provide a fast and efficient classification of the polarity of the tags that also considers the available information on the resources during the process.

Keywords: Folksonomy, opinion tags, Amazon.

Índice

<i>Agradecimentos</i> -----	<i>iii</i>
<i>Resumo</i> -----	<i>v</i>
<i>Abstract</i> -----	<i>vii</i>
<i>Índice</i> -----	<i>ix</i>
<i>Índice de Figuras</i> -----	<i>xiii</i>
<i>Índice de Tabelas</i> -----	<i>xv</i>
1 <i>Introdução</i> -----	1
1.1 Enquadramento e motivações -----	1
1.2 Objectivos e apresentação do trabalho -----	2
1.3 Considerações de escrita -----	4
1.4 Estrutura do documento -----	4
2 <i>Tags</i> -----	7
2.1 <i>Tagging</i> -----	7
2.1.1 <i>Definição</i> -----	7
2.1.2 <i>Categorização de tags</i> -----	8
2.2 Folksonomia -----	10
2.2.1 <i>Definição</i> -----	10
2.2.2 <i>Características</i> -----	11
2.2.3 <i>Vantagens e desvantagens</i> -----	13
2.2.4 <i>Folksonomias gerais e específicas</i> -----	13
2.3 Serviços de informação colaborativa -----	15
2.3.1 <i>Bookmarking social</i> -----	15
2.3.2 <i>Comércio na Web</i> -----	17
3 <i>Análise de sentimentos</i> -----	19
3.1 Objectivos -----	19
3.2 Áreas de aplicação -----	20

3.3	Técnicas -----	20
3.3.1	Técnicas de aprendizagem de máquina -----	20
3.3.2	Técnicas simbólicas -----	21
3.3.2.1	Recursos léxicos -----	21
3.4	Classificação da polaridade -----	23
3.5	Tags de opinião ou sentimentais -----	23
3.6	Estudos sobre opinião -----	25
3.6.1	Estudos relacionados com a Amazon-----	25
3.6.2	Outros estudos-----	26
4	<i>Estudo de caso</i> -----	27
4.1	Amazon -----	27
4.1.1	Caracterização da Amazon-----	27
4.1.2	Escolha da Amazon -----	29
4.2	Amostragem -----	30
4.2.1	Processo não aleatório -----	31
4.3	Base de dados -----	33
4.4	Classificação de tags -----	34
4.4.1	Processo de classificação de tags -----	34
4.4.2	Classificação semi-automática -----	36
5	<i>Análise dos resultados</i> -----	41
5.1	Número de palavras nas tags -----	41
5.2	Contabilização de tags e pontuações -----	43
5.3	Tags com uso da negação ou intensificação -----	45
5.4	Tags indefinidas ou imprecisas -----	46
5.5	Utilização das tags -----	46
5.5.1	Análise comparativa entre tags e pontuação média -----	47
5.5.2	Análise da correlação entre tipos de tags e pontuação média -----	53
5.6	Utilização das tags de opinião -----	54
5.6.1	Análise comparativa entre tags de opinião e pontuação média -----	54

5.6.2	Análise da correlação entre tags de opinião e pontuação média -----	58
5.6.3	Tags de opinião mais usadas-----	58
5.6.4	Análise da dependência entre tags de opinião e pontuação média -----	64
6	<i>Conclusões</i> -----	69
6.1	Considerações finais -----	69
6.2	Limitações e dificuldades encontradas -----	70
6.3	Trabalho futuro -----	71
	<i>Bibliografia</i> -----	73
Anexo 1	<i>Dados recolhidos da Amazon</i> -----	77
1.1	Produtos -----	77
1.2	Tags -----	78

Índice de Figuras

Figura 1 - Esquema dos processos realizados para análise dos dados da Amazon	4
Figura 2 - Exemplo de uma pesquisa por tag através da Galeria de Fotografias do Windows Live	8
Figura 3 - Modelo de categoria funcional de tags	9
Figura 4 - Modelo de sistema tagging social	11
Figura 5 - Produtos da Amazon com as tags thriller e adventure	12
Figura 6 - Ilustração do funcionamento das folksonomias broad e narrow	14
Figura 7 - Pesquisa de tag e apresentação das tags relacionadas no Del.icio.us	16
Figura 8 - Pesquisa de tag e apresentação das tags relacionadas no Del.icio.us (novo design)	16
Figura 9 - Um produto com as tags associadas e a possibilidade de adicionar ou pesquisar tags na Amazon	17
Figura 10 - Tags sentimentais utilizadas no serviço social de marcação japonês Hatena Bookmarks	24
Figura 11 - Sugestão de produtos com base nas compras de outros utilizadores na Amazon	27
Figura 12 - Pontuações atribuídas a um produto na Amazon	28
Figura 13 - Atribuir tag a um produto na Amazon	28
Figura 14 - Funcionalidade de atribuir tag a um produto na Amazon através de sugestões	29
Figura 15 - Representação do funcionamento do processo de amostragem	32
Figura 16 - Esquema das tabelas para registo dos dados recolhidos dos produtos	34
Figura 17 - Funcionamento do classificador automático de tags	38
Figura 18 - Tabelas auxiliares para a classificação das tags	39
Figura 19 - Percentagem de tags constituídas por uma ou várias palavras na amostra da Amazon	42
Figura 20 - Representação da percentagem de tags por quantidade de palavras usadas na amostra da Amazon	42
Figura 21 - Totais de tags e pontuações atribuídas aos produtos da amostra da Amazon para comparativo	44
Figura 22 - Análise do número de tags positivas, negativas e neutras atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação negativa	49
Figura 23 - Análise da percentagem da pontuação (1 a 5) por pontuação atribuída (correspondência de produto pela pontuação do gráfico da Figura 22)	49
Figura 24 - Análise do número de tags positivas, negativas e neutras atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação intermédia	50
Figura 25 - Análise da percentagem da pontuação (1 a 5) por pontuação atribuída (correspondência de produto pela pontuação do gráfico da Figura 24)	50

Figura 26 - Análise do número de tags positivas, negativas e neutras atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação positiva _____	51
Figura 27 - Análise da percentagem da pontuação (1 a 5) por pontuação atribuída (correspondência de produto pela pontuação do gráfico da Figura 26) _____	52
Figura 28 - Análise do número de tags de opinião atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação negativa _____	56
Figura 29 - Análise do número de tags de opinião atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação intermédia _____	57
Figura 30 - Análise do número de tags de opinião atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação positiva _____	57
Figura 31 - Tags de opinião mais usadas da amostra (Amazon) _____	62
Figura 32 - Classes gramaticais das tags de opinião mais usadas da amostra (Amazon) _____	63

Índice de Tabelas

Tabela 1 - Classificação da tag funny em diferentes recursos léxicos _____	22
Tabela 2 - Dados sobre o uso de algumas tags de opinião na Amazon _____	25
Tabela 3 - Recolha dos totais de tags e pontuações atribuídas aos produtos da amostra da Amazon para comparativo _____	44
Tabela 4 - Tags com uso da negação da amostra da Amazon _____	45
Tabela 5 - Tags indefinidas ou imprecisas da amostra da Amazon _____	46
Tabela 6 - Recolha de dados de produtos da Amazon _____	47
Tabela 7 - Recolha de dados sobre as pontuações de produtos da Amazon _____	53
Tabela 8 - Correlação entre percentagem de tags (por tipo) e média ponderada das pontuações atribuídas pelos utilizadores _____	54
Tabela 9 - Recolha de dados de produtos da Amazon considerando só tags de opinião _____	55
Tabela 10 - Correlação entre percentagem de tags de opinião e média das pontuações atribuídas pelos utilizadores _____	58
Tabela 11 - Tags de opinião mais usadas da amostra (Amazon) _____	59
Tabela 12 - Frequência de utilização de algumas tags considerado o tipo de pontuação média atribuída aos recursos _____	61
Tabela 13 - Tags de opinião atribuídas a mais recursos da amostra da Amazon _____	64
Tabela 14 - Tabela de contingência para avaliação de utilização da tag horrible e pontuação média (amostra Amazon) _____	66
Tabela 15 - Tabela de contingência para avaliação de utilização da tag boring e pontuação média (amostra Amazon) _____	66
Tabela 16 - Tabela de contingência para avaliação de utilização da tag book recommendations e pontuação média (amostra Amazon) _____	67
Tabela 17 - Tabela de contingência para avaliação de utilização da tag funny e pontuação média (amostra Amazon) _____	67
Tabela 18 - Dados dos produtos recolhidos da Amazon _____	77
Tabela 19 - Tags classificadas da amostra da Amazon _____	78

1 Introdução

Este capítulo pretende fornecer uma visão global do trabalho apresentado neste documento. O tema do trabalho é “tags de opinião” e neste capítulo é feito o enquadramento do tema, são apresentadas as motivações para a sua realização e os objectivos definidos.

Neste mesmo capítulo são ainda descritas as considerações de escrita do documento. Estas foram identificadas e aplicadas para facilitar a leitura deste documento. Também é descrita a estrutura do documento, com uma breve descrição de cada um dos capítulos.

1.1 Enquadramento e motivações

Este trabalho foi desenvolvido no âmbito da área de especialização em tecnologias do conhecimento e decisão e o mesmo aborda conceitos e estudos relevantes para um conhecimento mais aprofundado da área. O trabalho foi proposto pelo Núcleo de Investigação GILT¹ (Graphics, Interaction and Learning Technologies) que tem desenvolvido projectos que incluem a utilização de tags, nomeadamente em repositórios de recursos educativos.

Os serviços de *bookmarking* social, característicos da Web 2.0, permitem que os utilizadores guardem recursos através de perfis pessoais e possam marcar os recursos através de tags. Assim, os utilizadores auxiliam na indexação dos recursos através do uso de tags. A esta forma de marcar e indexar os recursos é muitas vezes chamada de *tagging* e o resultado desse processo é uma folksonomia.

Considerando que nos sistemas que permitem a atribuição de tags não existe normalmente um vocabulário controlado ou categorias pré-definidas, diversos autores estudaram quais as categorias de tags mais habituais. Uma das categorias sugeridas por vários autores engloba as tags de opinião.

As tags de opinião podem ser classificadas como negativas ou positivas porque através de palavras expressam sentimentos relativamente aos recursos. As atribuições de tags de opinião podem ser consideradas formas de recomendar recursos e como as recomendações normalmente influenciam as compras de outros utilizadores, é importante avaliar se essa influência corresponde realmente ao valor do recurso. Principalmente para lojas *online*, torna-se importante verificar se a opinião expressa pela atribuição de tags tem alguma correspondência com a opinião emitida de outras formas.

Uma das formas de classificação mais conhecida nos sistemas *online* é a atribuição de pontuações (com escala definida) para avaliação de recursos. Estas normalmente são acompanhadas de texto com

¹ <http://gilt.isep.ipp.pt/>

os comentários dos utilizadores que atribuíram as pontuações. A atribuição de tags comparada com a atribuição de pontuações pode trazer algumas vantagens relativamente ao uso:

- As pontuações atribuídas podem não ter o mesmo significado para todas as pessoas, seja para quem atribuí e para quem avalia o recurso consoante a pontuação atribuída;
- As tags permitem indexar e representar informação dos recursos;
- Quando os utilizadores ficam muito satisfeitos ou muito insatisfeitos com um determinado recurso podem expressar facilmente, rapidamente e livremente a sua opinião.

Por outro lado, as pontuações atribuídas aos recursos permitem determinar de uma forma rápida qual a opinião geral mediante a apresentação da pontuação média atribuída pelos utilizadores.

Os sistemas com a funcionalidade de *tagging* estão a aumentar e torna-se necessário compreender e analisar padrões da utilização de tags para que estes sistemas se tornem mais eficientes e as tags atribuídas possam ser também utilizadas de forma menos tradicional.

Esta é uma área de interesse pois mostra-se ser uma área muito actual, pois as tags são cada vez mais usadas e a sua utilização é muito comum em serviços *online* utilizados pelas pessoas praticamente diariamente.

1.2 Objectivos e apresentação do trabalho

Neste ponto são apresentados os objectivos genéricos e específicos que se pretenderam atingir no decorrer do trabalho apresentado nesta dissertação de mestrado. Também é feita uma breve descrição do trabalho realizado.

Os objectivos genéricos definidos para este trabalho foram os seguintes:

- Estudo de folksonomia e sistemas com a funcionalidade de *tagging*;
- Análise dos diferentes tipos de tags, incidindo nas tags de opinião ou sentimentais.

Como objectivos específicos foram definidos os seguintes:

- Analisar especificamente o funcionamento do sistema de *tagging* da Amazon e as tags fornecidas pelos utilizadores segundo diferentes aspectos;
- Usar um conjunto de dados para comparar as tags de opinião utilizadas pelos utilizadores (mais positivas ou negativas) com a classificação fornecida directamente pelos utilizadores que escolhem um número de estrelas (de uma a cinco estrelas).

Para uma análise mais aprofundada sobre o comportamento dos utilizadores ao atribuírem tags de opinião foi necessário realizar um estudo sobre os dados de um sistema com a funcionalidade de *tagging*. A Amazon² foi a escolhida para o estudo e foi analisada mais profundamente pois é uma loja *online* conhecida e mundialmente utilizada. O objectivo do estudo é analisar se as tags de opinião podem efectivamente servir para estimar a opinião dos utilizadores em relação aos recursos. Portanto foi aplicado um processo não aleatório para recolha de uma amostra de dados para comparar as tags de opinião utilizadas pelos utilizadores (mais positivas ou negativas) com a classificação fornecida directamente pelos utilizadores que escolhem um número de estrelas (de uma a cinco estrelas). Através da amostra recolhida foi também possível analisar os dados e identificar padrões na utilização das tags da Amazon. A Amazon foi também escolhida para estudo por permitir os dois tipos de classificação de produtos referidos. As duas formas de classificar produtos permitem uma forma de avaliação dos produtos à venda, visto que as ambas são efectuadas pelos utilizadores. Este tipo de análise é importante pois a opinião pode influenciar outros utilizadores.

Em termos de metodologia o estudo realizado neste trabalho pode ser caracterizado com um estudo de caso sobre a Amazon. Segundo o autor Ponte um estudo de caso “É uma investigação que se assume como particularística, isto é, que se debruça deliberadamente sobre uma situação específica que se supõe ser única ou especial, pelo menos em certos aspectos, procurando descobrir a que há nela de mais essencial e característico e, desse modo, contribuir para a compreensão global de um certo fenómeno de interesse” (Ponte, 2006)

A Figura 1 apresenta de forma gráfica os processos realizados para análise dos dados do estudo de caso. Através da recolha da amostra de dados sobre a Amazon foi possível obter informações de alguns produtos, tags e pontuações atribuídas pelos utilizadores. Para analisar os dados foi necessário classificar as tags como positivas, negativas ou neutras.

Numa primeira fase esta classificação foi um processo manual, todas as tags foram avaliadas e classificadas de acordo com o seu significado e atribuição ao recurso. Numa fase seguinte, depois de conhecidas as “regras” de classificação de tags e depois de todas as tags classificadas foi construído um classificador semi-automático de tags.

² <http://www.amazon.com>

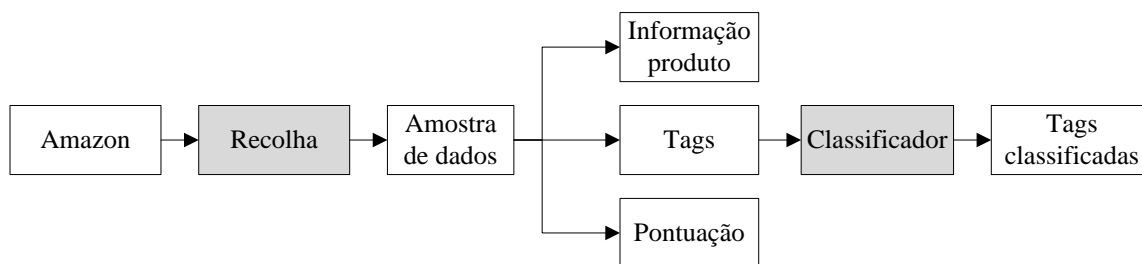


Figura 1 - Esquema dos processos realizados para análise dos dados da Amazon

1.3 Considerações de escrita

Foram consideradas algumas normas na escrita deste documento, nomeadamente as seguintes:

- Evitou-se o uso de estrangeirismos. Estes são utilizados nos casos em que seja mesmo essencial a sua utilização para melhor compreensão e identificação dos mesmos. As palavras estrangeiras aparecem em itálico no texto, com excepção das palavras tag e tags pela frequência das mesmas e por serem normalmente identificadas desta forma;
- Em cada primeira referência a um acrónimo é mencionado o seu significado entre parêntesis;
- Durante o texto, para fácil leitura e identificação, os conteúdos importantes estão a negrito;
- As citações surgem entre aspas;
- As tags de exemplo aparecem durante o texto entre aspas para fácil identificação e diferenciação do restante texto;
- Os sítios Web de alguns sistemas referenciados no documento estão em notas de rodapé para fácil identificação dos mesmos e para evitar essa referência no próprio texto, facilitando assim a sua leitura.

1.4 Estrutura do documento

Nesta secção são apresentadas breves descrições dos capítulos que constituem este documento.

O primeiro capítulo é introdutório. Neste é feito o enquadramento do tema, são apresentadas as motivações para a sua realização e é apresentado o resumo do trabalho realizado. Pretende-se enquadrar e apresentar o trabalho, incluindo os objectivos pretendidos.

No capítulo seguinte, o segundo, são apresentados os conceitos mais importantes para o trabalho realizado. *Tagging*, folksonomia e serviços de informação colaborativa são descritos nesse capítulo para estudo da utilização de tags.

No capítulo 3 descreve-se a área de análise de sentimentos considerando as áreas de aplicação e técnicas utilizadas na área. Também no mesmo capítulo, caracteriza-se a forma de classificação da

polaridade e introduz-se a aplicação da polaridade às tags de opinião ou sentimentais. Para a área em questão são apresentados estudos de análise de polaridade realizados por vários autores.

O capítulo 4 inclui o estudo de caso sobre a loja *online* Amazon. Neste capítulo descreve-se a Amazon como o sistema escolhido para o estudo e as razões da sua escolha. Descreve-se o processo de amostragem realizado e respectivo processo de classificação das tags para a amostra considerada. Neste capítulo é descrito o processo semi-automático desenvolvido.

Foram analisados os dados da amostra sobre a Amazon, o que permitiu encontrar padrões na utilização das tags. Estas análises e respectivos resultados obtidos são apresentados no Capítulo 5.

No último capítulo são apresentadas as conclusões do trabalho realizado. Apresentam-se os objectivos atingidos, respectivo impacto nos serviços de informação colaborativa, limitações e dificuldades sentidas durante a realização deste trabalho. São ainda apresentadas algumas propostas de trabalho futuro.

Existe um anexo (Anexo 1) que descreve alguns aspectos relevantes da amostra recolhida da Amazon e complementa a informação apresentada no Capítulo 4.

2 Tags

Cada vez mais existem sistemas que permitem a atribuição de tags. Neste capítulo são apresentados os conceitos relacionados com tags mais relevantes para o estudo de caso apresentado neste trabalho. *Tagging*, folksonomia e serviços de informação colaborativa são os principais conceitos discutidos neste capítulo.

2.1 *Tagging*

Tags são etiquetas ou marcas que são utilizadas no *tagging*. Nesta secção é apresentado o conceito de *tagging* e as categorizações de tags referidas por alguns autores.

2.1.1 Definição

Tagging é a acção de atribuir tags aos recursos de um sistema, ou seja, é a acção de etiquetar os recursos. Essas tags são atribuídas pelos utilizadores mesmo que estes tenham diferentes interesses, necessidades, objectivos e motivações. Normalmente o interesse comum é o de marcar o recurso para posteriormente encontrá-lo ou para partilhá-lo. Um recurso pode ser um livro, uma fotografia, um vídeo, uma página Web, entre outros. As tags atribuídas aos recursos são palavras-chave escolhidas livremente pelos utilizadores. Pode-se dizer que as tags são metadados sobre os recursos, pois segundo a organização NISO³ (National Information Standards Organization), metadados são “informações estruturadas que descrevem, explicam, localizam, ou tornam mais fácil de recuperar, usar ou gerir uma fonte de informação” (NISO, 2004). Assim, metadados são dados sobre dados.

As tags podem ser atribuídas livremente pelos utilizadores, mas cada sistema de marcação tem as suas regras, desde quem pode marcar, o que pode ser marcado e quais os tipos de tags que podem ser usadas (Smith, 2008).

Uma das grandes vantagens de sistemas com a funcionalidade de *tagging* disponível para os seus utilizadores, é a possibilidade dos mesmos organizarem os recursos de uma forma individualizada. Um exemplo de um sistema com este tipo de organização é a Galeria de Fotografias do Windows Live. Esta aplicação permite gerir fotografias e inclusive usar tags. Assim, um utilizador pode marcar fotografias com os termos considerados apropriados e depois pesquisar, por exemplo, por tag e encontrar rapidamente a(s) fotografia(s) que pretende, como se pode ver na Figura 2.

³ <http://www.niso.org>

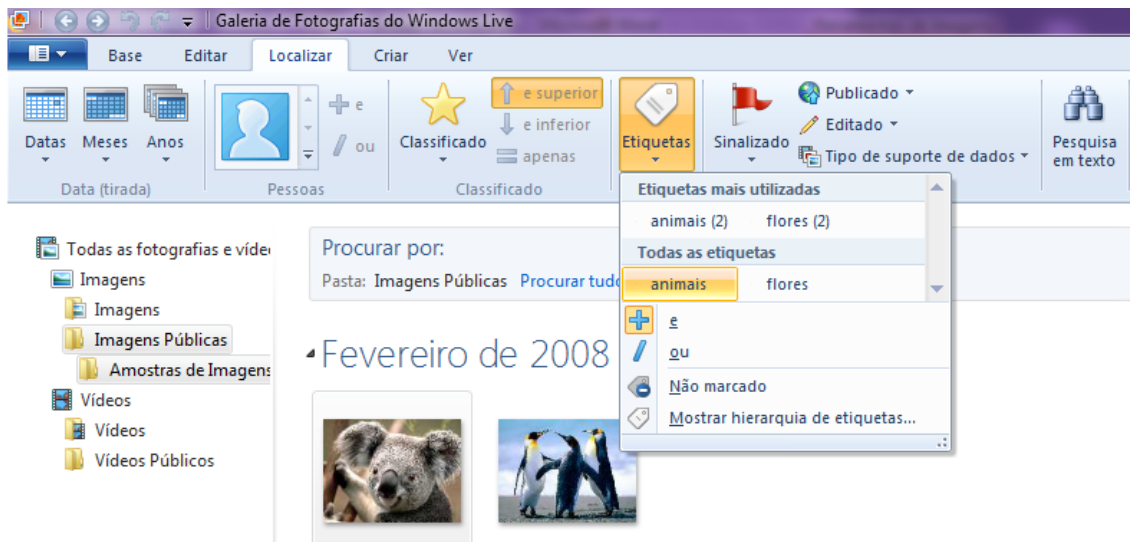


Figura 2 - Exemplo de uma pesquisa por tag através da Galeria de Fotografias do Windows Live

2.1.2 Categorização de tags

Nesta secção são apresentadas algumas categorias de tags, mencionado estudos de diversos autores.

Nos sistemas que permitem a atribuição de tags aos recursos, não existe habitualmente um vocabulário controlado ou categorias pré-definidas. Golder e Huberman identificaram sete categorias de tags no sistema del.icio.us⁴ num estudo concluído em 2006 (Golder & Huberman, 2006):

- Tópicos que agregam descrições genéricas para determinar de que se trata;
- Categorização do tipo para indicar sobre o que é o recurso;
- Identificação de quem é ou quem criou o recurso;
- Categoria que refina outra categoria;
- Identificação de qualidades ou características. Adjectivos que podem ser opiniões sobre os recursos;
- Auto-referência, por exemplo, a tag “meu” (*my*) é usada para identificar o conteúdo em termos da relação com quem fez a marcação;
- Organização de tarefas, que agrega as tags relacionadas com uma tarefa a ser realizada através do recurso, como “para ler” (*toread*), ou relacionadas com uma tarefa a ser realizada pelos utilizadores, como “procura de emprego” (*jobsearch*).

⁴ <http://del.icio.us>

Outro estudo sobre o sistema MoviesLens⁵ sugeriu três categorias mais gerais relacionadas com tarefas específicas do utilizador a saber (Sen, et al., 2006):

- Identificação de factos através de tags, por exemplo, pessoas, lugares ou conceitos sobre um filme. São tags que descrevem os recursos de forma normalmente reconhecida;
- Tags subjectivas que expressam opiniões relativamente aos recursos. Podem ser usadas para ajudar a avaliar uma recomendação;
- Tags pessoais que são para o próprio utilizador. São frequentemente usadas para a organização do utilizador, mas também para auto-referência ou identificação de produto que lhe pertence.

O modelo de categoria funcional (Heckner, Mühlbacher, & Wolff, 2008) encontra-se representado na Figura 3. Nesse modelo as tags são divididas em dois grupos: tags relacionadas com o assunto e tags não relacionadas com o assunto. O primeiro aplica-se a tags usadas para descrever o recurso e conteúdo (autor, data, descrição do conteúdo, área de estudo e outros), enquanto o segundo grupo, aplica-se a tags de carácter pessoal. Estas últimas tags podem estar relacionadas com as actividades dos utilizadores (tempo e tarefas relacionadas) ou estado emocional (afectivo). Os mesmos autores referem o facto das tags que podem ser chamadas de afectivas manifestarem reacções positivas ou negativas sobre os recursos. Pertencente ainda a este segundo grupo, estão as tags que o autor intitula de tags sem objectivos, isto é, sem significado perceptível (por exemplo, “sem tag” ou “teste”). Estas tags são usadas mais frequentemente quando o sistema exige atribuição de tags aos recursos, os utilizadores atribuem estas tags para evitar a atribuição de outras tags.

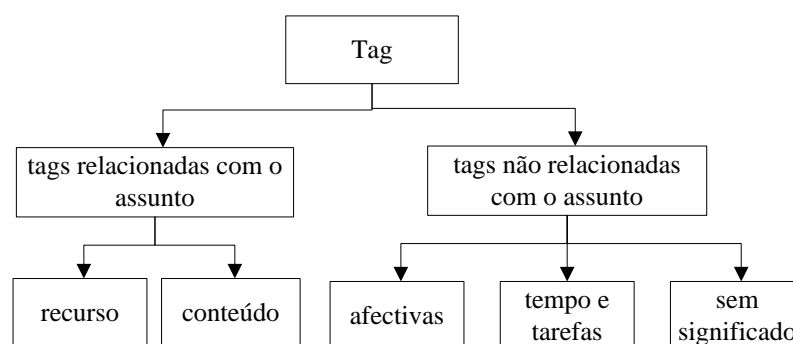


Figura 3 - Modelo de categoria funcional de tags

Adaptação da fonte: (Heckner, Mühlbacher, & Wolff, 2008)

⁵ <http://www.movielens.org>

Todos estes autores apresentam diferentes categorias de tags, mas todos os autores referem o uso de tags como forma de expressar a opinião ou de uso pessoal, e ambas podem ser consideradas formas de fornecer recomendações. Pois mesmo que as tags sejam utilizadas a nível pessoal, podem ser úteis para outros utilizadores. Os autores Golder e Huberman referem isso mesmo, por exemplo, se alguém marca um recurso com “para ler” (*toread*) pode ser do interesse de outras pessoas lerem também o mesmo livro (Golder & Huberman, 2006).

2.2 Folksonomia

Nesta secção é descrito o termo folksonomia, apresenta-se a sua definição, características da sua utilização em sistemas de classificação e quais as suas vantagens e desvantagens. Também nesta secção são caracterizadas e distinguidas as folksonomias gerais e específicas.

2.2.1 Definição

Existem várias definições do termo folksonomia. Para Wal, folksonomia é o resultado do processo de etiquetagem de recursos da Web (Vander Wal, 2007). Peterson refere que folksonomia é baseada numa metodologia de recuperação de informação da Web (Peterson, 2006). Alguns autores também usam outros termos para se referirem a folksonomia, como é o caso do termo “classificação social” (Hammond, Hannay, Lund, & Scott, 2005). A componente social de marcação é considerada na folksonomia, pois torna visíveis todas as tags para todos os utilizadores e existe a possibilidade de adicionar tags a todos os recursos (Vander Wal, 2005).

Folksonomias podem ser definidas de uma forma tripartida visto que existem três aspectos a serem considerados (Marlow, Naaman, boyd, & Davis, 2006):

- Recursos que devem ser referenciáveis;
- Tags que são atribuídas;
- Utilizadores que fazem a indexação.

A Figura 4 mostra um modelo representativo do funcionamento de sistemas de *tagging* social. Os utilizadores atribuem tags aos recursos, assim através das atribuições de tags, passam a existir ligações entre utilizadores e recursos.

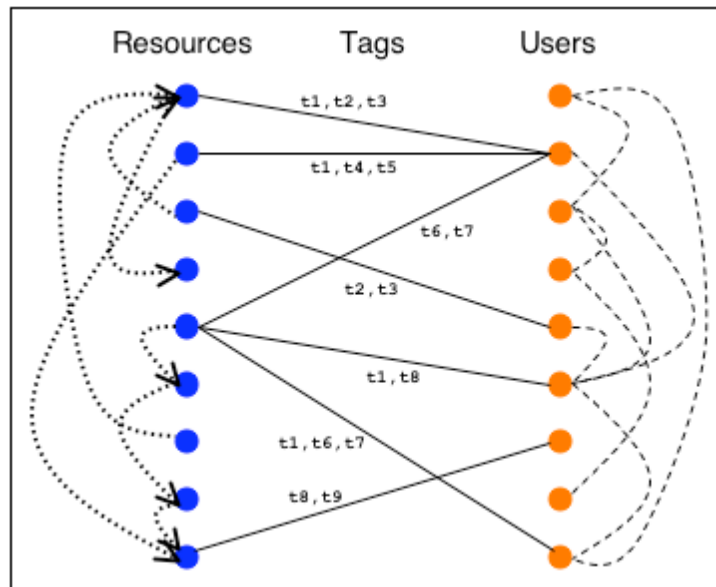


Figura 4 - Modelo de sistema *tagging* social

Fonte: (Marlow, Naaman, boyd, & Davis, 2006)

2.2.2 Características

Folksonomia é um termo utilizado para descrever os sistemas de classificação segundo a abordagem *bottom-up*, através do uso de *tagging* social. Gene Smith atribui quatro características às folksonomias, nomeadamente (Smith, 2008):

- *Tagging* de uma forma independente;
- Tags agregadas;
- Relacionamentos inferidos;
- Qualquer método de inferência é válido (embora alguns sejam melhores do que outros).

As características apresentadas são descritas de seguida. Os utilizadores são livres de escolher as tags que quiserem, mas poderão existir sugestões para ajudar os utilizadores a adicionar tags de forma mais fácil. Agregando as tags de muitos utilizadores, é criada uma folksonomia, pois a agregação significa juntar todas as tags automaticamente. Criar uma amostra de tags manualmente é criar uma taxonomia conceito distinto de folksonomia. Note-se que folksonomia é a junção das palavras pessoas (*folks*) e taxonomia (*taxonomy*). Numa folksonomia as relações entre tags são inferidas através de padrões de uso. Por exemplo, na loja *online* Amazon as tags “*thriller*” e “*adventure*” são usadas em alguns livros, como se pode verificar na Figura 5. Estas tags podem não ter nenhuma relação semântica mas estão relacionadas de alguma forma através da utilização conjunta em vários recursos.

thriller

Products (22,663) Contributors (32,032)

Narrow Your Results

3,987 products tagged (showing 1 - 20)

< Previous 1 2 ... 200 Next >

Showing products tagged

thriller and also

adventure

> Clear all filters

Narrow by popular tags

action (2167)

suspense (1635)

mystery (1315)

fiction (923)

science fiction (825)

fantasy (688)

action adventure (647)

1.		2.		3.		4.	
	The Pack ~ Les Turner (Kindle Edition) ★★★★☆ (1) \$6.99 \$0.99		Lethal Experiment ~ John Locke (Kindle Edition) ★★★★★ (133) \$6.99 \$0.99		Faithful Deceptions: A... ~ Melissa Ohnoutka (Kindle Edition) ★★★★★ (12) \$2.99 \$2.99		The Birth of the Peacekeepers. ~ Ricky Sides (Kindle Edition) ★★★★☆ (43) \$1.99 \$1.99

Figura 5 - Produtos da Amazon com as tags *thriller* e *adventure*

Existem formas de análise de tags para serem inferidas relações semânticas com base no comportamento de marcação. Uma das formas poderá ser com a contagem de utilização das tags, sabendo assim quais as mais populares, para que de uma maneira simples sejam encontrados padrões de marcação. Outra possível análise poderá ser através da frequência de tags utilizadas nos mesmos recursos, para determinar as tags utilizadas em conjunto. Por último, outra possível análise indicada por Smith é a utilização de um algoritmo de *cluster* para agrupar tags que tenham uma alta probabilidade de aparecerem juntas (Smith, 2008).

Com a Web 2.0 e através de vários serviços *online*, os utilizadores podem adicionar conteúdos, desde fotos, vídeos, texto, entre outros. Estes conteúdos gerados pelos utilizadores, o desenvolvimento de serviços de informação colaborativa e o uso de folksonomias são mutuamente dependentes. Isso significa que quantos mais serviços de informação colaborativa houver, mais conteúdos são gerados pelos utilizadores e mais tags haverá nas folksonomias (Peters, 2009). Também o número crescente de recursos de informação acrescentados pelos utilizadores necessita de serviços de informação cada vez mais colaborativos para indexação dos recursos armazenados e assim ser possível recuperá-los. Os utilizadores têm interesse em recuperar informação usando termos com os quais estão familiarizados e este processo de classificação pode permitir essa funcionalidade. O sucesso deste tipo de serviços faz com que existam mais recursos e com maior indexação, o que cria um caminho com acesso aos recursos de várias formas através das tags (Peters, 2009).

2.2.3 Vantagens e desvantagens

As características das folksonomias podem ser apontadas como vantagens ou desvantagens ao uso das folksonomias, algumas das mais relevantes são as seguintes:

- A maior vantagem no uso das folksonomias é a parte social, em que os próprios utilizadores podem organizar conteúdos dos recursos da Web e partilhar as tags com outros utilizadores. Isto faz com que o utilizador encontre outros utilizadores com os mesmos interesses, o que permite que sejam criadas comunidades com interesses comuns;
- Outra vantagem a apontar é não existir normalmente controlo do vocabulário usado na criação de tags, assim os utilizadores podem expressar livremente algo sobre o conteúdo dos recursos. Isto permite que haja liberdade de expressão e que os diversos factores, entre os quais intelectual e cultural das pessoas sejam contemplados;
- Ainda outra vantagem é os recursos etiquetados estarem disponíveis na Web, assim pode-se aceder a essa informação através de qualquer computador com ligação à Internet. Portanto, facilmente encontram-se outra vez os recursos marcados;
- A desvantagem do uso das folksonomias é não existir controlo do vocabulário usado na criação de tags. A liberdade na atribuição de tags faz com que haja pouca precisão na recuperação da informação, pois a mesma tag pode ter significados diferentes o que torna muitas vezes as tags ambíguas, demasiado personalizadas e inexactas. Para diminuir este problema os sistemas que usam folksonomias podem tentar melhorar a qualidade na indexação dos recursos, uma possibilidade é existir no sistema a funcionalidade de recomendação de tags durante a atribuição como no sistema Del.icio.us.

2.2.4 Folksonomias gerais e específicas

As folksonomias permitem que os recursos tenham tags associadas que fazem com que os utilizadores encontrem novamente os recursos ou que outros utilizadores encontrem os recursos.

Normalmente os autores distinguem dois tipos de folksonomias: as gerais e as mais específicas, mais conhecidas como folksonomias *broad* e *narrow*.

Os dois tipos são distinguidos pelas seguintes características (Sen, et al., 2006):

- **Folksonomia geral:** Atribuição de tags pelos utilizadores aos vários recursos (por exemplo, o sistema Del.icio.us);
- **Folksonomia específica:** Conjunto de tags por recurso, normalmente tags criadas pelo próprio autor do recurso, mas permite que outros utilizadores atribuam tags (por exemplo, o serviço Flickr⁶).

Na Figura 6 são apresentados esquemas do funcionamento dos dois tipos de folksonomias. Verifica-se que na folksonomia geral os utilizadores podem adicionar várias tags e cada uma das tags pode ser atribuída várias vezes. Por outro lado, na folksonomia específica é o autor do recurso a atribuir as primeiras tags sendo possível os outros utilizadores atribuírem tags. Para este tipo de folksonomia, o exemplo mostra que o grupo “A”, “C” e “D” não atribuíram tags, mas usaram as tags adicionadas por outros para encontrarem o recurso em causa. A folksonomia específica pode ser mais útil nos casos em que os recursos sejam difíceis de encontrar e assim os próprios autores podem escolher tags que permitam que outros utilizadores encontrem mais facilmente esses recursos.

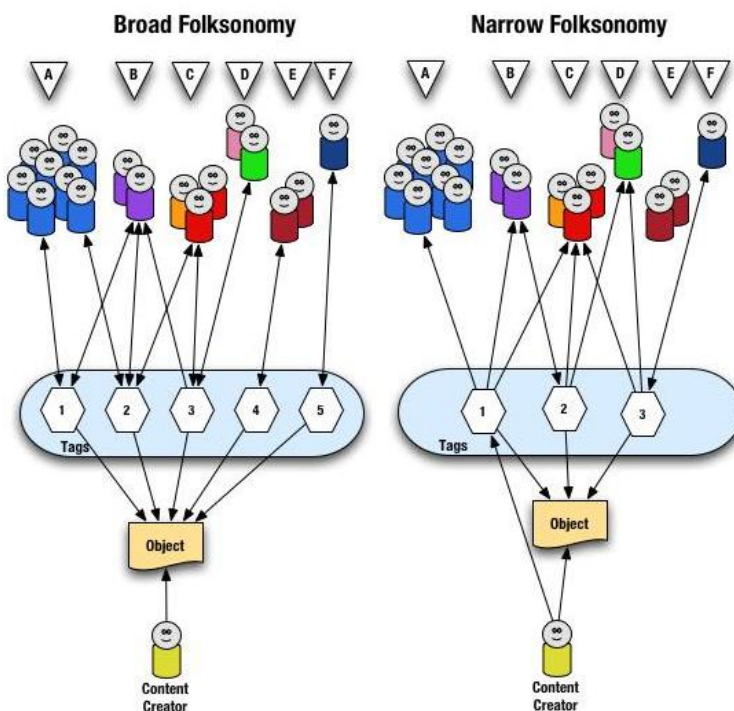


Figura 6 - Ilustração do funcionamento das folksonomias *broad* e *narrow*

Fonte: (Vander Wal, 2005)

⁶ <http://www.flickr.com>

2.3 Serviços de informação colaborativa

Os serviços de informação colaborativa permitem a gestão de recursos pessoais, mas permitem também a criação de dados públicos. Os recursos são indexados pelos utilizadores e essa indexação é construída através do uso de tags na marcação dos recursos. Por isso, os serviços de informação colaborativa também podem ser denominados por sistemas de marcação ou sistemas de *tagging*. Os utilizadores colaboram na criação do serviço de informação, pois fornecem acessos aos recursos e organizam-nos de forma a localizar, partilhar e encontrar informação. A criação de acessos e a organização dos recursos são dois aspectos que distinguem as aplicações de software social, e estas são funções que permitem a criação de uma base de conhecimentos.

O autor Smith distingue cinco categorias de sistemas de *tagging* (Smith, 2008):

- **Gestão de informações pessoais** (por exemplo, email ou galeria de fotografias que permitam *tagging*);
- **Bookmarking social** (por exemplo, o sistema Del.icio.us⁷);
- **Recolha e partilha de objectos digitais** (por exemplo, o serviço Flickr⁸);
- **Comércio na Web** (por exemplo, a loja *online* Amazon⁹);
- **Outros**. Por exemplo, existe um jogo chamado ESP¹⁰. O objectivo do jogo é criar metadados sobre imagens. Neste jogo participam dois utilizadores e os dois vêem a mesma imagem e escolhem tags. Quando chegam a acordo sobre a tag a atribuir seguem para a próxima imagem e pontuam.

A categoria *bookmarking* social e comércio na Web podem ser chamados de serviços de *bookmarking* social, como muitas vezes são conhecidos. Para estas categorias são apresentadas, nos pontos seguintes, pequenas descrições e exemplos, como forma de distinguir a categorização.

2.3.1 *Bookmarking* social

Os sistemas de *bookmarking* social permitem que os utilizadores classifiquem os recursos e partilhem informação com outros utilizadores. O Del.icio.us é um sistema de *bookmarking* social e foi o primeiro sítio Web a disponibilizar a funcionalidade de *tagging*. Este serviço *online* permite adicionar e pesquisar sítios Web sobre qualquer assunto e permite partilhar essa informação com outros.

⁷ <http://del.icio.us>

⁸ <http://www.flickr.com>

⁹ <http://www.amazon.com>

¹⁰ <http://www.gwap.com>

A Figura 7 exemplifica uma funcionalidade do sistema Del.icio.us: a possibilidade de pesquisar por uma tag (“software” neste exemplo) com indicação das tags relacionadas.

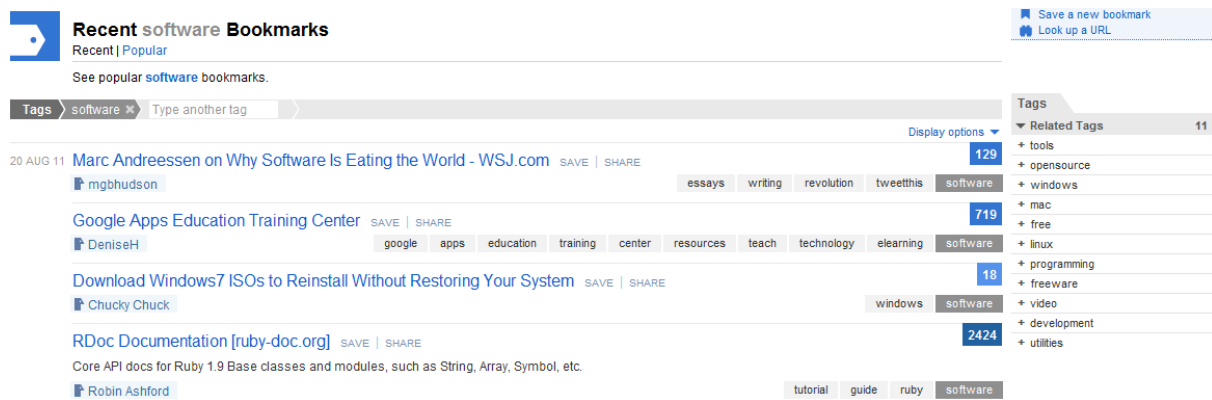


Figura 7 - Pesquisa de tag e apresentação das tags relacionadas no Del.icio.us

No final de Setembro de 2011 o Del.icio.us modificou a imagem, actualmente pesquisar por uma tag (“android” neste exemplo) aparece com o *design* apresentado na Figura 8.

results for 'android'

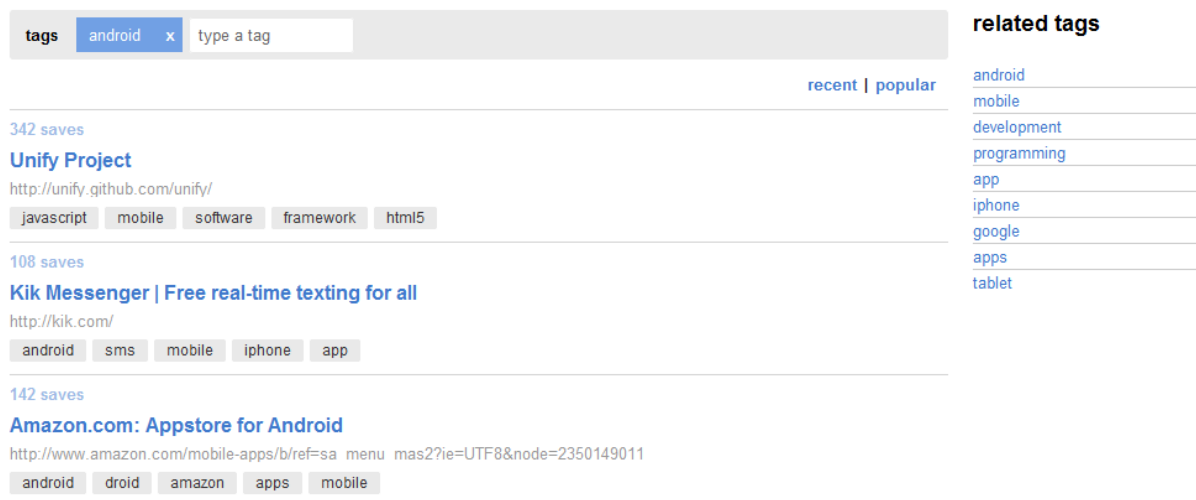


Figura 8 - Pesquisa de tag e apresentação das tags relacionadas no Del.icio.us (novo design)

2.3.2 Comércio na Web

As lojas *online* têm normalmente os produtos separados por categorias, mas também permitem a utilização de tags, o que possibilita a pesquisa de produtos e a navegação utilizando essas tags. Uma loja *online* popular é a Amazon que utiliza a funcionalidade de *tagging*, como se pode verificar na Figura 9. A Amazon é descrita em maior detalhe em outro capítulo deste documento, isto por ter sido o sistema escolhido para realizar o estudo de caso.



Click to **LOOK INSIDE!**

THE #1 NEW YORK TIMES BESTSELLER

MICHAEL CONNELLY

THE LINCOLN LAWYER

THE LINCOLN LAWYER [Mass Market Paperback]
Michael Connelly (Author) | **Like** (3)

★★★★☆ (510 customer reviews) | **Like** (3)

Price: **\$7.99** & eligible for **FREE Super Saver Shipping** on orders over \$25. [Details](#)
[Special Offers Available](#)

In Stock.

Tags Customers Associate with This Product ([What's this?](#))
Click on a tag to find related items, discussions, and people.

Check the boxes next to the tags you consider relevant or enter your own tags in the field below.

<input type="checkbox"/> michael connelly (11)	<input type="checkbox"/> legal mystery (2)	<input type="checkbox"/> la (1)
<input type="checkbox"/> book recommendations (3)	<input type="checkbox"/> california (1)	<input type="checkbox"/> los angeles (1)
<input type="checkbox"/> courtroom drama (3)	<input type="checkbox"/> dead souls and dark alleys (1)	Agree with these tags?
<input type="checkbox"/> crime and punishment (2)	<input type="checkbox"/> fiction (1)	See all 13 tags...

Search Products Tagged with
 GO

[See most popular Tags](#)

Figura 9 - Um produto com as tags associadas e a possibilidade de adicionar ou pesquisar tags na Amazon

3 Análise de sentimentos

Na Web, várias opiniões são disponibilizadas sobre muitos assuntos, pois as pessoas expressam as suas opiniões e comentários durante a utilização de serviços *online*. Estas opiniões podem ser muito importantes e podem ser compreendidas e analisadas com propósitos distintos.

No presente capítulo descrevem-se alguns objectivos da análise de sentimentos, bem como as aplicações mais comuns. Referem-se duas categorias de análise de sentimentos, as técnicas de aprendizagem de máquina e as técnicas simbólicas. Importantes na classificação da polaridade são as considerações relevantes no processo de classificação, estas são descritas neste capítulo. São também abordadas com alguma profundidade as tags sentimentais ou de opinião, que são as tags com polaridade associada. Ainda neste capítulo apresentam-se alguns estudos da área relevantes para o trabalho.

3.1 Objectivos

A área de análise de sentimentos tem como objectivo reconhecer e extrair opiniões, sentimentos e emoções expressas em texto. As opiniões normalmente são classificadas relativamente à polaridade (positiva, negativa e neutra) e intensidade (Pang & Lee, 2008).

Portanto, esta área caracteriza-se pela aplicação de técnicas que permitem atingir um ou vários dos seguintes objectivos:

- **Identificar opiniões:** separar o que são opiniões do que não são opiniões (o que não são opiniões têm polaridade neutra);
- **Avaliar a polaridade das opiniões:** com os dados disponíveis classificar como positiva ou negativa;
- **Avaliar a intensidade da polaridade das opiniões:** identificar o grau de negatividade ou positividade;
- **Classificar documentos:** classificação de acordo com os pontos de vista ou perspectivas;
- **Identificar emoções ou humor:** encontrar emoções ou humor em textos.

3.2 Áreas de aplicação

Com a crescente geração de conteúdos na Web torna-se útil analisar as opiniões das pessoas, de forma a permitir mostrar e compreender as opiniões dos outros utilizadores. A análise de sentimento pode ser aplicada a diversas áreas, entre as quais:

- **Económica e institucional:** empresas e consumidores. Por exemplo, análise de empresas na bolsa de valores ou análise de um produto de uma empresa através da opinião dos utilizadores sobre o produto;
- **Política:** entidades, partidos políticos e eleitores. Por exemplo, análise de políticos, pois os eleitores podem conseguir identificar qual a opinião de outros eleitores sobre um determinado candidato político;
- **Social:** tendências de opinião, identificação de comunidades e outros. Por exemplo, análise de lugares, pois uma pessoa que vai viajar pode utilizar as opiniões de outras para fazer escolhas sobre o roteiro da viagem.

3.3 Técnicas

Nesta secção são apresentadas duas categorias de análise de sentimentos: técnicas de aprendizagem de máquina e simbólicas. De referir que também a estatística pode ser uma técnica utilizada para análise de sentimentos. Sistemas de análise estatística podem ser utilizados para encontrar padrões e explicar os padrões através de modelos estatísticos, em conjunto com outras técnicas de análise de sentimentos.

3.3.1 Técnicas de aprendizagem de máquina

A aprendizagem de máquina é um ramo da inteligência artificial de desenvolvimento de algoritmos que permitem a automatização de um processo de aprendizagem. A aprendizagem é possível através de inferência indutiva baseada na observação de dados que representam informações incompletas sobre a estatística, existindo reconhecimento de padrões.

Podem ser utilizados dados de treino para treinar um classificador para reconhecimento de sentimentos em textos. Esta abordagem com necessidade de recorrer a dados de treino é chamada de aprendizagem supervisionada (Boiy, Hens, Deschacht, & Moens, 2007).

A aprendizagem supervisionada recorre aos dados de treino mas estes podem não estar disponíveis e construí-los demora bastante tempo. Quando não se recorre a dados de treino, a aprendizagem chama-se aprendizagem não supervisionada. Estas podem ser usadas para construir os dados de treino que posteriormente podem ser usados pela aprendizagem supervisionada (Boiy, Hens, Deschacht, & Moens, 2007).

3.3.2 Técnicas simbólicas

As técnicas simbólicas de análise de sentimentos usam recursos léxicos ou regras.

Através do uso de recursos léxicos avaliam-se as palavras de forma a determinar a polaridade das mesmas, desta forma as palavras são consideradas individualmente. Para completar esta abordagem podem ser calculados os valores para determinar o sentimento nas frases através da agregação de palavras.

A técnica apresentada não contempla as relações das palavras na frase, por isso pode realizada a avaliação das relações entre as palavras e analisar a componente gramatical da frase, assim esta é uma abordagem lexical e gramatical (Boiy, Hens, Deschacht, & Moens, 2007).

3.3.2.1 Recursos léxicos

Alguns métodos têm sido explorados para classificação de sentimentos, estes abrangem técnicas de processamento de linguagem natural, recuperação de informação e recursos léxicos.

Os recursos léxicos são compostos por um grande conjunto de palavras de uma determinada língua e com polaridade atribuída. Alguns dos recursos léxicos consideram que a mesma palavra pode ter variados significado e classes gramaticais diferentes, portanto podem ter várias polaridades atribuídas.

Os recursos léxicos permitem a identificação da polaridade em textos, de seguida são apresentados alguns:

- **General Inquirer**¹¹: A versão java *online* do General Inquirer¹² identifica cerca de 13000 palavras e utiliza 6336 regras. Podem ser atribuídas várias tags a cada palavra, estas tags são consideradas categorias no recurso léxico. Exemplos de tags que podem ser atribuídas às palavras: “*Pos*” (positiva) e “*Neg*” (negativa);
- **OpinionFinder**¹³: Este sistema processa documentos e identifica as opiniões e sentimentos presentes no texto bem como o tipo de polaridade (fraca ou forte) e classe gramatical. Parte integrante deste sistema é o Subjectivity Lexicon constituído por palavras classificadas quanto à polaridade e este contém 8221 entradas de palavras;
- **SentiWordNet**¹⁴: Este recurso além de determinar a polaridade identifica a intensidade da polaridade, esta pode ser: fraca, média ou forte (Esuli & Sebastiani, 2006). A versão actual deste recurso, SentiWordNet 3.0, atribui a intensidade da polaridade através de um valor numérico, por exemplo a tag “*good*” tem intensidade positiva de 0.875 (entre 0 e 1)

¹¹ <http://www.wjh.harvard.edu/~inquirer/>

¹² <http://www.webuse.umd.edu:9090/>

¹³ <http://www.cs.pitt.edu/mpqa/>

¹⁴ <http://sentiwordnet.isti.cnr.it/>

(Baccianella, Esuli, & Sebastiani, 2010). Também permite verificar as classes gramaticais das palavras. O recurso disponível para descarregamento (através de registo) tem cerca de 117659 entradas de palavras e existe uma versão disponível *online*.

É conhecido que os dois últimos recursos consideram a possibilidade das palavras terem mais do que uma polaridade ou intensidade de polaridade, muitas vezes estas considerações estão relacionadas com a classe gramatical das palavras ou com o significado das mesmas. Portanto, estes recursos podem ter várias entradas para as mesmas palavras.

Na Tabela 1 são apresentados os resultados das classificações da palavra *funny* atribuídas pelos recursos léxicos descritos. O primeiro recurso atribuiu tags para classificar a palavra, estas significam que a polaridade é positiva e indica uma avaliação de aprovação (*virtue*). O recurso Subjectivity Lexicon tem duas entradas para esta palavra, uma para identificar a palavra como adjetivo e outra para dizer que pode estar em qualquer parte escrita (*anypos*), ambas foram classificadas com polaridade positiva. No recurso SentiWordNet 3.0 existem várias entradas para a mesma palavra, estas têm significados e classes gramaticais diferentes (“a” para adjetivo e “n” para nome), a negrito encontra-se a intensidade relativa à polaridade. Por exemplo, para a primeira entrada a intensidade positiva é 0.125 e a negativa é 0.375, contudo o significado atribuído nesta situação é relativo a desvio do habitual, enquanto com o significado de cómico tem só uma polaridade positiva de 0.5. Este recurso disponibiliza o significado de cada grupo de palavras, este não é apresentado devido ser extenso.

Tabela 1 - Classificação da tag *funny* em diferentes recursos léxicos

General Inquirer (versão <i>online</i>)				Subjectivity Lexicon	SentiWordNet 3.0
Tag	N	%	Words		
Pos	1	100	FUNNY#1=1	type=strongsubj len=1 word1=funny pos1=adj stemmed1=n priorpolarity=positive	a 00968010 0.125 0.375 singular#2 rummy#1 rum#1 queer#1 peculiar#1 odd#4 funny#2 curious#1
Pstv	1	100	FUNNY#1=1	type=strongsubj len=1 word1=funny pos1=anypos stemmed1=y priorpolarity=positive	a 01265308 0.5 0 risible#1 mirthful#2 laughable#2 funny#1 comical#1 comic#1 amusing#2
Virtue	1	100	FUNNY#1=1	type=strongsubj len=1 word1=funny pos1=anypos stemmed1=y priorpolarity=positive	a 01917594 0 0.5 suspicious#2 suspect#1 shady#3 funny#3 fishy#2 a 02545023 0 0.625 funny#4 n 06779310 0 0 good_story#1 funny_story#1 funny_remark#1 funny#1

3.4 Classificação da polaridade

Esta secção pretende descrever as situações a considerar na classificação e análise da polaridade em textos.

Na análise de sentimentos identificam-se as opiniões positivas e negativas, emoções e avaliações (Wilson, Wiebe, & Hoffmann, 2005). Os mesmos autores dizem que é necessário considerar a polaridade contextual da frase, pois a polaridade da palavra pode ser diferente se esta estiver sozinha ou quando aparece inserida numa frase. As palavras podem ter polaridade positiva, mas podem não ser usadas para expressar sentimentos positivos (por exemplo, “*compromise*” tem um sentido positivo, mas não expressa sentimento). Pode também existir o uso da negação que provoca polaridade negativa contextual (por exemplo, “*not good*”). Se for uma expressão pode não ser usada para expressar sentimento, assim a polaridade contextual será neutra. Devem então ser considerados vários aspectos na análise de frases.

Outros aspectos a considerar na análise da polaridade:

- Negação local (por exemplo, “*not adorable*”);
- Dependências através de negação da proposição (por exemplo, “*does not look very adorable*”);
- Negação do sujeito (por exemplo, “*no one thinks that it’s adorable*”);
- Certas frases podem conter palavras de negação, mas para intensificar (por exemplo, “*not only adorable but much adorable*”).

Além dos aspectos descritos e a considerar para determinar a polaridade, alguns outros podem trazer problemas à classificação da polaridade:

- Uso de sarcasmo e ironia nos textos;
- Textos com palavras mal escritas ou sintaxe errada;
- Uso de termos que podem ser consideradas expressões que só numa determinada comunidade é reconhecida (por exemplo, termos usuais somente no país do utilizador);
- Uso de abreviaturas.

3.5 Tags de opinião ou sentimentais

Na análise da polaridade das tags é também necessário considerar os aspectos anteriormente descritos na sua classificação relativamente à polaridade.

Em sistemas que permitem a atribuição de tags aos recursos, não existe habitualmente um vocabulário controlado ou categorias pré-definidas. As tags são escolhidas de forma livre, o que pode ser considerada uma vantagem relativamente a outras formas de classificação (Sen, et al., 2006).

Alguns recursos são mais propensos a despertar sentimentos nos utilizadores (Newhagen, 1998) e outros autores falam sobre tags de opinião ou sentimentais (Yanbe, Jatowt, Nakamura, & Tanaka, 2007). Relativamente a este tipo de tags foi desenvolvido um estudo em 2007 que analisou o serviço de *bookmarking* social Japonês Hatena Bookmarks, que identificou as 54 tags sentimentais mais usadas, divididas em positivas e negativas por frequência, como se pode verificar na Figura 10.

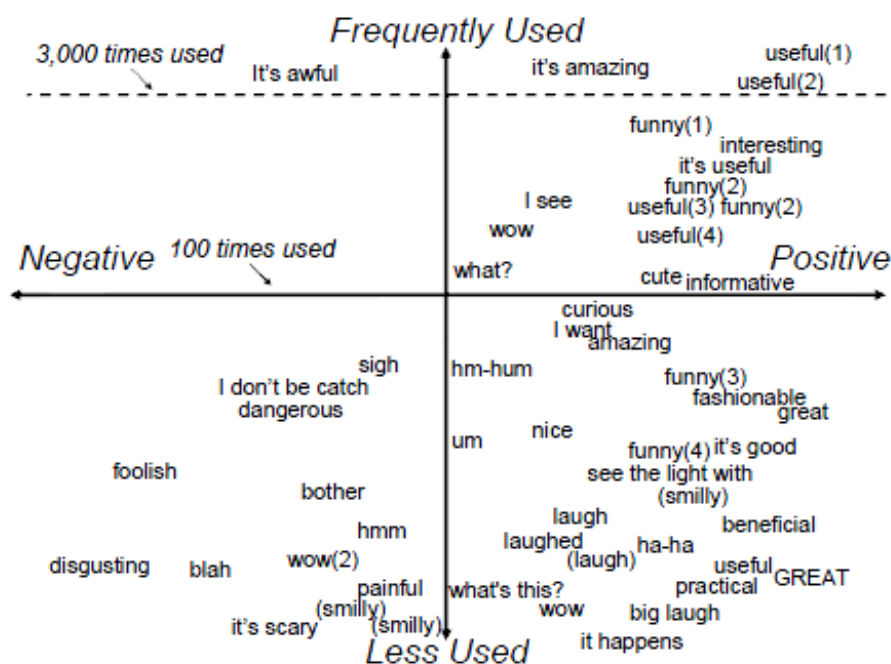


Figura 10 - Tags sentimentais utilizadas no serviço social de marcação japonês Hatena Bookmarks

Fonte: (Yanbe, Jatowt, Nakamura, & Tanaka, 2007)

Os autores do estudo concluíram que as tags negativas são menos usadas que as positivas, mas dados da Tabela 2 sugerem que o uso das tags de opinião está a aumentar. Os dados de 2007 apresentados nessa tabela foram recolhidos por o autor Zollers da loja *online* Amazon (Zollers, 2007) e foram também recolhidos dados de 2011 durante este estudo. Aparentemente cada vez mais os utilizadores, ao se depararem com um recurso que lhes desperta um sentimento positivo ou negativo, utilizam uma ou mais tags para expressar a sua opinião.

Tabela 2 - Dados sobre o uso de algumas tags de opinião na Amazon

Tag de opinião	Dados da utilização das tags nos produtos (2007)	Dados da utilização das tags nos produtos (2011)
<i>awesome</i>	691 utilizadores usaram a tag em 855 produtos	7748 utilizadores usaram a tag em 4822 produtos
<i>beautiful</i>	299 utilizadores usaram a tag em 356 produtos	4516 utilizadores usaram a tag em 3332 produtos
<i>cool</i>	404 utilizadores usaram a tag em 574 produtos	3901 utilizadores usaram a tag em 3295 produtos
<i>crap</i>	663 utilizadores usaram a tag em 698 produtos	-
<i>garbage</i>	433 utilizadores usaram a tag em 391 produtos	4628 utilizadores usaram a tag em 3396 produtos
<i>great</i>	461 utilizadores usaram a tag em 866 produtos	5631 utilizadores usaram a tag em 5826 produtos
<i>lame</i>	138 utilizadores usaram a tag em 166 produtos	1697 utilizadores usaram a tag em 1417 produtos
<i>boring</i>	-	3399 utilizadores usaram a tag em 3422 produtos

3.6 Estudos sobre opinião

A opinião dos utilizadores pode influenciar outros utilizadores, nesta secção são apresentados alguns estudos realizados sobre análise de sentimentos.

3.6.1 Estudos relacionados com a Amazon

São apresentados alguns estudos sobre a opinião realizados no âmbito da loja *online* Amazon.

O estudo dos autores Li e Hitt analisa como as preferências dos compradores pode influenciar a longo prazo as compras de novos utilizadores. Neste estudo são usados dados sobre os livros da Amazon. Estes dizem que os utilizadores consideram a classificação e opinião dos compradores para avaliar a qualidade dos livros com base nas classificações atribuídas através de estrelas e dos dados de vendas dos livros da Amazon (Li & Hitt, 2008).

Na Amazon é necessário escrever um comentário quando o utilizador fez uma compra e pretende atribuir uma classificação ao produto, mesmo com este trabalho adicional os compradores manifestam

a sua opinião. Os autores Wu e Huberman realizaram dois estudos sobre a expressão de opinião *online*: um aos livros mais vendidos da loja *online* Amazon e outro aos filmes do sistema IMDB¹⁵ para compreenderem como as opiniões *online* são criadas e evoluem ao longo do tempo. Estes indicam que os compradores têm mais vontade de expressar a sua opinião se ficaram muito satisfeitos ou muito insatisfeitos com a compra do produto. Nestas situações, os utilizadores não se importam de escrever o comentário quando atribuem uma classificação e contribuem com as suas opiniões mesmo depois de já existirem opiniões adicionadas por outros compradores (Wu & Huberman, 2010).

Estes estudos não se referem à expressão da opinião ou análise de sentimentos através do uso de tags. Não foram encontradas descrições na literatura com essa particularidade.

3.6.2 Outros estudos

Na área de análise de sentimentos têm sido realizados alguns estudos mais genéricos sobre a aplicação das técnicas existentes, não apenas relacionados com a Amazon.

O autor Lin e colegas classificaram os documentos de acordo com os pontos de vista defendidos e até 94,93%, dos 594 artigos disponíveis no sítio Web bitterlemons¹⁶, foram classificados correctamente. As publicações bitterlemons permitem troca de opiniões sobre o conflito entre Israel e Palestina. Com este estudo os autores pretendiam descobrir se os computadores podem aprender a identificar a perspectiva de um documento e aplicaram técnicas de aprendizagem para descobrir padrões (Lin, Wilson, Wiebe, & Hauptmann, 2006).

Os autores Yun e Hatzivassiloglou conseguiram distinguir opiniões e factos através das notícias do Wall Street Journal¹⁷, com uma taxa de 91% de sucesso na classificação de frases e uma taxa de 97% de precisão na classificação dos documentos (Yu & Hatzivassiloglou, 2003). Estes apresentam vários modelos para fazer essa distinção entre opiniões e factos e entre opiniões negativas e positivas.

Ainda outro estudo, não menos relevante, o do autor Pang e colegas, sobre as recomendações dos filmes do IMDB. Neste estudo, os dados existentes foram considerados como difíceis para análise de sentimentos nomeadamente quando existe indicação de sentimentos opostos no mesmo texto, mas mesmo assim atingiram uma taxa de sucesso de 82,7% na classificação da polaridade (Pang, Lee, & Vaithyanathan, 2002).

As análises de documentos por pessoas atingiria resultados superiores, mas nem sempre é possível pois normalmente são muitos documentos a serem analisados.

¹⁵ <http://www.imdb.com>

¹⁶ <http://www.bitterlemons.net/>

¹⁷ <http://trec.nist.gov/>

4 Estudo de caso

Em termos de metodologia, este trabalho inclui o que pode ser caracterizado como um estudo de caso (Kumar, 2005). O mesmo foi utilizado para analisar se as tags de opinião podem efectivamente servir para estimar a opinião dos utilizadores em relação aos recursos. Foi escolhida a loja *online* Amazon para o estudo de caso.

Neste capítulo é apresentada a Amazon e as razões para escolha da loja *online* para estudo de caso. Descreve-se o processo de amostragem e respectivo processo de classificação das tags para a amostra considerada.

4.1 Amazon

Neste ponto é apresentada a loja *online* Amazon que foi escolhida para realizar o estudo de caso neste trabalho. São apresentadas as razões da escolha da Amazon.

4.1.1 Caracterização da Amazon

A loja *online* Amazon¹⁸ faz uso do comportamento de compras e navegação dos seus clientes para partilhar as impressões dos produtos com outros utilizadores, desde a sua fundação em 1994. Por exemplo, na Figura 11 é possível visualizar informação sobre as compras efectuadas por quem comprou o livro *The Lincoln Lawyer*.



Figura 11 - Sugestão de produtos com base nas compras de outros utilizadores na Amazon

Além disso, os utilizadores registados podem escrever e publicar opiniões sobre os produtos comprados na Amazon e pontuar os mesmos (1 até 5 estrelas) – ver Figura 12 para um exemplo relacionado com livro mencionado anteriormente.

¹⁸ <http://www.amazon.com>

Customer Reviews



Figura 12 - Pontuações atribuídas a um produto na Amazon

Desde finais de 2005 é possível o uso de tags na Amazon para descrever e categorizar produtos. Só os utilizadores registados podem marcar os produtos, usando as tags que pretenderem. Os utilizadores podem marcar os recursos com tags já utilizadas anteriormente ou podem adicionar novas tags. As tags podem ter mais do que uma palavra. Tem como regra de formatação a conversão das tags para letras minúsculas. A Figura 13 exemplifica a atribuição de uma tag ao livro referido anteriormente.

Tags Customers Associate with This Product [\(What's this?\)](#)

Click on a tag to find related items, discussions, and people.

Check the boxes next to the tags you consider relevant or enter your own tags in the field below.

- | | | |
|---|---|---|
| <input type="checkbox"/> michael connelly (5) | <input type="checkbox"/> adventure (1) | <input type="checkbox"/> friendship (1) |
| <input type="checkbox"/> jake lassiter (2) | <input type="checkbox"/> book recommendations (1) | <input type="checkbox"/> kennedy (1) |
| <input type="checkbox"/> legal thriller (2) | <input type="checkbox"/> character driven (1) | Agree with these tags? |
| <input type="checkbox"/> paul levine (2) | <input type="checkbox"/> courtroom drama (1) | See all 24 tags... |

Your tags:

(Press the "Add" button to add your tag to the list.)

Tag	Usages
adventure	149682 usages
action adventure	37749 usages
adventure scifi	3128 usages
adventure games	3072 usages
action and adventure	3071 usages
adventure fantasy	1991 usages
action- adventure	1763 usages
archaeology adventure	1308 usages
adventure fiction	990 usages
adventure mystery	893 usages
adventures	719 usages
adventures	552 usages
Topic: action adventure movies	414 usages

Figura 13 - Atribuir tag a um produto na Amazon

No seguimento deste exemplo é apresentada na Figura 14 a funcionalidade de sugestões para atribuição de tags na Amazon.



Figura 14 - Funcionalidade de atribuir tag a um produto na Amazon através de sugestões

4.1.2 Escolha da Amazon

Muitos são os serviços *online* com a funcionalidade de *tagging*, mas foi o serviço da Amazon o escolhido. Neste ponto são apresentadas as razões da escolha da Amazon como estudo de caso para este trabalho.

Sendo possível na Amazon classificar produtos através da atribuição de uma a cinco estrelas e através da atribuição de tags, isso possibilita uma base de comparação entre os dois tipos de classificação de produtos. Esta foi uma das razões principais da escolha da Amazon como caso de estudo. As duas formas de classificar produtos permitem uma forma de avaliação dos produtos para venda, visto que se pode considerar que as duas formas baseiam-se na opinião dos utilizadores. Como a opinião dos utilizadores sobre os produtos pode influenciar a compra por parte de outros utilizadores, é importante verificar a coerência entre as duas formas de avaliação.

Note-se que algumas lojas *online*, como a Fnac *online*¹⁹, a Wook²⁰ e a Javari *online*²¹, permitem a adição de comentários para os utilizadores registados e uma atribuição de pontuação aos recursos de 1 (mau) a 10 (muito bom), no caso da Fnac *online*, e de 1 a 5 no caso da Wook, mas não existe a possibilidade de atribuição de tags.

¹⁹ <http://fnac.pt>

²⁰ <http://wook.pt>

²¹ <http://www.javari.co.uk>

Outro factor considerado na escolha da loja *online* Amazon para este estudo de caso foi o facto de ser uma loja conhecida e utilizada mundialmente, o que permite ao estudo uma outra abrangência. Por último, uma outra razão para a escolha da Amazon é a possibilidade de qualquer pessoa registada poder marcar os recursos e não apenas os próprios autores. Logo, não deve existir nenhuma influência de marcação, por tags atribuídas pelos próprios autores. Isto porque em sistemas como o Flickr²², por exemplo, os autores que partilham fotografias normalmente adicionam tags enquanto na Amazon os recursos não são criados por utilizadores da loja *online*.

4.2 Amostragem

A amostragem é o processo de selecção de uma amostra de um grupo maior, para ser uma base para estimar ou prever algo com a informação recolhida. Através da amostra recolhida é possível fazer análises: da previsão de informação desconhecida ou até mesmo fazer estimativas sobre o grupo maior (Kumar, 2005).

A amostragem pode ser tipificada em dois grupos, estes são processos de selecção de uma amostra e são os seguintes:

- **Amostragem aleatória:** neste tipo de amostragem cada elemento do grupo tem a mesma probabilidade de ser seleccionado para pertencer à amostra e a escolha de um elemento não depende da escolha de outro elemento;
- **Amostragem não aleatória:** este tipo de amostragem não usa a probabilidade para escolher elementos, mas depende de outras considerações para seleccionar elementos do grupo.

Neste estudo é necessária uma amostragem, e para a amostra representar com qualidade o conjunto total é necessário usar um bom método para a seleccionar de uma forma correcta e válida, isto é, com diversidade adequada. Quando se aumenta a amostra e esta não significa descoberta de mais informação diz-se que se chegou a um ponto de saturação, em que não se obtém novos dados ou a nova informação é insignificante. Chegado a esse ponto não será necessário aumentar a amostra e está identificado o tamanho máximo ou necessário da amostra.

²² <http://www.flickr.com>

4.2.1 Processo não aleatório

A amostra considerada neste estudo é obtida através de um processo não aleatório, pois existem considerações na escolha dos produtos da loja *online* Amazon.

Podem ser considerados quatro processos não aleatórios diferentes, estes são característicos processos de amostragem. Destes o que melhor identifica o processo usado neste estudo é a amostragem por quotas.

De seguida são apresentados os quatro tipos de amostragem diferentes:

- **Amostragem por quotas:** a amostra é seleccionada através da escolha de elementos com as características necessárias, isto até chegar ao número de elementos necessários. Em primeiro lugar desenvolvem-se categorias de elementos e depois seleccionam-se elementos da amostra com base acidental ou intencional;
- **Amostragem acidental:** os elementos da amostra são escolhidos por conveniência, estes podem ser pessoas conhecidas, mas os resultados desta amostragem aplicam-se só a esta amostra;
- **Amostragem intencional:** amostragem usada para descrever algo ou desenvolver o que é pouco conhecido, para isso o pesquisador usa o julgamento para seleccionar os elementos.
- **Amostragem bola de neve:** processo que usa redes como processo de selecção de elementos para constituir a amostra, de um grupo inicial de elementos identifica-se outros de interesse e assim sucessivamente obtendo-se informações a partir de informações.

Como o autor Kumar indica, na selecção de uma amostra deve-se obter o máximo de precisão nas estimativas pretendidas para uma determinada dimensão da amostra e evitar distorções na selecção da amostra (Kumar, 2005). Então, para o objectivo pretendido não teria interesse considerar na recolha todos os tipos de produtos existentes na Amazon. Além disso, também não tem interesse ser considerada uma amostra com muitos produtos de classificação alta, pois não estaríamos a avaliar todos os tipos de classificação. Relativamente às tags, se os produtos tiverem poucas tags atribuídas, pouca informação pode ser obtida desses produtos. O interesse do estudo não é avaliar o número de tags atribuídas de forma geral, mas terá interesse a análise dos produtos com informação suficiente, isto para avaliar se o uso das tags reflecte a classificação atribuída através de estrelas. Se considerarmos a recolha por produto, por classificação e os que tenham um número significativo de tags, assim são necessárias menos recolhas.

O estudo realizado tem em conta os dois sistemas de classificação utilizados na Amazon. Dos vários tipos de produtos disponíveis na Amazon, como peças de roupa, artigos de joalharia, filmes,

artigos para casa, entre outros, optou-se por incluir no estudo livros, visto ser um dos tipos de produtos há mais tempo comercializado pela Amazon.

Procedeu-se à recolha de tags e pontuações atribuídas (1 a 5 estrelas) a 17 livros. Esses foram escolhidos para a amostra de forma a contemplar as diferentes pontuações possíveis de serem atribuídas. Estrategicamente foram escolhidos, numa primeira instância, livros dos mais comprados e que já estivessem no topo das vendas há algum tempo, pois esses normalmente tinham muitos utilizadores a associar uma pontuação aos artigos, o que era importante para este estudo. Como estes livros tinham maioritariamente uma classificação positiva, decidiu-se numa segunda fase, e depois de avaliadas as diferentes tags que eram atribuídas, recolher também dados de livros com pontuação inferior. Assim, desta forma conseguiu-se garantir que na amostra considerada estavam representados livros com uma pontuação média positiva e negativa. Todos os critérios de restrição da amostra descritos foram considerados no processo não aleatório de recolha da amostra.

O mesmo processo foi seguido para a recolha de tags e pontuações atribuídas de dispositivos electrónicos, isto para aumentar a amostra com produtos de outra categoria. Devido a serem uns dos mais comprados e comentados, esta recolha baseou-se na amostra de 5 auscultadores existentes na Amazon.

Na Figura 15 é apresentado um esquema representativo do funcionamento do processo de amostragem usado para recolha da amostra da Amazon. De forma geral foram seleccionados os produtos com a informação considerada suficiente de tags e pontuações, e para existir uma variedade de produtos, foram escolhidos os produtos com pontuações médias diversas (entre 1 e 5 estrelas).

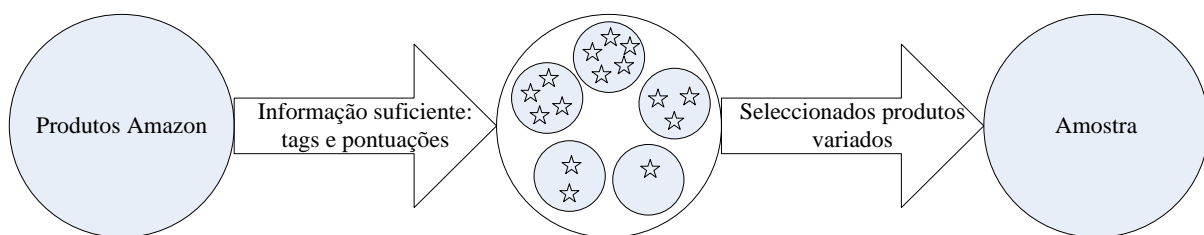


Figura 15 - Representação do funcionamento do processo de amostragem

4.3 Base de dados

Os dados recolhidos sobre os produtos incluídos na amostra da Amazon necessitam de registo para posteriores análises. Foi criada uma base de dados para registo dos dados e preparação da estrutura de forma a contemplar também a classificação das tags apresentada na secção seguinte.

Na Figura 16 apresenta-se o esquema das tabelas usadas para o registo dos dados recolhidos e tratados dos produtos da Amazon.

Existe registo dos dados considerados importantes para o estudo em causa, entre os quais:

- Dados representativos dos produtos (tabela *product*), que inclui: identificação do produto (*ID*), título do produto (*title*), pontuação média atribuída através de estrelas (*stars*), número de utilizadores que atribuíram a pontuação (*reviews*) e departamento à qual o produto pertence nomeadamente, livros ou dispositivos electrónicos (*department*);
- Detalhe dos produtos relativamente à pontuação atribuída (tabela *productNumberStars*), isto é, por tipo de pontuação (1 a 5 estrelas - *numberStars*) indicação do número de utilizadores que atribuíram essa pontuação (*numberReviews*) ao respectivo produto (*ID*);
- Tags atribuídas por produto (tabela *productTag*) com indicação do número de atribuições (*reviewsTag*), em que pode ou não existir uma classificação específica (*classification*) para a tag associada ao produto (*ID*). Estes são os casos em que a polaridade depende do recurso, por exemplo, a tag é classificada como neutra se o conteúdo desta fizer parte do título do recurso e noutros casos a tag pode ter outra polaridade que surge como classificação genérica da tag;
- Classificação das tags (tabela *tags*).

No Anexo 1 são apresentados alguns aspectos relevantes da amostra recolhida da Amazon, de uma forma mais pormenorizada.

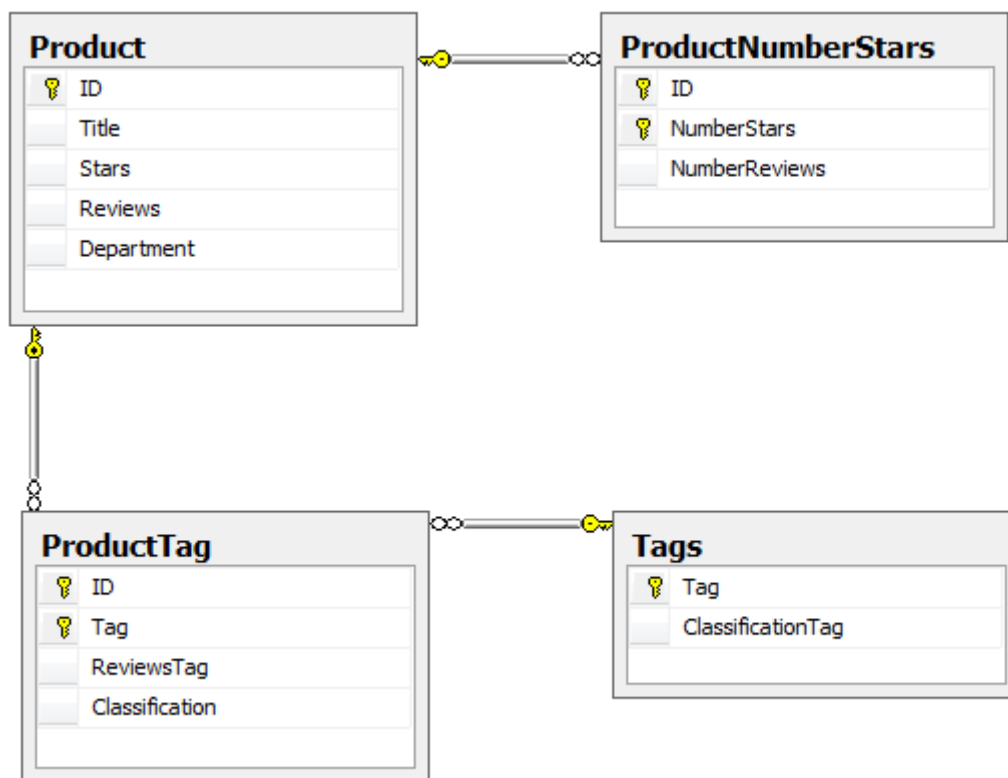


Figura 16 - Esquema das tabelas para registo dos dados recolhidos dos produtos

4.4 Classificação de tags

Nesta secção é apresentada o processo realizado para ser possível a classificação de tags necessária para o estudo em causa. Também são apresentadas as particularidades encontradas nesse processo. A classificação de tags foi realizada através de um processo manual, mas depois do conhecimento de alguns padrões avançou-se para a implementação de um processo semi-automático de classificação de tags.

4.4.1 Processo de classificação de tags

Relativamente à amostra contemplada, foram analisadas as tags recolhidas, de forma a proceder-se à respectiva classificação e diferenciação entre tags de opinião e as tags que não expressavam a opinião sobre os recursos. Quando as tags expressam opinião é necessário determinar a polaridade das mesmas.

De acordo com as categorizações das tags já referidas, concluí-se que para este estudo de caso é importante considerar três grupos principais de tags:

- **Tags positivas**, usadas pelos utilizadores para expressar uma opinião ou avaliação positiva sobre o recurso ou sentimentos de agrado que foram desencadeados. Algumas delas podem estar associadas a acções a realizar pelo utilizador (por exemplo, “*to read list*”), pelas razões que a seguir são apresentadas;
- **Tags negativas**, contrariamente às tags positivas, são usadas pelos utilizadores para expressar uma opinião ou avaliação negativa sobre o recurso ou sentimentos negativos que foram desencadeados;
- **Tags neutras** são todas as que não são positivas nem negativas.

Alguns investigadores referem-se às tags como “*to_read*” como tags funcionais (Kroski, 2005), tags de sinalização (Dennis, 2006), ou ainda tags relacionadas com o tempo e tarefas (Kipp M. E., 2007). Os recursos marcados são susceptíveis de terem sido considerados úteis pois expressam uma resposta do utilizador ao invés de uma declaração do tema do documento, são intrinsecamente sensíveis ao tempo, sugerem um envolvimento activo com o texto no qual o utilizador está a ligar a questão percebida com uma tarefa específica ou um conjunto específico de interesses (Kipp & Campbell, 2006).

O autor Bal identificou mais um grupo de tags nomeadas “*undefined*”, neste trabalho também foi necessário considerar as tags indefinidas ou imprecisas (Bal, 2009). As tags classificadas desta forma são apresentadas no capítulo que apresenta os resultados do estudo. Classificar desta forma as tags desconhecidas ou que o significado depende do contexto, permite que o grau de incerteza do sentido do uso da tag não influencie os restantes grupos de tags.

Algumas questões tiveram que ser consideradas na classificação das tags, entre as quais, a análise da polaridade pois é necessário considerar o contexto relativamente ao recurso, porque a mesma tag pode ter classificação diferente consoante o recurso em questão. Por exemplo, a classificação da tag “lixo” (“*garbage*”) pode estar dependente do recurso em análise, pois se esta for adicionada a um livro sobre reciclagem, a polaridade será neutra devido a tratar-se do assunto do mesmo. No entanto, se aplicada a um livro sobre política estará a ser utilizada para expressar um sentimento negativo. Mas a tag poderá fazer parte da designação do produto, neste caso esta teria de ser classificada como neutra já que pertence a uma característica intrínseca do produto. Também, por exemplo, a tag “hilariante” (“*hilarious*”) se utilizada num livro cómico, pode estar a expressar uma opinião positiva enquanto se utilizada num livro sobre política será para exprimir uma opinião negativa.

Na análise da polaridade das tags foi considerada a polaridade contextual da frase, entre as quais, a identificação do uso da negação, existente nas tags constituídas por várias palavras (por exemplo, “*not good*”), para identificação das tags negativas.

Na classificação das tags participaram duas pessoas e apenas quando ambas concordavam com a polaridade a atribuir é que a mesma se aplicava. Em casos de dúvida, esta foi classificada como neutra. Assim o grau de incerteza quanto à classificação de tags de opinião diminuiu. De referir que cada uma das tags foi analisada e interpretada para garantir que esta ficaria bem classificada. Para isso, foi necessário utilizar dicionários *online* e até mesmo pesquisas *online* para tentar perceber o sentido da atribuição das tags, isto para ser atribuída a classificação às mesmas.

4.4.2 Classificação semi-automática

Como parte integrante deste estudo, foi criado um classificador que através das tags associadas aos recursos classifica-os de uma forma quase automática. Assim, consegue-se obter uma base de opiniões sobre os recursos, de forma a ser possível proceder-se a análise dos dados.

Como parte do mecanismo de classificação automática foi usado um recurso léxico, dos vários recursos disponíveis e já identificados no ponto 3.3.2.1 deste documento foi escolhido o Subjectivity Lexicon (parte integrante do OpinionFinder) devido a estar disponível um ficheiro para descarregamento com os termos e a sua classificação. Outra razão da sua escolha é a sua simplicidade em relação aos outros recursos léxicos. Poucos termos existiam em classes gramaticais diferentes com polaridade diversa, os casos encontrados foram tratados para o contexto em questão, isto é, de forma que faça sentido na utilização de tags. O tratamento do conteúdo do recurso para ser interpretada a classificação das palavras foi necessário, para ficar disponível para uso na classificação das tags.

A polaridade dos termos pode ser positiva, negativa, ambas, ou neutro. Wilson e colegas indicam situações em que são utilizados termos com polaridades diferentes, fornecendo também alguns exemplos transcritos (Wilson, Wiebe, & Hoffmann, 2005):

- **Termos com polaridade positiva** descrevem ou caracterizam emoções (por exemplo, “*I’m happy*”), avaliações (por exemplo, “*Great idea!*”) e opiniões (por exemplo, “*She supports the bill*”) positivas;
- **Termos com polaridade negativa** caracterizam emoções (por exemplo, “*I’m sad*”), avaliações (por exemplo, “*Bad idea!*”) e opiniões negativas (por exemplo, “*She’s against the bill*”);
- **Dupla polaridade** aparece em expressões que têm polaridade positiva e negativa (por exemplo, “*good and evil*”);

- **Polaridade neutra** nos restantes casos: onde existe, por exemplo, especulação ou quando a polaridade não é positiva nem negativa.

Os tipos de classificação citados são as classificações incluídas no recurso léxico. Relativamente à dupla polaridade, esta classificação é semelhante à considerada na classificação já descrita como tags indefinidas ou imprecisas (“*undefined*”).

A integração e utilização do recurso léxico foi útil na classificação das tags, mas foi necessário rever a atribuição de tags, pois muitas das palavras existentes têm, por exemplo, um sentido positivo, mas não servem para exprimir uma opinião positiva relativamente a um recurso, como é o caso da palavra “*promise*” pertencente ao recurso léxico.

Outra situação a considerar é a referida polaridade contextual das tags, pois estas podem ser constituídas por várias palavras em vez de uma única palavra (por exemplo, “*new york times bestselling author*”), esta é uma dificuldade para uma classificação de tags automática, visto que, o recurso léxico contém apenas palavras e não frases. Quando numa tag composta por várias palavras existe, por exemplo, uma classificação positiva e neutra, a classificação final da tag é positiva. Foi também necessário considerar o uso da negação (por exemplo, “*no*” e “*not*”) nas tags compostas por várias palavras.

Outro caso a considerar é, por exemplo, quando uma tag positiva usada em um certo recurso pode ser neutra isto devido a representar o assunto do mesmo, estes casos pode-se identificar através do título do recurso. Esta identificação torna-se importante porque alguns autores identificaram que muitas vezes as tags utilizadas num recurso provêm do título do mesmo (Heckner, Neubauer, & Wolff, 2008).

Depois de referidos os casos contemplados no processamento da classificação automática de tags verifica-se que são necessárias duas fases: uma para classificação de tags de forma automática e uma segunda fase para revisão da classificação de forma manual.

Na Figura 17 está representado de forma geral o funcionamento do classificador automático de tags desenvolvido, este processamento é despoletado quando são inseridas tags associadas a produtos. O processo usa duas fontes: as tags já classificadas e os dados do recurso léxico. Para as tags associadas a produtos identifica quais já estão classificadas e em conjunto com a consulta ao recurso léxico aplica as regras para classificar as tags.

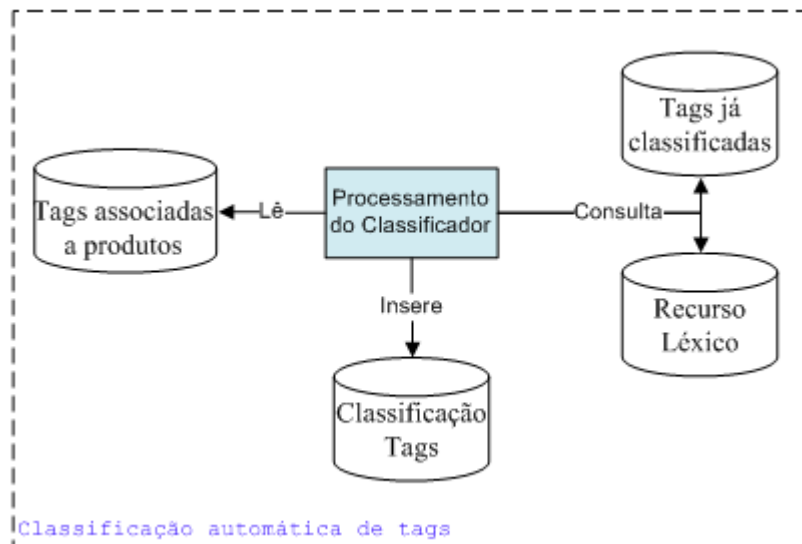


Figura 17 - Funcionamento do classificador automático de tags

Na Figura 18 estão representadas umas tabelas auxiliares existentes para classificação das tags, que distinguem-se da seguinte forma:

- Classificação temporária a ser validada, as tags classificadas pelo classificador são incluídas nesta tabela para passar pela validação manual. Depois de confirmada a classificação atribuída automaticamente, são actualizadas as tabelas definidas;
- Registo da classificação já atribuída às tags inseridas. Esta tabela permite validar se a mesma classificação genérica já atribuída à tag aplica-se ao recurso em questão, desta forma evita-se erros de classificação de tags. Logo este registo também é utilizado para validação da classificação, neste caso de classificação já atribuída;
- Tabela com a polaridade das palavras do recurso léxico que serve de apoio ao classificador.



Figura 18 - Tabelas auxiliares para a classificação das tags

Foram encontrados alguns problemas na utilização do classificador de forma totalmente automática, estes são apresentados de seguida:

- A base de dados com as tags classificadas ainda é pequena, logo não são muitas as tags já classificadas anteriormente antes da execução do classificador;
- Existindo revisão manual da classificação consegue-se identificar o contexto de utilização da tag e identificar a classificação consoante o recurso em questão, por exemplo através do assunto ou categoria do produto;
- O recurso léxico tem algumas limitações pois este não inclui todas as palavras que constituem as tags usadas e ainda a polaridade atribuída a algumas palavras pode ser aplicada de outra forma a tags. Por exemplo, a palavra poluidores (“*polluters*”) normalmente não será aplicada para expressar um sentimento, mas apenas como referência, mesmo assim no recurso léxico esta palavra tem polaridade negativa;
- Com a revisão manual existe a possibilidade de encontrar outros problemas de identificação da polaridade.

Foi utilizado o sistema de gestão de base de dados Microsoft SQL Server para criar a base de dados da amostra da Amazon. Todo o processo de classificação semi-automática foi realizado através deste sistema.

5 Análise dos resultados

A interação dos utilizadores influenciam outros utilizadores, isto está relacionado com o que o autor Stanam Alag chama de inteligência colectiva. Inteligência colectiva é quando pessoas interagem com outras pessoas e surge inteligência ou comportamento que de outra forma não existia (Alag, 2009). A partir da utilização de tags e da atribuição de pontuações aos recursos é possível extrair inteligência, até porque essas opiniões podem influenciar outros utilizadores. Através da inteligência extraída da utilização de tags e pontuações atribuídas pelos utilizadores é possível a realização de análises a esses conteúdos que servem para expressar a opinião dos utilizadores sobre os recursos da Amazon.

Neste capítulo são apresentadas algumas análises e avaliações aplicadas aos dados da base de dados construída através da amostra recolhida da loja *online* Amazon.

Uma das análises que será realizada é análise da relação entre os dois sistemas de classificação existentes na Amazon, esta será avaliada através da correlação de Spearman. O coeficiente de correlação de Spearman permite avaliar a dependência entre duas variáveis. Através dessa avaliação consegue-se determinar o grau de correlação entre as variáveis. A correlação (R_s) pode assumir valores entre -1 e +1 e o sinal da correlação indica a direcção da correlação entre as variáveis. Quanto mais perto R_s está de zero, mais fraca é a correlação entre as variáveis. É uma correlação com significado estatístico quando superior a ,5 (ou -,5) existindo associada a probabilidade de erro (p). As análises apresentadas neste capítulo, foram realizados através da ferramenta SPSS²³ (Statistical Package for the Social Sciences).

5.1 Número de palavras nas tags

A primeira análise será avaliar a normal utilização das tags na Amazon. Através da amostra recolhida é possível analisar as tags existentes e concluir qual é, de forma geral, a constituição das tags utilizadas na Amazon.

Para a amostra considerada, foram analisadas as tags de forma a verificar o número de palavras que constituem estas. Através do gráfico da Figura 19 verifica-se que são mais utilizadas as tags constituídas por mais do que uma palavra.

²³ <http://www.ibm.com/software/analytics/spss/products/statistics>

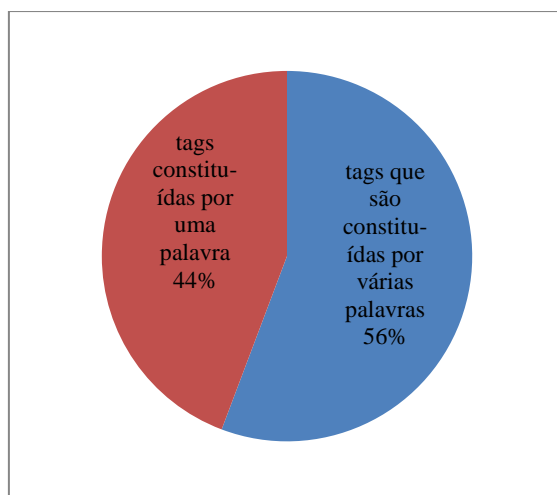


Figura 19 - Percentagem de tags constituídas por uma ou várias palavras na amostra da Amazon

Sendo mais utilizadas as tags que são constituídas por várias palavras podemos analisar as quantidades de palavras mais atribuídas. Na Figura 20 apresenta-se um gráfico que indica a percentagem de tags por número de palavras utilizadas. Verifica-se que, para a amostra considerada, existem tags com até 12 palavras.

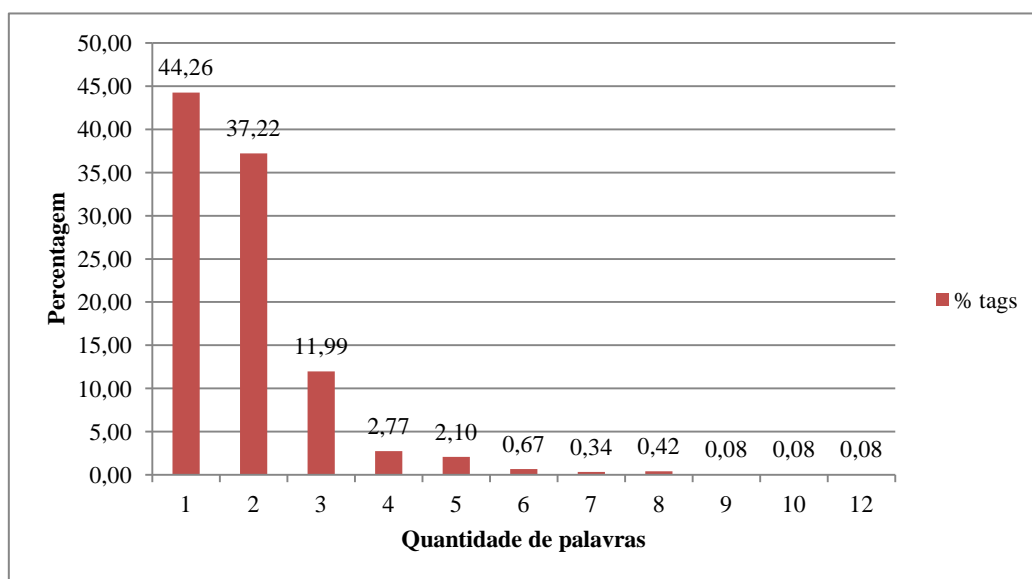


Figura 20 - Representação da percentagem de tags por quantidade de palavras usadas na amostra da Amazon

Algumas das tags constituídas mais palavras: *“how long can you read it with a straight face”*, *“ill bring tinker is headed off to the office - twill be another”* e *“skye the troll and other fairy tales for children”*. Todas as tags referidas são neutras relativamente à classificação.

As tags compostas por hífenes ou *underscores* não são corrigidos pela Amazon. Nesta amostra existem 56 tags com um ou mais hífenes que correspondem a 1,34% das tags consideradas. Relativamente ao uso de *underscores*, só uma das tags é que contém este símbolo. Estes símbolos ou espaços não são utilizados normalmente como forma de separação de tags na amostra considerada.

5.2 Contabilização de tags e pontuações

Com os dados recolhidos foi verificado se são mais atribuídas tags ou pontuações na Amazon.

Na Tabela 3 é apresentada a contabilização de tags e pontuações atribuídas por produto. Pode-se observar que em mais de metade dos produtos da amostra o número total de tags atribuídas é superior ao número total de pontuações (entre 1 a 5) atribuídas, esta observação é mais perceptível através da visualização do gráfico da Figura 21. Na mesma tabela referida é apresentado ainda o total de tags, mas só contabilizando as tags de opinião, neste caso já se verifica que somente em dois produtos este número é superior ao número total de pontuações atribuídas. De qualquer forma, verifica-se que a atribuição de tags é frequentemente utilizada, pelo menos comparativamente à atribuição de pontuações aos produtos da Amazon, isto na amostra considerada neste trabalho. Ainda na mesma tabela é apresentada a pontuação média para referência.

Tabela 3 - Recolha dos totais de tags e pontuações atribuídas aos produtos da amostra da Amazon para comparativo

Produto	Nº total de tags atribuídas	Nº total de pontuações atribuídas	Nº total de tags de opinião atribuídas	Pontuação média
1	444	135	26	3.4
2	1213	704	28	4.4
3	680	404	140	4.4
4	1785	643	22	4.5
5	72	746	5	2.0
6	122	122	46	2.0
7	778	274	29	2.9
8	140	397	2	3.0
9	356	531	5	3.4
10	13	52	7	1.5
11	175	375	12	4.1
12	300	138	30	4.5
13	485	225	6	4.3
14	440	342	79	4.8
15	63	400	3	3.1
16	1538	141	1372	1.2
17	208	22	110	1.4
18	1477	946	65	3.8
19	354	263	43	3.1
20	158	87	7	2.0
21	803	596	121	4.3
22	86	88	3	2.2

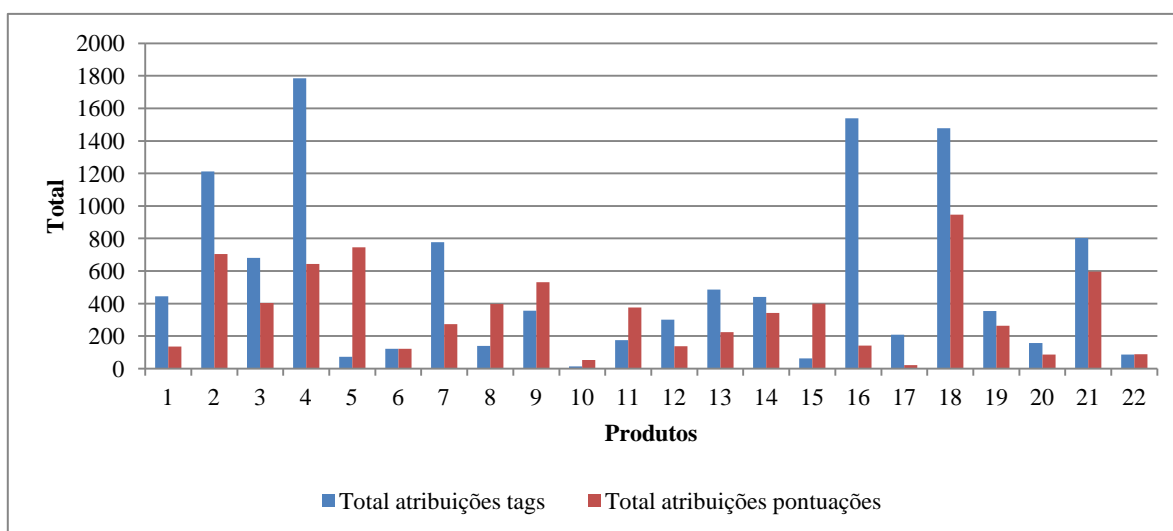


Figura 21 - Totais de tags e pontuações atribuídas aos produtos da amostra da Amazon para comparativo

5.3 Tags com uso da negação ou intensificação

Muitas das tags da amostra são constituídas por várias palavras e nestas situações pode-se analisar a polaridade contextual das mesmas. Em alguns casos é aplicado o uso da negação para provocar polaridade negativa e em outros são usados intensificadores para aumentar a intensidade da polaridade das tags.

Na Tabela 4 são apresentadas as tags da amostra com uso da negação na sua constituição. Palavras como “no” e “not” podem provocar polaridade global negativa considerados os outros termos da tag, o que faz com que as tags sejam globalmente classificadas como negativas.

Tabela 4 - Tags com uso da negação da amostra da Amazon

Tag	Classificação atribuída	Total de utilizações	Pontuação média
<i>no that funny</i>	negativa	1	4.4
<i>no good</i>	negativa	3	2.0
<i>no justice</i>	neutra	1	4.1
<i>not that great</i>	negativa	1	4.5
<i>no credibility</i>	negativa	1	1.2
<i>not worth it</i>	negativa	1	4.3
<i>not durable</i>	negativa	1	3.8

Observando os dados da tabela, identifica-se uma tag com classificação neutra atribuída, mas esta é uma exceção porque para o recurso em questão a tag está relacionada com o assunto do mesmo e isto significa que esta não representa uma emoção negativa neste contexto de utilização. Para os restantes recursos o uso da negação foi utilizado para expressar uma opinião negativa.

Note-se ainda que para as restantes tags, se não fosse utilizada a negação estas teriam polaridade positiva, por exemplo a palavra “good” representa nitidamente uma opinião positiva sobre o recurso. Portanto, o uso da negação contribuiu para a alteração da polaridade final das tags.

Na amostra considerada foram encontradas algumas palavras que utilizadas nas tags são usadas para intensificar a polaridade das mesmas. Estas têm este propósito quando utilizadas em conjunto com outras e quando a tag já teria a polaridade definida sem o uso do intensificador. A tag “highly recommended” tem polaridade positiva e se esta não utilizasse a palavra “highly” a polaridade continuaria a ser positiva, mas desta forma a intensidade da polaridade é superior. Também na amostra

considerada, foi encontrada a tag “*very funny*”, na qual o uso da palavra “funny” permite aumentar a intensidade da polaridade positiva.

5.4 Tags indefinidas ou imprecisas

Além de tags positivas, negativas e neutras foram classificadas na amostra algumas tags como indefinidas ou imprecisas (“*undefined*”). Foram poucas as tags classificadas desta forma, este grupo é constituído por apenas 9 tags na amostra considerada. Mas de qualquer forma foram assim classificadas para não influenciar a análise dos restantes grupos de tags.

Na Tabela 5 são apresentadas as tags indefinidas ou imprecisas encontradas na amostra considerada da Amazon e a pontuação média do recurso onde foram utilizadas. Cada uma delas só teve uma atribuição, mas algumas delas foram atribuídas aos mesmos recursos. Por coincidência, na tabela apresentada, as pontuações médias que coincidem são relativas ao mesmo recurso.

As tags indefinidas ou imprecisas são tags com vários sentidos (podem ser positivas ou negativas), significado desconhecido ou imperceptível.

Tabela 5 - Tags indefinidas ou imprecisas da amostra da Amazon

Tag	Pontuação média
<i>shocking</i>	4.5
<i>crapstacular</i>	1.2
<i>arm band for nana</i>	3.8
<i>equalizer</i>	3.8
<i>flat responce</i>	4.3
<i>foley work</i>	4.3
<i>grip</i>	4.3
<i>immersion</i>	4.3
<i>knock-off</i>	2.3

5.5 Utilização das tags

Considerando todas as tags da amostra da Amazon foram realizadas várias análises à utilização das tags. Estas análises baseiam-se na utilização das tags comparando com as pontuações atribuídas aos produtos. Estas análises permitem a comparação dos dois sistemas de classificação da Amazon: atribuição de tags e pontuações (estrelas de 1 a 5).

5.5.1 Análise comparativa entre tags e pontuação média

Neste ponto são apresentadas análises comparativas entre a utilização das tags e as pontuações atribuídas aos recursos da amostra.

Na Tabela 6 apresentam-se os dados relativos à recolha, de forma resumida. Para os produtos seleccionados são apresentados os totais de tags positivas, negativas e neutras, bem como as respectivas percentagens e pontuação média atribuída pelos utilizadores através do sistema de classificação que permite escolher entre 1 a 5 estrelas. Os últimos 5 produtos correspondem à recolha dos dispositivos electrónicos. A estes não são realizadas análises de forma particular porque verifica-se, através dos dados recolhidos, que o comportamento do uso de tags é semelhante nos diferentes produtos. Isto é, para produtos com pontuação média semelhante existem percentagens de tags semelhantes, não existindo grande discrepância entre os diferentes tipos de produtos.

Tabela 6 - Recolha de dados de produtos da Amazon

Produto	Total Tags Positivas	Total Tags Negativas	Total Tags Neutras	Percentagem Tags Positivas	Percentagem Tags Negativas	Percentagem Tags Neutras	Pontuação Média
1	131	9	540	19,26	1,32	79,41	4.4
2	0	7	6	0	53,85	46,15	1.5
3	22	6	1185	1,81	0,49	97,69	4.4
4	6	40	76	4,92	32,79	62,3	2.0
5	29	0	749	3,73	0	96,27	2.9
6	2	3	351	0,56	0,84	98,6	3.4
7	0	5	67	0	6,94	93,06	2.0
8	1	1	138	0,71	0,71	98,57	3.0
9	21	1	1762	1,18	0,06	98,77	4.5
10	0	3	60	0	4,76	95,24	3.1
11	9	3	163	5,14	1,71	93,14	4.1
12	0	110	98	0	52,88	47,12	1.4
13	79	0	361	17,95	0	82,05	4.8
14	29	1	270	9,67	0,33	90	4.5
15	1	25	418	0,23	5,63	94,14	3.4
16	11	1361	165	0,72	88,55	10,74	1.2

Produto	Total Tags Positivas	Total Tags Negativas	Total Tags Neutras	Percentagem Tags Positivas	Percentagem Tags Negativas	Percentagem Tags Neutras	Pontuação Média
17	6	0	479	1,24	0	98,76	4.3
18	112	9	678	14,02	1,13	84,86	4.3
19	53	12	1410	3,59	0,81	95,59	3.8
20	2	5	151	1,27	3,16	95,57	2.0
21	0	3	82	0	3,53	96,47	2.2
22	36	7	311	10,17	1,98	87,85	3.1

De seguida são apresentados gráficos de análise aos dados da amostra, estes consideram a variação da percentagem de tags positivas, negativas e neutras e a variação da percentagem das possíveis pontuações (1 a 5) relativamente à pontuação atribuída. Para melhor visualização dos gráficos estes são apresentados pelo agrupamento da pontuação atribuída aos produtos: pontuação negativa (menor que 3), intermédia (3) e positiva (maior que 3). Além de favorecer a visualização dos gráficos, esta divisão permite análises mais específicas por tipo de pontuação média. Fazer esta análise especificamente por produto não se mostrou necessário porque seria demasiado detalhe e seria difícil tirar conclusões da comparação dos dois sistemas de classificação como expressão da opinião dos utilizadores.

Como alguns autores verificaram, normalmente existem mais tags neutras atribuídas, as tags de opinião são menos atribuídas (Yanbe, Jatowt, Nakamura, & Tanaka, 2007). Mesmo assim através da análise dos gráficos podem-se tirar conclusões sobre a expressão da opinião dos utilizadores.

No gráfico apresentado na Figura 22 pode-se verificar a variação da percentagem de tags positivas, negativas e neutras relativamente à pontuação negativa atribuída. Neste gráfico as pontuações são baixas e consoante a pontuação diminui também as tags negativas são mais atribuídas existindo poucas tags positivas atribuídas a estes produtos.

É apresentada no gráfico da Figura 23 a variação da percentagem das possíveis pontuações (1 a 5) por pontuação negativa atribuída. A pontuação está ordenada da mesma forma que a do gráfico da Figura 22, para se fazer a correspondência e comparar os valores. Através do gráfico pode-se verificar que quando a pontuação média atribuída é menor, maior é a percentagem da estrela 1. A estrela 2 também é considerada uma pontuação negativa, mas como a maior parte das atribuições é de 1 estrela, as restantes foram atribuídas de uma forma menos intensa ou distribuída. Isto acontece porque a pontuação atribuída é a média das pontuações (1 a 5), isto verifica-se através das distribuições das atribuições das pontuações.

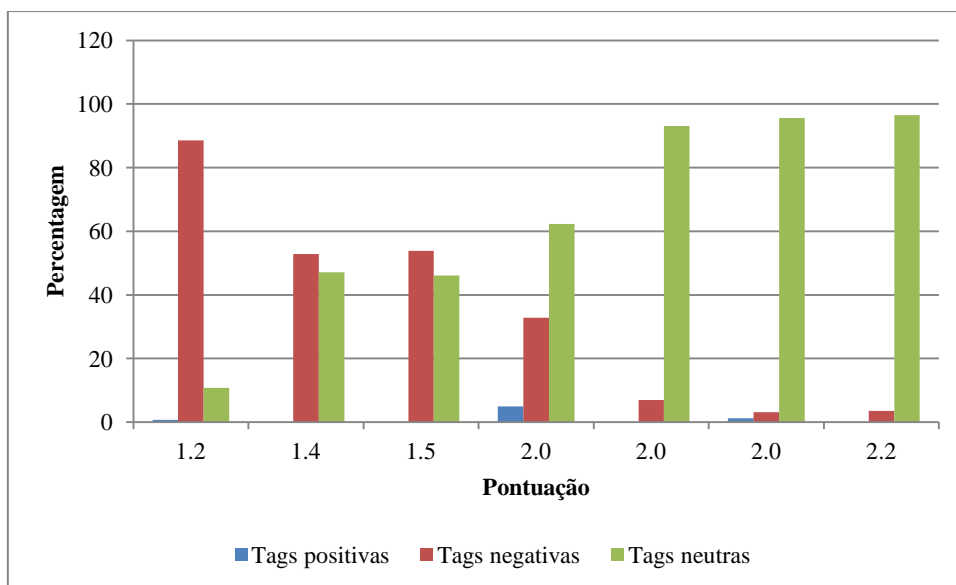


Figura 22 - Análise do número de tags positivas, negativas e neutras atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação negativa

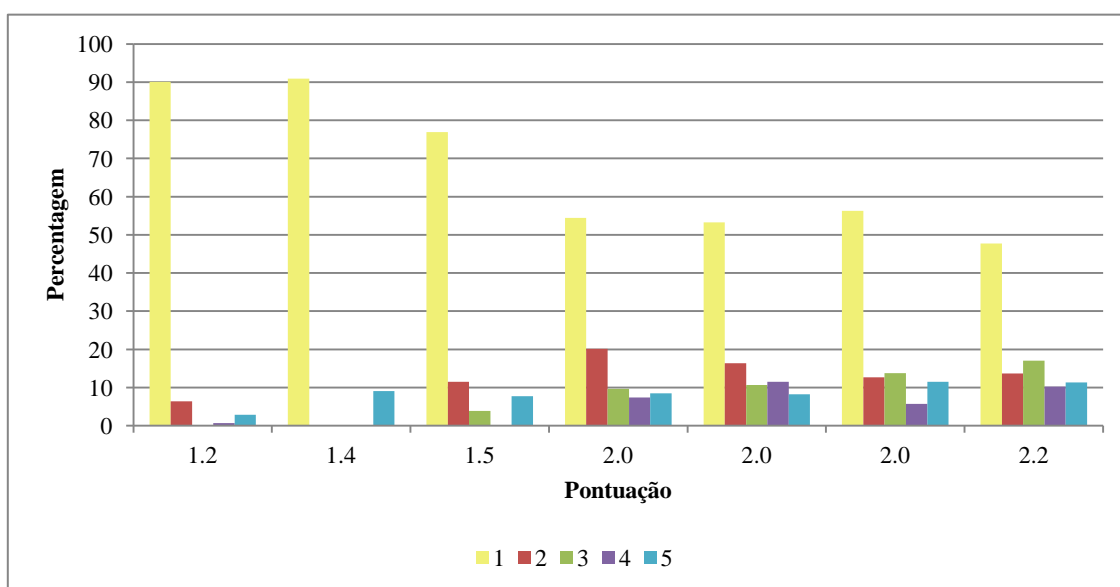


Figura 23 - Análise da percentagem da pontuação (1 a 5) por pontuação atribuída (correspondência de produto pela pontuação do gráfico da Figura 22)

Na Figura 24 é apresentada a variação da percentagem de tags positivas, negativas e neutras relativamente à pontuação intermédia atribuída. Estes são os casos em que a pontuação é mediana (por volta dos três) e verifica-se que não existe um grande número de tags de opinião, o que é compreensível visto que esses recursos não devem ter causado uma forte impressão nos utilizadores, predominando as tags neutras.

É apresentada no gráfico da Figura 25 a variação da percentagem das pontuações (1 a 5) por pontuação intermédia atribuída aos mesmos produtos referidos na Figura 24. Nestes casos em que a pontuação média é cerca de 3 estrelas, nota-se uma grande variação de percentagens nas pontuações possíveis de atribuir. Neste caso, ou existe maior atribuição de pontuações inferiores (1 a 2) ou superiores (4 a 5), ou então existe uma distribuição das percentagens por todos os tipos de estrelas. Assim sendo verifica-se que houveram diversas opiniões que fizeram com que a pontuação média atribuída a estes recursos ficasse intermédia.

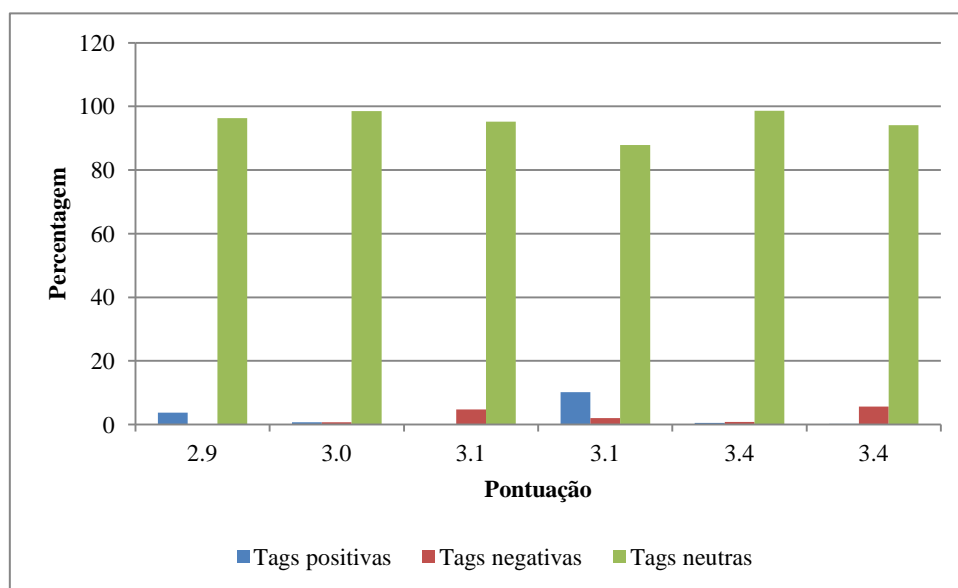


Figura 24 - Análise do número de tags positivas, negativas e neutras atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação intermédia

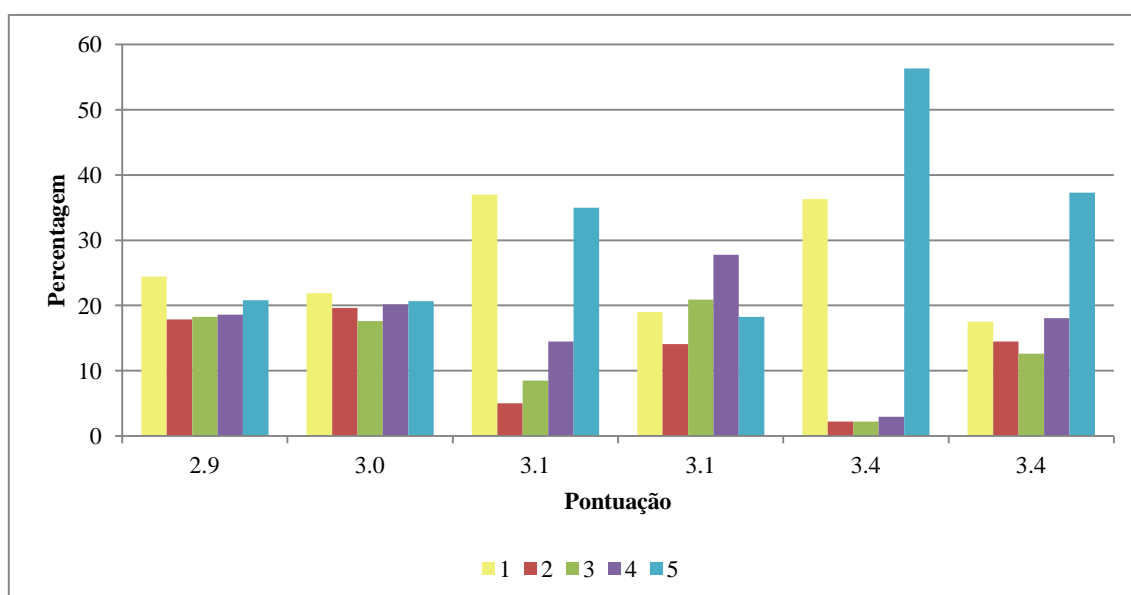


Figura 25 - Análise da percentagem da pontuação (1 a 5) por pontuação atribuída (correspondência de produto pela pontuação do gráfico da Figura 24)

Analisadas as pontuações médias negativas e intermédias atribuídas aos produtos da amostra falta somente analisar as positivas. Na Figura 26 é apresentada a variação da percentagem das tags por tipo relativamente à pontuação positiva atribuída. Para estes produtos que têm uma pontuação média positiva verifica-se que existem mais tags positivas e menos tags negativas atribuídas. Nota-se que as tags neutras continuam a ser as mais atribuídas.

Através do gráfico da Figura 27 pode-se analisar a variação da percentagem das pontuações (1 a 5) por pontuação positiva atribuída aos mesmos produtos indicados na Figura 26. Quando a pontuação média atribuída é maior ou considerada positiva, é maior a percentagem das estrelas 4 ou 5, enquanto as restantes percentagens são inferiores.

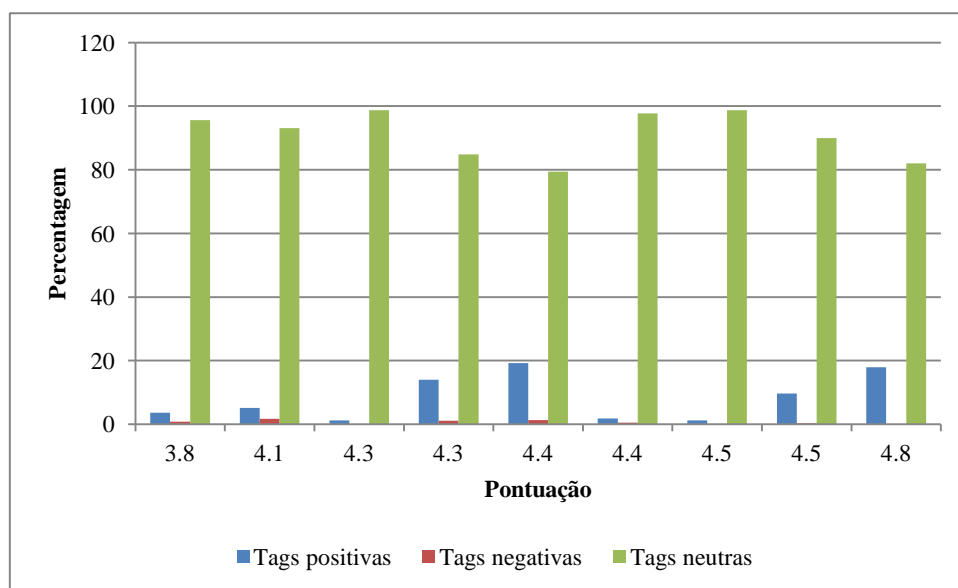


Figura 26 - Análise do número de tags positivas, negativas e neutras atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação positiva

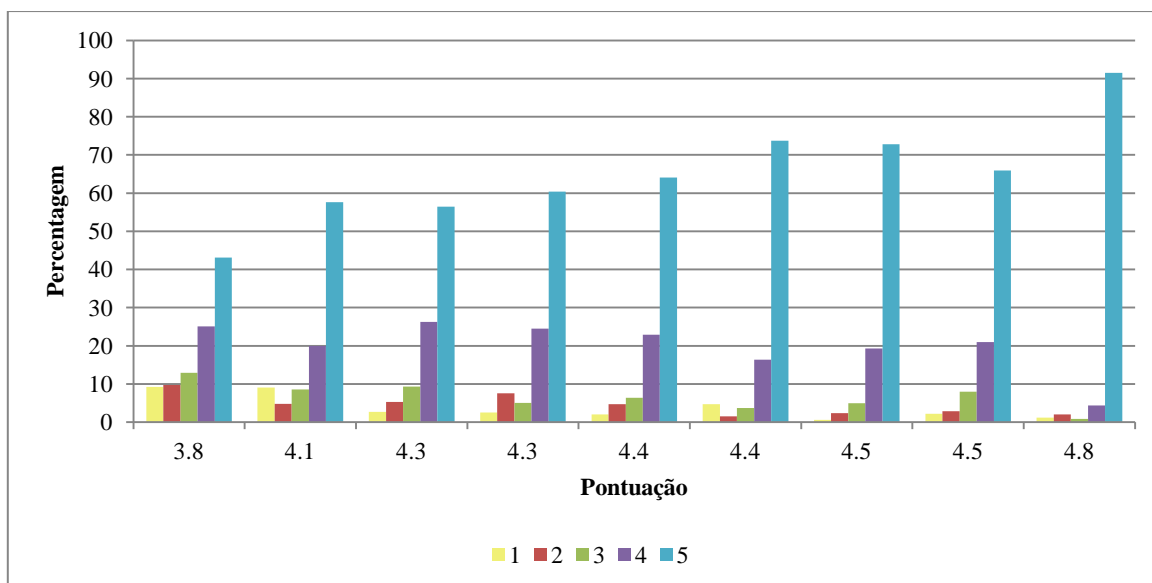


Figura 27 - Análise da percentagem da pontuação (1 a 5) por pontuação atribuída (correspondência de produto pela pontuação do gráfico da Figura 26)

Depois desta análise de comparação das diferentes opiniões, seja através das pontuações atribuídas ou da atribuição de tags, verifica-se que existe uma correspondência entre os dois sistemas de avaliação e classificação dos recursos. Isto é, verifica-se que através de sistemas de classificação distintos encontram-se, de forma geral, os mesmos padrões de opinião.

Os dados da recolha que deram origem aos gráficos anteriores sobre a análise da percentagem da pontuação (1 a 5) por pontuação atribuída são apresentados na Tabela 7. São apresentados os totais e percentagens de estrelas atribuídas por tipo de estrela (1 a 5) e a pontuação média atribuída pelos utilizadores.

Tabela 7 - Recolha de dados sobre as pontuações de produtos da Amazon

Produto	Total *	Total **	Total ***	Total ****	Total *****	% *	% **	% ***	% ****	% *****	Pontuação média
1	49	3	3	4	76	36	2,22	2,22	2,96	56,3	3.4
2	14	33	45	161	451	2	4,69	6,39	22,87	64,06	4.4
3	19	6	15	66	298	5	1,49	3,71	16,34	73,76	4.4
4	4	15	32	124	468	1	2,33	4,98	19,28	72,78	4.5
5	406	150	72	55	63	54	20,11	9,65	7,37	8,45	2.0
6	65	20	13	14	10	53	16,39	10,66	11,48	8,2	2.0
7	67	49	50	51	57	24	17,88	18,25	18,61	20,8	2.9
8	87	78	70	80	82	22	19,65	17,63	20,15	20,65	3.0
9	93	77	67	96	198	18	14,5	12,62	18,08	37,29	3.4
10	40	6	2	0	4	77	11,54	3,85	0	7,69	1.5
11	34	18	32	75	216	9	4,8	8,53	20	57,6	4.1
12	3	4	11	29	91	2	2,9	7,97	21,01	65,94	4.5
13	6	12	21	59	127	3	5,33	9,33	26,22	56,44	4.3
14	4	7	3	15	313	1	2,05	0,88	4,39	91,52	4.8
15	148	20	34	58	140	37	5	8,5	14,5	35	3.1
16	127	9	0	1	4	90	6,38	0	0,71	2,84	1.2
17	20	0	0	0	2	91	0	0	0	9,09	1.4
18	87	92	122	237	408	9	9,73	12,9	25,05	43,13	3.8
19	50	37	55	73	48	19	14,07	20,91	27,76	18,25	3.1
20	49	11	12	5	10	56	12,64	13,79	5,75	11,49	2.0
21	15	45	30	146	360	3	7,55	5,03	24,5	60,4	4.3
22	42	12	15	9	10	48	13,64	17,05	10,23	11,36	2.2

5.5.2 Análise da correlação entre tipos de tags e pontuação média

Neste ponto são realizadas análises da correlação entre a atribuição de tags e pontuações.

O coeficiente de correlação de Spearman, como já indicado, permite medir o grau de associação entre duas variáveis. Este foi utilizado para ser verificado o nível de correlação entre percentagem de tags (por tipo) e a pontuação média atribuída pelos utilizadores. Os resultados obtidos, e apresentados na Tabela 8, permitem afirmar que as percentagens de tags de opinião estão relacionadas com a pontuação média fornecida directamente pelos utilizadores. A correlação com mais significado é entre a percentagem de tags negativas e a pontuação, pois também apresenta uma probabilidade de erro nula ($p=,000$). Assim, em termos percentuais quanto mais tags negativas são atribuídas a um produto, menor é realmente a pontuação média do produto.

Tabela 8 - Correlação entre percentagem de tags (por tipo) e média ponderada das pontuações atribuídas pelos utilizadores

	% tags positivas		% tags negativas		% tags neutras		Pontuação média	
	Rs	p	Rs	p	Rs	p	Rs	p
% tags positivas	1		-,502	,017	-,179	,426	,594	,004
% tags negativas	-,502	,017	1		-,617	,002	-,774	,000
% tags neutras	-,179	,426	-,617	,002	1		,353	,107

Foi analisando o nível de correlação entre o número total de tags e a pontuação média atribuída pelos utilizadores, e verifica-se que para este caso existe uma correlação de $R_s=,477$ com probabilidade de erro de $p=,025$. A correlação encontrada não é forte, mas sustenta o facto de que quanto maior o número de tags atribuídas, maior é também a pontuação média.

5.6 Utilização das tags de opinião

Nesta secção são consideradas as tags classificadas como tags de opinião, isto é, as análises incidem nas tags positivas e negativas.

5.6.1 Análise comparativa entre tags de opinião e pontuação média

Contabilizando apenas as tags que efectivamente expressam uma opinião, apresentam-se na Tabela 9 os totais de tags de opinião, as percentagens de tags de opinião e respectiva pontuação média atribuída para cada recurso considerado na amostra.

Observando a tabela realçam percentagens muito altas e a maioria é consoante a pontuação média. Se a pontuação média é alta também a percentagem de tags positivas é alta e se a pontuação média é baixa a percentagem de tags negativas é alta. Algumas das percentagens tem valor zero, isto acontece à percentagem de tags positivas quando a pontuação média é baixa e sucede-se o mesmo à percentagem de tags negativas quando a pontuação média é alta. Portanto, por observação a percentagem de atribuição de tags de opinião tem comportamento diferente consoante a pontuação média dos recursos.

Tabela 9 - Recolha de dados de produtos da Amazon considerando só tags de opinião

Produto	Total tags positivas	Total tags negativas	Percentagem tags positivas	Percentagem tags negativas	Pontuação média
1	131	9	93,57	6,43	4.4
2	0	7	0	100	1.5
3	22	6	78,57	21,43	4.4
4	6	40	13,04	86,96	2.0
5	29	0	100	0	2.9
6	2	3	40	60	3.4
7	0	5	0	100	2.0
8	1	1	50	50	3.0
9	21	1	95,45	4,55	4.5
10	0	3	0	100	3.1
11	9	3	75	25	4.1
12	0	110	0	100	1.4
13	79	0	100	0	4.8
14	29	1	96,67	3,33	4.5
15	1	25	3,85	96,15	3.4
16	11	1361	0,8	99,2	1.2
17	6	0	100	0	4.3
18	112	9	92,56	7,44	4.3
19	53	12	81,54	18,46	3.8
20	2	5	28,57	71,43	2.0
21	0	3	0	100	2.2
22	36	7	83,72	16,28	3.1

De seguida são apresentados gráficos de análise aos dados da amostra, considerando a variação da percentagem tags positivas e negativas relativamente à pontuação média atribuída pelos utilizadores, isto por produto. Para melhor análise são apresentados gráficos por pontuação negativa (menor que 3), intermédia (3) e positiva (maior que 3) de forma a facilitar a avaliação.

Na Figura 28 é apresentado um gráfico com a variação da percentagem de tags positivas e negativas relativamente à pontuação média negativa atribuída pelos utilizadores. Notoriamente são

atribuídas maioritariamente tags negativas. Um dos recursos ainda tem uma percentagem considerável de tags positivas, mas foram atribuídas mais tags negativas.

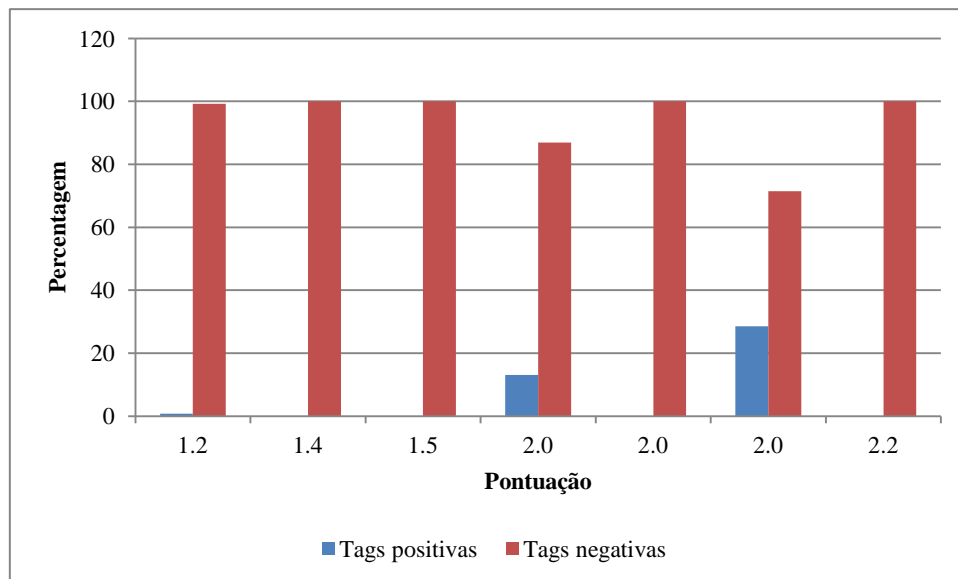


Figura 28 - Análise do número de tags de opinião atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação negativa

Quando a pontuação média é intermédia, isto é, é cerca de 3 estrelas, existem diferentes opiniões e consequentemente também existem diferentes utilizações das tags positivas e negativas como se pode verificar na Figura 29. Em alguns casos só foram atribuídas tags negativas ou positivas. Mas de notar que se forem verificadas as quantidades de tags através da Tabela 9 conclui-se que neste grupo são pouco utilizadas as tags de opinião, os totais destas são inferiores comparativamente à quantidade de tags de opinião para as pontuações maiores ou menores.

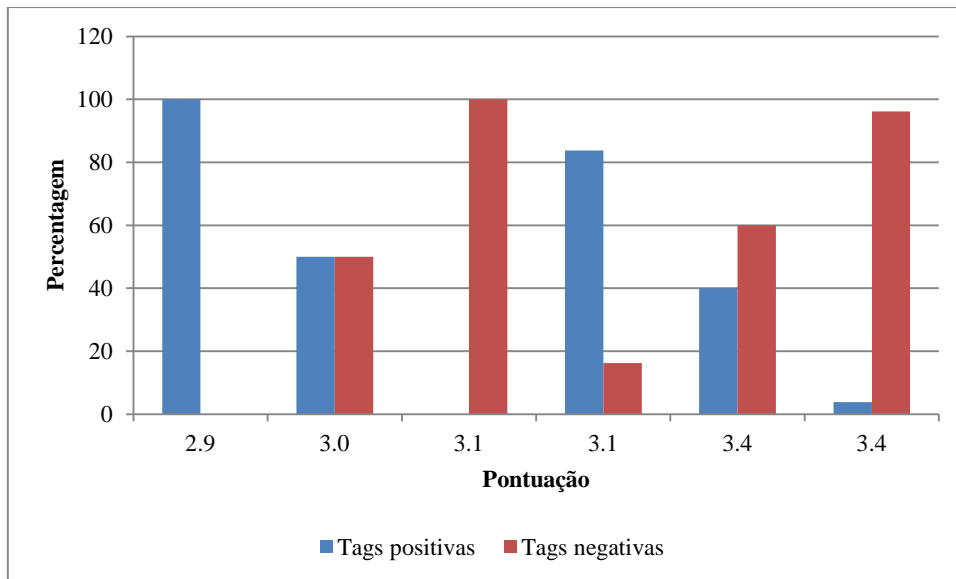


Figura 29 - Análise do número de tags de opinião atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação intermédia

Através do gráfico da Figura 30 é possível analisar a variação da percentagem de tags positivas e negativas relativamente à pontuação média positiva atribuída pelos utilizadores na amostra considerada. As tags de opinião atribuídas a produtos com pontuação positiva são maioritariamente tags positivas. Assim, verifica-se que a atribuição de tags negativas a alguns dos produtos pode estar relacionada com a pontuação atribuída. Um produto que cause algum desagrado não terá uma pontuação de 5 estrelas, e poderá ter tags negativas atribuídas.

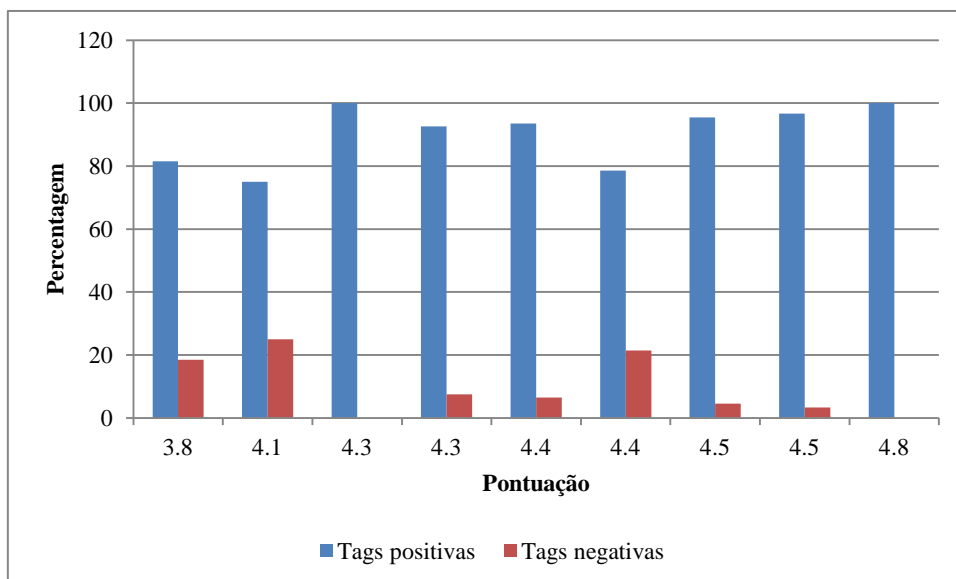


Figura 30 - Análise do número de tags de opinião atribuídas (valores percentuais) relativamente à pontuação de cada produto - pontuação positiva

Depois da análise comparativa entre a variação da percentagem das tags positivas e negativas relativamente à pontuação média atribuída pelos utilizadores concluí-se que quando a pontuação média é menor são mais usadas as tags negativas e quando a pontuação média é maior são mais usadas as tags positivas.

5.6.2 Análise da correlação entre tags de opinião e pontuação média

Nesta secção avalia-se a correlação entre as percentagens de tags positivas e negativas e as pontuações médias atribuídas, através da correlação de Spearman.

O coeficiente de correlação de Spearman foi calculado e os resultados são apresentados na Tabela 10. Considerando somente as tags de opinião, verifica-se que as percentagens de tags positivas e negativas estão relacionadas com a pontuação média fornecida directamente pelos utilizadores. Isto porque entre a utilização destes dois tipos de tags e a pontuação média, existe uma correlação de $R_s = ,756$ com probabilidade de erro de nula ($p = ,000$). Assim, pode-se afirmar que a pontuação está relacionada com a percentagem de tags de opinião atribuídas.

Tabela 10 - Correlação entre percentagem de tags de opinião e média das pontuações atribuídas pelos utilizadores

	% tags positivas		% tags negativas		Pontuação média	
	R_s	p	R_s	p	R_s	p
% tags positivas	1		-1		,756	,000
% tags negativas	-1		1		-,756	,000

5.6.3 Tags de opinião mais usadas

Considerando as tags de opinião, foram realizadas algumas análises específicas que utilizaram as tags de opinião mais usadas na amostra considerada, porque apresentam um peso considerável relativamente a toda a amostra. No fim desta secção são apresentadas as tags atribuídas a mais recursos.

Na Tabela 11 apresentam-se algumas das tags de opinião mais utilizadas na amostra considerada neste estudo. É descrita a respectiva frequência de utilização das tags nos produtos da Amazon, existe a indicação do número de recursos em que cada tag é usada na amostra considerada e ainda é apresentada a polaridade e frequência de utilização de cada tag. De referir que a polaridade das tags apresentadas, por acaso só têm uma polaridade pois a polaridade pode ser diferente consoante o recurso, por exemplo, “*garbage*” se aplicada a um livro sobre lixo seria classificada como neutra.

Mesmo consideradas todas as tags de opinião, verifica-se que para a amostra considerada, as tags positivas foram as menos atribuídas com uma utilização de 4,7% em relação ao total de tags. As tags negativas têm uma taxa de utilização de 13,8% em relação ao número total de tags utilizadas.

Tabela 11 - Tags de opinião mais usadas da amostra (Amazon)

Tag de Opinião	Frequência na amostra considerada	Polaridade atribuída na amostra	Número de recursos em que é usada	Dados da utilização das tags nos produtos
<i>ignorant</i>	250	negativa	1	969 utilizadores usaram a tag em 208 produtos
<i>biased</i>	195	negativa	2	742 utilizadores usaram a tag em 268 produtos
<i>lies</i>	166	negativa	1	4644 utilizadores usaram a tag em 1673 produtos
<i>fail</i>	158	negativa	1	1775 utilizadores usaram a tag em 799 produtos
<i>oversimplified</i>	154	negativa	1	168 utilizadores usaram a tag em 11 produtos
<i>lacks originality</i>	118	negativa	1	121 utilizadores usaram a tag em 3 produtos
<i>funny</i>	115	positiva	3	11290 utilizadores usaram a tag em 6328 produtos
<i>good price</i>	51	positiva	2	897 utilizadores usaram a tag em 557 produtos
<i>garbage</i>	50	negativa	4	4628 utilizadores usaram a tag em 3396 produtos
<i>book recommendations</i>	43	positiva	6	19180 utilizadores usaram a tag em 16463 produtos
<i>comfortable</i>	43	positiva	2	7142 utilizadores usaram a tag em 5656 produtos
<i>value</i>	35	positiva	3	2304 utilizadores usaram a tag em 1566 produtos
<i>funny mystery</i>	29	positiva	1	1488 utilizadores usaram a tag em 599 produtos
<i>poor writing</i>	24	negativa	2	128 utilizadores usaram a tag em 85 produtos
<i>trash</i>	22	negativa	3	3074 utilizadores usaram a tag em 2489 produtos
<i>terrible</i>	21	negativa	3	2801 utilizadores usaram a tag em 2131 produtos
<i>cliche</i>	18	negativa	2	186 utilizadores usaram a tag em 168 produtos
<i>hilarious</i>	16	positiva	1	4064 utilizadores usaram a tag em 1895 produtos
<i>laugh</i>	16	positiva	1	536 utilizadores usaram a tag em 237 produtos

Tag de Opinião	Frequência na amostra considerada	Polaridade atribuída na amostra	Número de recursos em que é usada	Dados da utilização das tags nos produtos
<i>horrible</i>	15	negativa	5	3275 utilizadores usaram a tag em 2219 produtos
<i>junk</i>	15	negativa	2	5780 utilizadores usaram a tag em 4607 produtos
<i>bigot</i>	14	negativa	1	931 utilizadores usaram a tag em 107 produtos
<i>delight</i>	14	positiva	1	65 utilizadores usaram a tag em 46 produtos
<i>hypocrisy</i>	14	negativa	1	873 utilizadores usaram a tag em 313 produtos
<i>useless crap</i>	14	negativa	1	270 utilizadores usaram a tag em 232 produtos

Para melhor análise da utilização das tags da tabela anterior, na Tabela 12 apresentam-se os totais e percentagens de utilização das tags consideradas em relação ao tipo de pontuação média. A pontuação média foi dividida em 3 tipos genéricos de pontuação média, estes podem ser considerados média negativa (<3), positiva (>3) e intermédia (=3). Assim é possível avaliar a utilização das tags consoante a pontuação média atribuída. Pode-se verificar por exemplo, que a tag “*funny*” foi utilizada somente em recursos com uma pontuação média considerada positiva.

Tabela 12 - Frequência de utilização de algumas tags considerado o tipo de pontuação média atribuída aos recursos

Tag	Pontuação média						Total
	<3		>3		=3		
	Total	%	Total	%	Total	%	
<i>ignorant</i>	250	100%	0	0%	0	0%	250
<i>biased</i>	194	99,5%	0	0%	1	0,5%	195
<i>lies</i>	166	89,2%	0	0%	20	10,8%	186
<i>fail</i>	158	100%	0	0%	0	0%	158
<i>oversimplified</i>	154	100%	0	0%	0	0%	154
<i>lacks originality</i>	118	100%	0	0%	0	0%	118
<i>funny</i>	0	0%	115	100%	0	0%	115
<i>good price</i>	0	0%	50	98%	1	2%	51
<i>garbage</i>	47	94%	2	4%	1	2%	50
<i>book recommendations</i>	0	0%	42	100%	0	0%	42
<i>comfortable</i>	0	0%	43	100%	0	0%	43
<i>value</i>	0	0%	3	8,6%	32	91,4%	35
<i>funny mystery</i>	0	0%	0	0%	29	100%	29
<i>poor writing</i>	24	100%	0	0%	0	0%	24
<i>trash</i>	20	90,9%	0	0%	2	9,1%	22
<i>terrible</i>	20	95,2%	1	4,8%	0	0%	21
<i>cliche</i>	18	100%	0	0%	0	0%	18
<i>hilarious</i>	0	0%	16	0%	0	0%	16
<i>laugh</i>	0	0%	16	100%	0	0%	16
<i>horrible</i>	14	93,3%	1	6,7%	0	0%	15
<i>junk</i>	14	93,3%	0	0%	1	6,7%	15
<i>bigot</i>	14	100%	0	0%	0	0%	14
<i>delight</i>	0	0%	14	100%	0	0%	14
<i>hypocrisy</i>	14	100%	0	0%	0	0%	14
<i>useless crap</i>	14	100%	0	0%	0	0%	14
Total	1239	76,1%	303	18,6%	87	5,3%	1629

Como forma de melhorar a visualização da frequência de uso das tags de opinião, são apresentadas na forma de gráfico as tags de opinião mais usadas da amostra na Figura 31, de forma similar à utilizada no estudo do serviço Japonês Hatena Bookmarks apresentado na secção 3.5 deste documento. Constituem o gráfico as tags mais usadas com frequência superior a 10 utilizações da amostra. Neste gráfico além de existir referência à frequência é também considerada a intensidade da polaridade atribuída às tags e divididas entre mais ou menos positiva ou negativa, a análise da intensidade foi obtida através do recurso SentiWordNet disponível *online* (ver ponto 3.3.2.1 para mais

informações sobre o recurso léxico). O levantamento da intensidade da polaridade foi realizado por palavra, mas quando a tag é constituída por mais do que uma palavra foi usada a intensidade da palavra com a mesma polaridade atribuída à tag (por exemplo, “*funny mystery*”) ou da palavra com intensidade superior quando ambas têm a mesma polaridade (por exemplo, “*useless crap*”). Quando existiam diversas intensidades de polaridade disponíveis para a mesma palavra foi utilizada a superior, isto porque não é possível determinar o significado e contexto de utilização da palavra.

Relativamente à amostra considerada, verifica-se que foram mais utilizadas as tags negativas. No estudo Japonês referido, os autores concluíram que as tags negativas são menos usadas que as positivas, mas na amostra da Amazon utilizada neste estudo isso não se verificou.

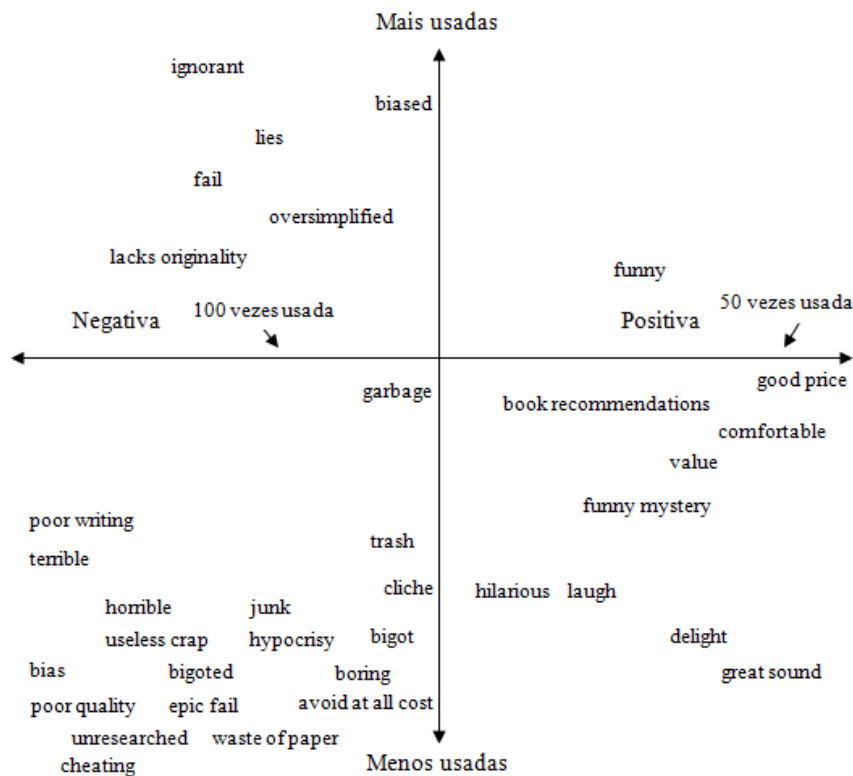


Figura 31 - Tags de opinião mais usadas da amostra (Amazon)

Através do mesmo recurso e para as mesmas tags mais usadas da amostra, foram analisadas as classes gramaticais das palavras que constituem as tags também através do SentiWordNet. Muitas das palavras que constituem as tags consideradas pertencem a várias classes gramaticais e não é possível escolher uma única porque as palavras podem ter diversos sentidos e significados que não é possível determinar através da associação das tags aos recursos. Assim, para os casos em que a classe da tag não pode ser determinada porque as mesmas palavras podem pertencer a classes gramaticais

diferentes, é atribuída uma classe indefinida a essa tag. Por exemplo, na língua inglesa a palavra *delight* é um substantivo mas também é um verbo, o que também se verifica para a palavra *junk*.

Para as tags mais usadas, apresentadas na Tabela 11, as classes usadas são: adjetivos, substantivos, verbos e indefinidos. Assim sendo apresenta-se no gráfico da Figura 32 as divisões das classes gramaticais e verifica-se que os adjetivos são os mais usados como tags de opinião, excluindo os indefinidos. Através das categorizações de tags apresentadas por alguns autores eram esperados bastantes adjetivos utilizados, mas também verifica-se que palavras pertencentes a outras classes gramaticais também são utilizadas pois também permitem expressar a opinião sobre os recursos.

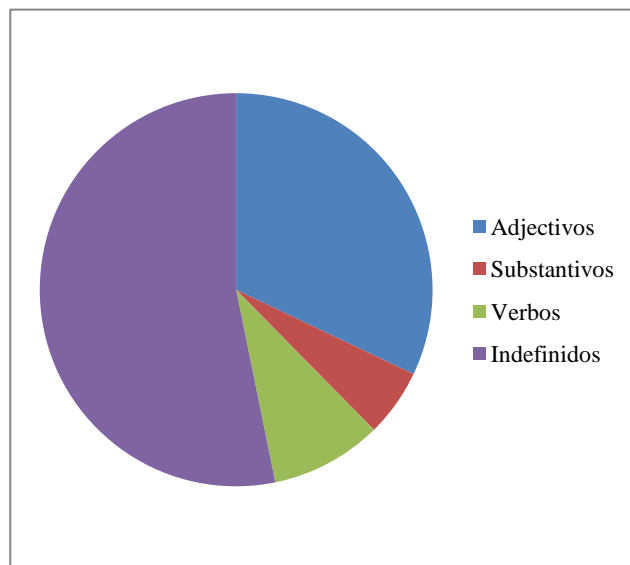


Figura 32 - Classes gramaticais das tags de opinião mais usadas da amostra (Amazon)

Outra forma de verificar as tags mais usadas será analisar as tags de opinião atribuídas a mais recursos da amostra considerada. Estas tags são apresentadas na Tabela 13 com os respectivos dados da utilização das tags nos produtos da Amazon, indicação da frequência de utilização das tags nos recursos contabilizados, polaridade atribuída e ainda é apresentado o número de recursos em que a tag é utilizada. Foram incluídas nesta tabela as tags usadas em mais de dois recursos. Por observação desta tabela e da Tabela 11 onde são apresentadas as tags mais usadas, podem-se fazer as seguintes comparações: algumas tags são comuns às duas tabelas, na Tabela 13 estão a negrito as tags que coincidentes; e continuam a ser as tags com polaridade negativa as mais atribuídas.

Tabela 13 - Tags de opinião atribuídas a mais recursos da amostra da Amazon

Tag de Opinião	Número de recursos em que é usada	Polaridade atribuída na amostra	Frequência na amostra nos recursos	Dados da utilização das tags nos produtos
<i>book recommendations</i>	6	positiva	43	19180 utilizadores usaram a tag em 16463 produtos
<i>horrible</i>	5	negativa	15	3275 utilizadores usaram a tag em 2219 produtos
<i>boring</i>	5	negativa	12	3399 utilizadores usaram a tag em 3422 produtos
<i>bad</i>	4	negativa	4	3859 utilizadores usaram a tag em 3133 produtos
<i>garbage</i>	4	negativa	50	4628 utilizadores usaram a tag em 3396 produtos
<i>funny</i>	3	positiva	115	11290 utilizadores usaram a tag em 6328 produtos
<i>complete crap</i>	3	negativa	3	1310 utilizadores usaram a tag em 656 produtos
<i>bad book</i>	3	negativa	4	351 utilizadores usaram a tag em 246 produtos
<i>terrible</i>	3	negativa	21	2801 utilizadores usaram a tag em 2131 produtos
<i>trash</i>	3	negativa	22	3074 utilizadores usaram a tag em 2489 produtos
<i>value</i>	3	positiva	35	2304 utilizadores usaram a tag em 1566 produtos
<i>worthless</i>	3	negativa	9	1763 utilizadores usaram a tag em 1495 produtos

5.6.4 Análise da dependência entre tags de opinião e pontuação média

Através de métodos de análise de dados pode-se estudar a dependência entre as tags de opinião e a pontuação média atribuída aos recursos. Através da observação da utilização de algumas tags de opinião pretende-se avaliar se existe associação entre a utilização ou a não utilização dessas tags e a pontuação média atribuída dos recursos. É possível realizar testes que permitem observar em que casos as tags de opinião são ou não utilizadas.

Para estas análises são usadas tabelas de contingência. As tabelas de contingência permitem representar dados, principalmente quando estes podem ser classificados segundo dois critérios. Estas permitem verificar o relacionamento entre duas variáveis.

Nestas análises são usadas duas variáveis: utilização da tag de opinião e pontuação média. Cada uma das variáveis é analisada consoante as classes de interesse. Para a variável tag de opinião é contabilizado o número de recursos que “tem tag atribuída” e os que “não tem tag atribuída”. Comparativamente com a variável pontuação média em que é contabilizado o número de recursos com

pontuação média negativa “<3” e os que têm a pontuação média considerada intermédia e positiva “>=3”. A última classe é um conjunto constituído simultaneamente pela pontuação média intermédia e positiva para diminuir o erro na análise.

Várias análises podem ser realizadas usando as tabelas de contingência, mas para a análise pretendida e com os dados disponíveis, o teste considerado mais adequado é o teste exacto de Fisher. Foi este o teste aplicado, pelas seguintes razões:

- O teste mais conhecido é o Qui-quadrado, mas em amostras pequenas o erro é alto;
- O teste de Fisher permite calcular a probabilidade das características que estão em análise serem independentes, quando a mostra é pequena;
- Usa-se o teste de Fisher quando o número total da amostra é menor que 20 ou este é maior que 20 e menor que 40 e a menor frequência é menor que 5.

O teste de Fisher calcula a probabilidade com base no produto dos factoriais dos totais marginais da tabela pelo factorial do total geral multiplicado pelo inverso do produto dos factoriais dos valores observados em cada classe. Quando este valor é menor que 5% (nível de significância habitualmente considerado) diz-se que as variáveis são dependentes, isto é, a sua associação não é casual. Na análise de dependência entre variáveis este teste vai ser usado para testar a hipótese de as variáveis em questão não serem independentes. Para cada teste há duas hipóteses:

H_0 (hipótese nula): As variáveis são independentes.

H_1 (hipótese nula): As variáveis não são independentes

Como já referido, as tabelas de contingência permite a representação dos dados, estas apresentam os resultados das variáveis em estudo: mostram o número de observações considerando diferentes critérios e combinações. São apresentados quatro casos, isto é, quatro tags de opinião com utilização em recursos de variadas pontuações médias.

Na Tabela 14 apresentam-se os totais e percentagens de recursos que têm ou não a tag “*horrible*” atribuída em relação à pontuação média dos recursos. Como já referido, a pontuação média foi dividida em 2 tipos de pontuação média, esta pode ser considerada média negativa (<3) e média intermédia ou positiva (>=3). Com estes dados foi aplicado o teste exacto de Fisher que obteve o resultado de 4,31%. Com este resultado pode-se afirmar que existe uma relação estatisticamente significativa entre a utilização da tag “*horrible*” e a pontuação média dos recursos, isto é, a hipótese das variáveis não serem independentes é aceite. Esta análise permite afirmar a tag “*horrible*” é utilizada em recursos com a pontuação média inferior a 3 estrelas e a pontuação média é superior quando os recursos não têm a tag “*horrible*” atribuída. A tag foi classificada como negativa, por isso é

um dos casos que efectivamente a tag está a ser usada nos recursos com pontuação média considerada negativa.

Tabela 14 - Tabela de contingência para avaliação de utilização da tag *horrible* e pontuação média (amostra Amazon)

Utilização da tag “ <i>horrible</i> ” (negativa) aos recursos	Pontuação média				Total
	<3		≥3		
	Total	%	Total	%	
Tem tag atribuída	4	80%	1	20%	5
Não tem tag atribuída	4	23,5%	13	76,5%	17
Total	8	36,4%	14	63,6%	22

Por outro lado na Tabela 15, apresentam-se os totais e percentagens de recursos que têm ou não a tag “*boring*” atribuída em relação ao tipo de pontuação média atribuída a esses recursos. Para este caso e com estes dados foi aplicado o teste exacto de Fisher e este obteve o resultado de 23,3%. Portanto, a hipótese nula é rejeitada ao nível de significância de 5%. Por observação verifica-se que esta tag negativa é usada frequentemente em recursos com pontuação média negativa. E os recursos que não têm esta tag atribuída normalmente são recursos com pontuação média positiva.

Tabela 15 - Tabela de contingência para avaliação de utilização da tag *boring* e pontuação média (amostra Amazon)

Utilização da tag “ <i>boring</i> ” (negativa) aos recursos	Pontuação média				Total
	<3		≥3		
	Total	%	Total	%	
Tem tag atribuída	3	60%	2	40%	5
Não tem tag atribuída	5	29,4%	12	70,6%	17
Total	8	36,4%	14	63,6%	22

Foram também analisadas a atribuição de tags com polaridade positiva. Na Tabela 16 apresentam-se os totais e percentagens de recursos que têm ou não a tag “*book recommendations*” atribuída em relação ao tipo de pontuação média atribuída aos recursos. Foi aplicado o teste exacto de Fisher e este obteve o resultado de 4%. Com este resultado a hipótese nula é aceite ao nível de significância de 5%. Pode-se assim afirmar que a associação entre a utilização da tag nos recursos e a pontuação média dos mesmos não é casual. Portanto, a tag “*book recommendations*” é atribuída a recursos com a pontuação média igual ou superior a 3 estrelas e a pontuação média dos recursos é

inferior quando os recursos não têm esta tag atribuída. A tag tem polaridade positiva, logo é um dos casos em que efectivamente a tag está a ser utilizada nos recursos com pontuação média considerada positiva.

Tabela 16 - Tabela de contingência para avaliação de utilização da tag *book recommendations* e pontuação média (amostra Amazon)

Utilização da tag “ <i>book recommendations</i> ” (positiva) aos recursos	Pontuação média				Total
	<3		≥3		
	Total	%	Total	%	
Tem tag atribuída	0	0%	6	100%	6
Não tem tag atribuída	8	50%	8	50%	16
Total	8	36,4%	14	63,6%	22

Na Tabela 17, apresentam-se os totais e percentagens de recursos que têm ou não a tag “*funny*” atribuída em relação ao tipo de pontuação média atribuída aos recursos. Neste caso o teste exacto de Fisher obteve o resultado de 23,6%. Portanto, a hipótese das variáveis nula é rejeitada ao nível de significância de 5%. Também neste caso, por observação, verifica-se que esta tag positiva é usada frequentemente em recursos com pontuação média positiva. E os recursos que não têm esta tag atribuída normalmente são recursos com pontuação média negativa.

Tabela 17 - Tabela de contingência para avaliação de utilização da tag *funny* e pontuação média (amostra Amazon)

Utilização da tag “ <i>funny</i> ” (positiva) aos recursos	Pontuação média				Total
	<3		≥3		
	Total	%	Total	%	
Tem tag atribuída	0	0%	3	100%	6
Não tem tag atribuída	8	50%	11	50%	16
Total	8	36,4%	14	63,6%	22

Estas análises permitiram analisar a dependência de algumas tags em relação à pontuação média dos recursos.

6 Conclusões

Este capítulo apresenta os objectivos atingidos com o trabalho desenvolvido e o respectivo impacto nos serviços de informação colaborativa. Identificam-se as dificuldades e limitações deste trabalho. São ainda apresentadas algumas propostas de desenvolvimentos e considerações futuras que podem complementar este trabalho.

6.1 Considerações finais

Através dos dados recolhidos da loja *online* Amazon e das análises realizadas concluí-se que as tags de opinião atribuídas aos recursos podem verdadeiramente estar relacionadas com a expressão da opinião relativamente aos recursos. Isto porque, considerada a pontuação atribuída mais próxima da opinião geral dos utilizadores, verificou-se que existe correspondência entre as tags de opinião atribuídas e respectiva pontuação, com significado estatístico na amostra considerada no estudo de caso apresentado. Portanto, foram atingidos os objectivos do trabalho definidos no Capítulo 1 desta dissertação.

As principais contribuições do presente trabalho são as seguintes:

- Apresentação de um estudo detalhado sobre *tagging* e serviços de *bookmarking* social (Capítulo 2);
- Apresentação de uma revisão da literatura sobre tags de opinião (Capítulo 3);
- Apresentação de um estudo inovador sobre a utilização das tags de opinião numa loja *online*, segundo diversos aspectos (Capítulos 4 e 5);
- Proposta de um classificador semi-automático de tags em relação à polaridade (Capítulo 4).

As análises realizadas permitiram retirar algumas conclusões sobre a amostra considerada neste trabalho (apresentadas no Capítulo 5):

- A maioria das tags atribuídas são constituídas por mais do que uma palavra;
- Maior frequência na utilização de tags do que na recomendação através da atribuição das pontuações;
- Quanto maior o número de tags atribuídas maior é a pontuação média dos recursos;
- As tags mais usadas têm polaridade negativa;
- Existe dependência entre as tags de opinião e a pontuação média dos recursos, isto é o mesmo que dizer que existe correlação ou relação.

O trabalho apresentado poderá ser bastante útil para serviços de informação colaborativa, na medida em que pode ser utilizado para indicar se os recursos são globalmente considerados de forma positiva ou negativa. Mesmo o número de tags de um recurso poderá dar uma primeira indicação da polaridade da opinião dos utilizadores.

A inteligência é cada vez mais o futuro, principalmente para o futuro da Web, por isso tornam-se importantes os estudos da informação que permita obter inteligência. Relativamente ao uso de tags e atribuição de pontuações é possível analisar o comportamento dos utilizadores e indicar se estes estão a influenciar da forma correcta outros utilizadores. No caso de uma loja *online*, a opinião que influencia outros utilizadores é importante, pois as vendas são influenciadas e também a própria veracidade da informação da loja.

Vários trabalhos de investigação têm se dedicado ao estudo de folksonomias e às diversas aplicações que lhes estão associadas. No entanto, este trabalho debruça-se sobre um tipo de tag em particular, as tags sentimentais ou de opinião, como são utilizadas e se efectivamente podem ser utilizadas para inferir a opinião que os utilizadores têm sobre os recursos. Não foram encontrados estudos com esta particularidade. Os estudos encontrados sobre a expressão da opinião *online* referem-se apenas a comentários e classificações atribuídas pelos utilizadores, sem nenhuma tentativa de verificar se as mesmas podem ser utilizadas na identificação de uma opinião globalmente positiva ou negativa sobre os recursos.

Foi construído um classificador semi-automático de tags, este foi desenvolvido de forma progressiva durante este trabalho para que fossem identificados e analisados os problemas de classificação de tags.

Durante a realização deste trabalho foi submetido e aceite um artigo intitulado “Utilização de Tags de Opinião – um estudo de caso” na 6ª conferência Ibérica de Sistemas e Tecnologias de Informação. No artigo descreve-se este trabalho, mas está mais focado no estudo de caso sobre a loja *online* Amazon.

Este trabalho permitiu alargar conhecimentos na área de especialização em tecnologias do conhecimento e decisão do mestrado em engenharia informática, Por outro lado, esses conhecimentos foram utilizados em áreas novas, como análise sentimental e sistemas de informação colaborativos. Assim, foi um trabalho gratificante também a nível pessoal.

6.2 Limitações e dificuldades encontradas

Durante a realização deste trabalho foram encontradas algumas limitações e dificuldades que são de seguida apresentadas.

Houve alguma dificuldade em trabalhar com conceitos tão subjectivos, com a necessidade de considerar diversos aspectos e pesquisas complementares de forma a classificar correctamente as tags atribuídas aos recursos. Todo o processo é moroso e complexo. Essa dificuldade reflectiu-se na classificação das tags, nomeadamente na identificação do sentido da atribuição de tags pelos utilizadores. Nesse processo participaram duas pessoas e para algumas tags foram assumidos diferentes significados às mesmas, e só a consideração ampla de diversos factores permitiu que fosse percebido o possível sentido de algumas tags. Muitas vezes foram utilizados dicionários *online* para se perceber as diversas hipóteses de utilização dos termos, inclusive para a identificação de expressões idiomáticas, cujo significado pode não ser percebido pela mera análise das suas palavras constituintes. Por exemplo, a tag “*dystopian*” foi classificada de forma diferente pelas duas pessoas, isto aconteceu devido ao desconhecimento de este ser um género literário, depois da troca desta informação ambas a classificaram como neutra por esta não expressar uma opinião negativa ou positiva sobre o recurso.

Foi construído um classificador semi-automático que permitiu atenuar algumas das dificuldades na classificação das tags. Pois, este considera algumas características do recurso na sua classificação, nomeadamente o título do mesmo. O classificador também aplica as considerações relevantes da classificação da polaridade, como é o caso da polaridade contextual da frase em que existe a consideração do uso da negação quando as tags são constituídas por várias palavras.

Como principal limitação deste trabalho identifica-se a amostra ser pequena. O tempo disponível para a realização deste trabalho (cerca de um ano lectivo) e a complexidade inerente à classificação das tags e à construção de um sistema semi-automático para esse efeito foram os principais factores que conduziram a esta limitação do trabalho.

6.3 Trabalho futuro

As limitações mencionadas na secção anterior, constituem também oportunidades de investigação que poderão ser contempladas em trabalhos futuros.

Com os dados apresentados neste estudo verifica-se que será útil a integração de análises sobre a classificação de recursos de uma forma geral na loja *online* Amazon. Para isso será necessário analisar um extractor de dados Web, de forma a este extrair os dados necessários do sítio Web da loja *online* Amazon.

Pretende-se considerar de forma diferente tags como “*good*” e “*very good*”, por exemplo. Apesar de ambas serem consideradas tags de opinião positivas, pretende-se considerar intensidades diferentes de polaridade. Um recurso que já se identificou adequado para este propósito é o SentiWordNet 3.0, um recurso léxico que pode ser incorporado em aplicações relacionadas com classificação de sentimento e estimativa de opinião (Baccianella, Esuli, & Sebastiani, 2010).

A classificação de tags nos serviços *online* podem-se mostrar muito úteis, pois no futuro estes poderiam apresentar *rankings* através das tags atribuídas como forma de apresentação dos resultados das tags de opinião atribuídas pelos utilizadores. Por exemplo, para *ranking* dos livros mais recomendados, esta avaliação teria em conta a atribuição de tags e pontuações mais positivas.

O trabalho apresentado aplica-se a um sistema de língua estrangeira, mas um recurso léxico emocional para a língua portuguesa europeia (Rodrigues, 2009) poderá ser útil para a consideração de tags, tais como “gosto” ou “bom”. O mesmo poderá constituir uma base para a construção de um classificador de sentimentos para esta língua. Quando existirem sistemas em português com as características e dimensão da Amazon, poderão ser realizadas análises semelhantes às do presente trabalho e verificar as diferenças das utilizações de tags consoante a língua.

Bibliografia

- Alag, S. (2009).** *Collective Intelligence in Action*. Greenwich: Manning.
- Baccianella, S., Esuli, A., & Sebastiani, F. (2010).** SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proceedings of the Seventh Conference on International Language Resources and Evaluation* (pp. 2200-2204). Valletta, Malta: European Language Resources Association.
- Bal, B. K. (2009).** Towards an analysis of opinions in news editorials: how positive was the year? *Proceedings of the 8th International Conference in Computational Semantics*, (pp. 260-263). Tilburg.
- Boiy, E., Hens, P., Deschacht, K., & Moens, M.-F. (2007).** Automatic sentiment analysis in on-line text. *Proceedings of the 11th International Conference on Electronic Publishing*, (pp. 349-360). Vienna, Austria.
- Dennis, B. (2006).** Foragr: collaboratively tagged photographs and social information visualization. *Collaborative Web Tagging Workshop*. Edinburgh.
- Esuli, A., & Sebastiani, F. (2006).** SentiWordNet: A publicly available lexical resource for opinion mining. *Proceedings of the 5th Conference on Language Resources and Evaluation*, (pp. 417-422).
- Golder, S. A., & Huberman, B. A. (2006).** Usage patterns of collaborative tagging systems. *Journal of Information Science* , 32 (2), 198-208.
- Hammond, T., Hannay, T., Lund, B., & Scott, J. (2005).** *Social Bookmarking Tools (I)*. *D-Lib Magazine*, 11(4). Obtido em 2011, de <http://www.dlib.org/dlib/april05/hammond/04hammond.html>
- Heckner, M., Mühlbacher, S., & Wolff, C. (2008).** Tagging tagging. Analysing user keywords in scientific bibliography management systems. *Journal of Digital Information* , 9 (2), 1-19.
- Heckner, M., Neubauer, T., & Wolff, C. (2008).** Tree, funny, to_read, google: what are tags supposed to achieve? a comparative analysis of user keywords for different digital resource types. *Proceeding of the 2008 ACM workshop on Search in social media*, (pp. 3-10). New York.
- Kipp, M. E. (2007).** @toread and Cool : tagging for time, task and emotion. *Information Architecture Summit Las Vegas*, (pp. 22-26). Las Vegas, Nevada.
- Kipp, M., & Campbell, D. (2006).** Patterns and inconsistencies in collaborative tagging systems: an examination of tagging practices. *The 17th Annual General Meeting of the American Society for Information Science and Technology*. Austin, Texas.
- Kroski, E. (2005).** *The Hive Mind: Folksonomies and User-Based Tagging*. Obtido em 2011, de <http://infotangle.blogspot.com/2005/12/07/the-hive-mind-folksonomies-anduser-based-tagging/>

- Kumar, R. (2005).** *Research methology A step-by-step guide for begginers.* Sage.
- Li, X., & Hitt, L. M. (2008).** Self-selection and information role of online product reviews. *Information Systems Research*, (pp. 456–474).
- Lin, W.-H., Wilson, T., Wiebe, J., & Hauptmann, A. (2006).** Which side are you on? identifying perspectives at the document and sentence levels. *Proceedings of the 10th Conference on Computational Natural Language Learning.* New York City: 109–116.
- Marlow, C., Naaman, M., boyd, d., & Davis, M. (2006).** Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead. *Proceedings of the Collaborative Web Tagging Workshop.* Edinburgh, Scotland.
- Newhagen, J. E. (1998).** TV news images that induce anger, fear, and disgust: Effects on approach-avoidance and memory. *Journal of Broadcasting & Electronic Media* , 42 (2), 265-276.
- NISO. (2004).** Obtido em 2011, de <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>
- Pang, B., & Lee, L. (2008).** Opinion Mining and Sentiment Analysis. In *Foundations and Trends in Information Retrieval* 2(1-2) (pp. 1-135).
- Pang, B., Lee, L., & Vaithyanathan, S. (2002).** Thumbs up? sentiment classification using machine learning techniques. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, (pp. 79-86). Stroudsburg.
- Peters, I. (2009).** *Folksonomies indexing and retrieval in Web 2.0.* K G Saur Verlag.
- Peterson, E. (11 de 2006).** *Beneath the Metadata: some philosophical problems with Folksonomy.* *D-Lib Magazine*, v.12, n.11. Obtido em 05 de 2011, de <http://www.dlib.org/dlib/november06/peterson/11peterson.html>
- Ponte, J. P. (2006).** Estudos de caso em educação matemática. *Bolema*, 25 (Este artigo é uma versão revista e actualizada de um artigo anterior: Ponte, J. P. (1994). O estudo de caso na investigação em educação matemática. *Quadrante*, 3(1), pp3-18. (re-publicado com autorização)), pp. 105-132.
- Rodrigues, D. H. (2009).** *Construção automática de um dicionário emocional para o português.* Covilhã: Universidade da Beira Interior.
- Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., et al. (2006).** tagging, communities, vocabulary, evolution. *Computer Supported Cooperative Work* (pp. 181-190). Banff, Alberta, Canada: ACM.
- Smith, G. (2008).** *Tagging: People-Powered Metadata for the Social Web.* New Riders.

Vander Wal, T. (2005). *Explaining and Showing Broad and Narrow Folksonomies*. Obtido em 2011, de <http://www.vanderwal.net/random/entrysel.php?blog=1635>

Vander Wal, T. (2007). *Folksonomy: Folksonomy Coinage and Definition*. Obtido em 2011, de <http://www.vanderwal.net/folksonomy.html>

Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language*, (pp. 347–354). Vancouver.

Wu, F., & Huberman, B. A. (2010). Opinion Formation Under Costly Expression. *ACM Transactions on Intelligent Systems and Technology (TIST)* .

Yanbe, Y., Jatowt, A., Nakamura, S., & Tanaka, K. (2007). Can social bookmarking enhance search in the web? *Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries*, (pp. 107-116). Vancouver.

Yu, H., & Hatzivassiloglou, V. (2003). Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences. *Proceedings of the 2003 conference on empirical methods in natural language processing*. Stroudsburg.

Zollers, A. (2007). Emerging motivations for tagging: expression, performance, and activism. *Tagging and Metadata for Social Information Organization Workshop*. Banff, Canada.

Anexo 1 Dados recolhidos da Amazon

Os dados mais relevantes que constituem a amostra do estudo de caso da Amazon são apresentados neste anexo. Não são apresentados todos os dados devido à sua extensão.

1.1 Produtos

Os dados relativos aos produtos que constituem a amostra da Amazon são apresentados na Tabela 18. A identificação relativa aos livros corresponde ao ISBN (International Standard Book Number) dos mesmos, enquanto para os auscultadores corresponde ao ASIN (Amazon Standard Identification Number).

Tabela 18 - Dados dos produtos recolhidos da Amazon

<i>ID</i>	<i>Title</i>	<i>Stars</i>	<i>Reviews</i>	<i>Department</i>
9780061992704	<i>Sh*t My Dad Says [Hardcover]</i>	4.4	404	<i>Books</i>
9780071468725	<i>McGraw-Hills Conquering LSAT Logic Games [Paperback]</i>	1.5	52	<i>Books</i>
9780307269997	<i>The Girl Who Kicked the Hornets Nest Deckle Edge [Hardcover]</i>	4.4	704	<i>Books</i>
9780316003360	<i>Cleaving: A Story of Marriage, Meat, and Obsession [Hardcover]</i>	2.0	122	<i>Books</i>
9780375431883	<i>Visions of Sugar Plums (Random House Large Print (Cloth/Paper))[Large Print][Hardcover]</i>	2.9	274	<i>Books</i>
9780385528047	<i>The Confession: A Novel [Hardcover]</i>	3.4	531	<i>Books</i>
9780399150890	<i>Blow Fly [Hardcover]</i>	2.0	746	<i>Books</i>
9780425180631	<i>The Last Precinct (Scarpetta) [Paperback]</i>	3.0	397	<i>Books</i>
9780439023498	<i>Catching Fire (The Second Book of the Hunger Games) [Hardcover]</i>	4.5	643	<i>Books</i>
9780446364492	<i>The Bridges of Madison County [Mass Market Paperback]</i>	3.1	400	<i>Books</i>
9780446677387	<i>Kiss the Girls (2nd Alex Cross Novel) [Paperback]</i>	4.1	375	<i>Books</i>
9780615214795	<i>Maradonia and the Seven Bridges [Paperback]</i>	1.4	22	<i>Books</i>
9780765326263	<i>A Dogs Purpose [Hardcover]</i>	4.8	342	<i>Books</i>
9780810984912	<i>Diary of a Wimpy Kid: The Ugly Truth [Hardcover]</i>	4.5	138	<i>Books</i>
9781596986244	<i>Crimes Against Liberty: An Indictment of President Barack Obama [Hardcover]</i>	3.4	135	<i>Books</i>
9781599211794	<i>The Cult of Perfection: Making Peace with Your Inner Overachiever [Hardcover]</i>	1.2	141	<i>Books</i>
B001LDJIQM	<i>The Apothecarys Daughter [Kindle Edition]</i>	4.3	225	<i>Books</i>
B000065BPB	<i>Sennheiser HD-280 PRO Headphones</i>	4.3	596	<i>Electronics</i>
B000FOYMKU	<i>Coby CVE92 Isolation Stereo Earphones</i>	3.8	946	<i>Electronics</i>
B000OPCH52	<i>Stereo Earbud Headphone for Apple iPod nano/ iPod mini/ iPod video/ iPod shuffle</i>	2.0	87	<i>Electronics</i>
B001W3QW5G	<i>3G Iphone Earphones with Mic</i>	2.2	88	<i>Electronics</i>
B002SQK2F2	<i>iMicro SP-IM942 Headset with Microphone</i>	3.1	263	<i>Electronics</i>

1.2 Tags

Na Tabela 19 são apresentadas as tags da amostra da Amazon classificadas durante este trabalho, são apresentadas em várias colunas só para facilitar a visualização.

Tabela 19 - Tags classificadas da amostra da Amazon

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>Obamosque</i>	<i>neutral</i>	<i>fox news skill</i>	<i>neutral</i>	<i>obamanomics</i>	<i>neutral</i>
<i>1st person</i>	<i>neutral</i>	<i>fraudulent</i>	<i>negative</i>	<i>ok for the price</i>	<i>positive</i>
<i>2009 amazon frustration-fre e packaging revolution</i>	<i>neutral</i>	<i>free case</i>	<i>neutral</i>	<i>old</i>	<i>negative</i>
<i>286</i>	<i>neutral</i>	<i>free ipod</i>	<i>neutral</i>	<i>omg those poor trees</i>	<i>negative</i>
<i>2nd in series</i>	<i>neutral</i>	<i>freedom</i>	<i>neutral</i>	<i>opinionated</i>	<i>negative</i>
<i>3g iphone earphones with mic</i>	<i>neutral</i>	<i>freedom fight</i>	<i>neutral</i>	<i>oracle of devi</i>	<i>neutral</i>
<i>3g iphones with mic</i>	<i>neutral</i>	<i>fun and tough heroic heroine</i>	<i>positive</i>	<i>original ipod headphones</i>	<i>neutral</i>
<i>3rd book</i>	<i>neutral</i>	<i>fun read</i>	<i>positive</i>	<i>over indulged children</i>	<i>neutral</i>
<i>3rd generation ipod touch</i>	<i>neutral</i>	<i>funny</i>	<i>positive</i>	<i>over the head</i>	<i>neutral</i>
<i>5 star books</i>	<i>positive</i>	<i>funny as hell</i>	<i>positive</i>	<i>overachiever</i>	<i>neutral</i>
<i>9 99 boycott</i>	<i>neutral</i>	<i>funny book</i>	<i>positive</i>	<i>overrated</i>	<i>neutral</i>
<i>a dogs purpose</i>	<i>neutral</i>	<i>funny books</i>	<i>positive</i>	<i>oversimplified</i>	<i>negative</i>
<i>a scattered life</i>	<i>neutral</i>	<i>funny mystery</i>	<i>positive</i>	<i>over-the-ear</i>	<i>neutral</i>
<i>a secret in the attic</i>	<i>neutral</i>	<i>future</i>	<i>neutral</i>	<i>over-the-head headphones</i>	<i>neutral</i>
<i>absorbing</i>	<i>positive</i>	<i>futuristic</i>	<i>neutral</i>	<i>overwritten</i>	<i>neutral</i>
<i>abusive</i>	<i>negative</i>	<i>gadgets</i>	<i>neutral</i>	<i>page turner</i>	<i>positive</i>
<i>accessories</i>	<i>neutral</i>	<i>games</i>	<i>neutral</i>	<i>page-turner</i>	<i>positive</i>
<i>action</i>	<i>neutral</i>	<i>gaming headphones</i>	<i>neutral</i>	<i>palin 2012</i>	<i>neutral</i>
<i>action adventure</i>	<i>neutral</i>	<i>gaming headset</i>	<i>neutral</i>	<i>palin2012</i>	<i>neutral</i>
<i>action and adventure</i>	<i>neutral</i>	<i>garbage</i>	<i>negative</i>	<i>palm t3</i>	<i>neutral</i>
<i>action thriller</i>	<i>neutral</i>	<i>garbage can</i>	<i>negative</i>	<i>paranoia</i>	<i>negative</i>
<i>action-adventur e</i>	<i>neutral</i>	<i>garbage disposal</i>	<i>negative</i>	<i>paranormal romance</i>	<i>neutral</i>
<i>action-adventur e-romance</i>	<i>neutral</i>	<i>gay</i>	<i>neutral</i>	<i>parenting</i>	<i>neutral</i>
<i>addictive</i>	<i>positive</i>	<i>generic apple earphone</i>	<i>neutral</i>	<i>passive noise canceling</i>	<i>neutral</i>
<i>adolescence</i>	<i>neutral</i>	<i>george bush</i>	<i>neutral</i>	<i>pathetic</i>	<i>negative</i>
<i>adult</i>	<i>neutral</i>	<i>german movies</i>	<i>neutral</i>	<i>patricia cornwell</i>	<i>neutral</i>
<i>adult humor</i>	<i>neutral</i>	<i>gibraltar</i>	<i>neutral</i>	<i>patterson</i>	<i>neutral</i>
<i>adultery</i>	<i>neutral</i>	<i>girl with the dragon tadoo</i>	<i>neutral</i>	<i>peace</i>	<i>neutral</i>
<i>adultry</i>	<i>neutral</i>	<i>gloria teusch</i>	<i>neutral</i>	<i>peeta</i>	<i>neutral</i>
<i>adults</i>	<i>neutral</i>	<i>glorified fanfiction</i>	<i>neutral</i>	<i>pepsi</i>	<i>neutral</i>
<i>adventure</i>	<i>neutral</i>	<i>god save america</i>	<i>neutral</i>	<i>pepsi points</i>	<i>neutral</i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>adventure scifi</i>	<i>neutral</i>	<i>good</i>	<i>positive</i>	<i>pepsi stuff</i>	<i>neutral</i>
<i>advice</i>	<i>positive</i>	<i>good bargain</i>	<i>positive</i>	<i>perfect</i>	<i>positive</i>
<i>affair</i>	<i>neutral</i>	<i>good novel</i>	<i>positive</i>	<i>perfect for price</i>	<i>positive</i>
<i>affordable</i>	<i>positive</i>	<i>good price</i>	<i>positive</i>	<i>perfection</i>	<i>positive</i>
<i>a-khanh</i>	<i>neutral</i>	<i>good sound</i>	<i>positive</i>	<i>perfectionist</i>	<i>positive</i>
<i>alcoholic</i>	<i>neutral</i>	<i>good value</i>	<i>positive</i>	<i>permanent obscurity</i>	<i>negative</i>
<i>alcoholism</i>	<i>neutral</i>	<i>government conspiracy</i>	<i>neutral</i>	<i>personal transformation</i>	<i>neutral</i>
<i>alex cross</i>	<i>neutral</i>	<i>governmental corruption</i>	<i>neutral</i>	<i>pet loss</i>	<i>neutral</i>
<i>alex cross series - the best one</i>	<i>positive</i>	<i>grandfathers</i>	<i>neutral</i>	<i>pets</i>	<i>neutral</i>
<i>alternate history</i>	<i>neutral</i>	<i>grandparents</i>	<i>neutral</i>	<i>phone headset</i>	<i>neutral</i>
<i>alternate reality</i>	<i>neutral</i>	<i>graphic novel</i>	<i>neutral</i>	<i>piece of crap</i>	<i>negative</i>
<i>alternative</i>	<i>neutral</i>	<i>greasy</i>	<i>negative</i>	<i>piedmont reading club</i>	<i>neutral</i>
<i>alternative medicine</i>	<i>neutral</i>	<i>great</i>	<i>positive</i>	<i>pippi longstockings</i>	<i>neutral</i>
<i>amaliehoward</i>	<i>neutral</i>	<i>great bang for buck</i>	<i>positive</i>	<i>pirrung</i>	<i>neutral</i>
<i>amature</i>	<i>neutral</i>	<i>great books</i>	<i>positive</i>	<i>plum</i>	<i>neutral</i>
<i>amazing</i>	<i>positive</i>	<i>great ending</i>	<i>positive</i>	<i>pms</i>	<i>neutral</i>
<i>amazing earphones</i>	<i>positive</i>	<i>great noval</i>	<i>positive</i>	<i>pointless</i>	<i>negative</i>
<i>amazing sound reproduction</i>	<i>positive</i>	<i>great novel</i>	<i>positive</i>	<i>poirot</i>	<i>neutral</i>
<i>amazon</i>	<i>neutral</i>	<i>great price</i>	<i>positive</i>	<i>poirot and hastings</i>	<i>neutral</i>
<i>amazon digital services</i>	<i>neutral</i>	<i>great product</i>	<i>positive</i>	<i>police procedural</i>	<i>neutral</i>
<i>amazon preorder fail</i>	<i>neutral</i>	<i>great sound</i>	<i>positive</i>	<i>political corruption</i>	<i>neutral</i>
<i>andrew breitbart</i>	<i>neutral</i>	<i>great sounds</i>	<i>positive</i>	<i>politics</i>	<i>neutral</i>
<i>andy brazil</i>	<i>neutral</i>	<i>great value</i>	<i>positive</i>	<i>polyphony be you and accept me too</i>	<i>neutral</i>
<i>angst</i>	<i>neutral</i>	<i>great writer</i>	<i>positive</i>	<i>pompous</i>	<i>negative</i>
<i>animal behavior</i>	<i>neutral</i>	<i>greg heffley</i>	<i>neutral</i>	<i>poop</i>	<i>negative</i>
<i>animal lovers</i>	<i>neutral</i>	<i>grief therapy</i>	<i>positive</i>	<i>poop-dick</i>	<i>negative</i>
<i>animal shelters</i>	<i>neutral</i>	<i>grigham review</i>	<i>neutral</i>	<i>poor</i>	<i>negative</i>
<i>animal stories</i>	<i>neutral</i>	<i>grimmy-rejected</i>	<i>neutral</i>	<i>poor quality</i>	<i>negative</i>
<i>anthony earphones</i>	<i>neutral</i>	<i>grip</i>	<i>undefined</i>	<i>poor research</i>	<i>negative</i>
<i>anti-christian</i>	<i>neutral</i>	<i>grisham</i>	<i>neutral</i>	<i>poor service</i>	<i>neutral</i>
<i>anti-marxist</i>	<i>neutral</i>	<i>grisham novel</i>	<i>neutral</i>	<i>poor sound quality</i>	<i>negative</i>
<i>a-phuong</i>	<i>neutral</i>	<i>growing up</i>	<i>neutral</i>	<i>poor trees</i>	<i>negative</i>
<i>apothecary</i>	<i>neutral</i>	<i>guy</i>	<i>neutral</i>	<i>poor writing</i>	<i>negative</i>
<i>apple</i>	<i>neutral</i>	<i>hack</i>	<i>neutral</i>	<i>poorly written</i>	<i>negative</i>
<i>apryl skies</i>	<i>neutral</i>	<i>hackers</i>	<i>neutral</i>	<i>portable</i>	<i>neutral</i>
<i>arena</i>	<i>neutral</i>	<i>handsfree</i>	<i>neutral</i>	<i>pos</i>	<i>neutral</i>
<i>arm band for nana</i>	<i>undefined</i>	<i>harbinger of doom</i>	<i>neutral</i>	<i>post traumatic stress</i>	<i>neutral</i>
<i>around-the-ear headphones</i>	<i>neutral</i>	<i>hard to put down</i>	<i>positive</i>	<i>post-apocalypti c</i>	<i>neutral</i>
<i>article 2</i>	<i>neutral</i>	<i>hd280</i>	<i>neutral</i>	<i>power</i>	<i>neutral</i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>arrogant</i>	<i>negative</i>	<i>hd 280</i>	<i>neutral</i>	<i>post-apocalyptic dystopian action-adventure</i>	<i>neutral</i>
<i>ashland book shoppe</i>	<i>neutral</i>	<i>hd-280</i>	<i>neutral</i>	<i>pre ordering books from amazon is horrible</i>	<i>neutral</i>
<i>asinine</i>	<i>negative</i>	<i>hd-280 headphone</i>	<i>neutral</i>	<i>presidency</i>	<i>neutral</i>
<i>aspergers</i>	<i>neutral</i>	<i>hd595</i>	<i>neutral</i>	<i>president</i>	<i>neutral</i>
<i>asphyxia</i>	<i>neutral</i>	<i>head phones</i>	<i>neutral</i>	<i>president obama</i>	<i>neutral</i>
<i>audacity</i>	<i>neutral</i>	<i>headache</i>	<i>negative</i>	<i>pre-teen</i>	<i>neutral</i>
<i>audio</i>	<i>neutral</i>	<i>headphones</i>	<i>neutral</i>	<i>prim</i>	<i>neutral</i>
<i>audio accessories</i>	<i>neutral</i>	<i>headset</i>	<i>neutral</i>	<i>production</i>	<i>neutral</i>
<i>audiophile</i>	<i>neutral</i>	<i>headset 1</i>	<i>neutral</i>	<i>promises</i>	<i>neutral</i>
<i>autism</i>	<i>neutral</i>	<i>healing</i>	<i>neutral</i>	<i>proselytizing</i>	<i>neutral</i>
<i>autobiography</i>	<i>neutral</i>	<i>heavy</i>	<i>negative</i>	<i>pseudo-science</i>	<i>neutral</i>
<i>average</i>	<i>neutral</i>	<i>help</i>	<i>positive</i>	<i>psudo-science</i>	<i>neutral</i>
<i>avoid</i>	<i>negative</i>	<i>hemi-sync</i>	<i>neutral</i>	<i>psychology</i>	<i>neutral</i>
<i>avoid at all costs</i>	<i>negative</i>	<i>her daughter s dream</i>	<i>neutral</i>	<i>puberty</i>	<i>neutral</i>
<i>awesome</i>	<i>positive</i>	<i>herbal medicine</i>	<i>neutral</i>	<i>publish my review</i>	<i>neutral</i>
<i>awesome headphones</i>	<i>positive</i>	<i>herbal remedies</i>	<i>neutral</i>	<i>punk</i>	<i>neutral</i>
<i>awful</i>	<i>negative</i>	<i>hi fi</i>	<i>neutral</i>	<i>purpose</i>	<i>neutral</i>
<i>b001hbibny</i>	<i>neutral</i>	<i>hi-fi</i>	<i>neutral</i>	<i>pwned</i>	<i>neutral</i>
<i>bad</i>	<i>negative</i>	<i>high quality</i>	<i>positive</i>	<i>quality</i>	<i>positive</i>
<i>bad advice</i>	<i>negative</i>	<i>high quality headphones</i>	<i>positive</i>	<i>quality tones</i>	<i>positive</i>
<i>bad book</i>	<i>negative</i>	<i>highly recommended</i>	<i>positive</i>	<i>quick read</i>	<i>neutral</i>
<i>bad sound</i>	<i>negative</i>	<i>hilarious</i>	<i>positive</i>	<i>quiet earphones</i>	<i>positive</i>
<i>bad writing</i>	<i>negative</i>	<i>hip hop quality headphones</i>	<i>positive</i>	<i>quirky humor</i>	<i>positive</i>
<i>baloney</i>	<i>negative</i>	<i>historical</i>	<i>neutral</i>	<i>quotations</i>	<i>neutral</i>
<i>banal</i>	<i>negative</i>	<i>historical fiction</i>	<i>neutral</i>	<i>quotes</i>	<i>neutral</i>
<i>barack obama</i>	<i>neutral</i>	<i>historical fiction- england</i>	<i>neutral</i>	<i>race conflict</i>	<i>neutral</i>
<i>bargain</i>	<i>positive</i>	<i>historical romance</i>	<i>neutral</i>	<i>race relations</i>	<i>neutral</i>
<i>barry soetoro</i>	<i>neutral</i>	<i>historical romance with unique setting</i>	<i>neutral</i>	<i>raceciever podcasts</i>	<i>neutral</i>
<i>bass</i>	<i>neutral</i>	<i>history</i>	<i>neutral</i>	<i>racism</i>	<i>neutral</i>
<i>battle royale</i>	<i>neutral</i>	<i>history of medicine</i>	<i>neutral</i>	<i>raymond masters</i>	<i>neutral</i>
<i>beautiful lies</i>	<i>neutral</i>	<i>hit list</i>	<i>positive</i>	<i>read read read</i>	<i>positive</i>
<i>beautiful story</i>	<i>positive</i>	<i>homeschool</i>	<i>neutral</i>	<i>read this</i>	<i>positive</i>
<i>bet me</i>	<i>neutral</i>	<i>honest</i>	<i>positive</i>	<i>reading is fun</i>	<i>positive</i>
<i>bethany house publishers</i>	<i>neutral</i>	<i>honesty</i>	<i>positive</i>	<i>reality</i>	<i>neutral</i>
<i>better than overpriced apple</i>	<i>positive</i>	<i>hope</i>	<i>neutral</i>	<i>reality shows</i>	<i>neutral</i>
<i>beware of canned reviews</i>	<i>negative</i>	<i>horrible</i>	<i>negative</i>	<i>rebellion</i>	<i>neutral</i>
<i>beyond funny</i>	<i>positive</i>	<i>horrible false idiot</i>	<i>negative</i>	<i>recomiendo earphones coby</i>	<i>positive</i>
<i>beyond salvation</i>	<i>neutral</i>	<i>horror</i>	<i>neutral</i>	<i>recommended</i>	<i>positive</i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>bias</i>	<i>negative</i>	<i>hoser</i>	<i>negative</i>	<i>recording</i>	<i>neutral</i>
<i>biased</i>	<i>negative</i>	<i>how long can you read it with a straight face</i>	<i>neutral</i>	<i>regency romance</i>	<i>neutral</i>
<i>biases</i>	<i>negative</i>	<i>hubris</i>	<i>neutral</i>	<i>rehashed</i>	<i>negative</i>
<i>big and uncomfortable</i>	<i>negative</i>	<i>human-animal bond</i>	<i>neutral</i>	<i>reincarnation</i>	<i>neutral</i>
<i>big government</i>	<i>neutral</i>	<i>humor</i>	<i>neutral</i>	<i>relationships</i>	<i>neutral</i>
<i>bigot</i>	<i>negative</i>	<i>humor advise instruction laugh dad son</i>	<i>positive</i>	<i>reluctant hero</i>	<i>neutral</i>
<i>bigoted</i>	<i>negative</i>	<i>humorous father and son</i>	<i>positive</i>	<i>reluctant readers</i>	<i>neutral</i>
<i>bigotry</i>	<i>negative</i>	<i>hunger</i>	<i>neutral</i>	<i>remarriage</i>	<i>neutral</i>
<i>biography</i>	<i>neutral</i>	<i>hunger game</i>	<i>neutral</i>	<i>repetitive</i>	<i>negative</i>
<i>blomkvist</i>	<i>neutral</i>	<i>hunger games</i>	<i>neutral</i>	<i>repression</i>	<i>neutral</i>
<i>bloomquist</i>	<i>neutral</i>	<i>hunger games series</i>	<i>neutral</i>	<i>responsibility</i>	<i>neutral</i>
<i>blues blueshaq rick shaq goldstein</i>	<i>neutral</i>	<i>hunger gamnes</i>	<i>neutral</i>	<i>revenge</i>	<i>neutral</i>
<i>book</i>	<i>neutral</i>	<i>hypocrisy</i>	<i>negative</i>	<i>review</i>	<i>neutral</i>
<i>book of fiction</i>	<i>neutral</i>	<i>hypocrit</i>	<i>negative</i>	<i>review by clayton bye</i>	<i>neutral</i>
<i>book recommendations</i>	<i>positive</i>	<i>hypocrite</i>	<i>negative</i>	<i>reviews</i>	<i>neutral</i>
<i>book review</i>	<i>neutral</i>	<i>i own it</i>	<i>neutral</i>	<i>revolution</i>	<i>neutral</i>
<i>books for boys</i>	<i>neutral</i>	<i>i stanked my pants</i>	<i>neutral</i>	<i>right wing propaganda</i>	<i>neutral</i>
<i>books for girls</i>	<i>neutral</i>	<i>i will wait for you</i>	<i>neutral</i>	<i>right wing radicalism</i>	<i>neutral</i>
<i>books turned into to tv shows</i>	<i>neutral</i>	<i>idiot</i>	<i>negative</i>	<i>ripoff</i>	<i>neutral</i>
<i>booksilly</i>	<i>negative</i>	<i>idiotic</i>	<i>negative</i>	<i>rivercrap</i>	<i>negative</i>
<i>booze</i>	<i>neutral</i>	<i>iem</i>	<i>neutral</i>	<i>robert james waller</i>	<i>neutral</i>
<i>bored</i>	<i>negative</i>	<i>ignorance</i>	<i>negative</i>	<i>romance</i>	<i>neutral</i>
<i>bor-hbk-0810</i>	<i>neutral</i>	<i>ignorant</i>	<i>negative</i>	<i>romantic</i>	<i>neutral</i>
<i>boring</i>	<i>negative</i>	<i>ill bring tinker is headed off to the office - twill be another</i>	<i>neutral</i>	<i>rowley jefferson</i>	<i>neutral</i>
<i>bounty hunter</i>	<i>neutral</i>	<i>immersion</i>	<i>undefined</i>	<i>rubbish</i>	<i>negative</i>
<i>broken</i>	<i>negative</i>	<i>impeach oama</i>	<i>neutral</i>	<i>rude</i>	<i>negative</i>
<i>bruce cameron</i>	<i>neutral</i>	<i>in ear</i>	<i>neutral</i>	<i>rush limbaugh</i>	<i>neutral</i>
<i>bud</i>	<i>neutral</i>	<i>in ear headphones</i>	<i>neutral</i>	<i>russion defector</i>	<i>neutral</i>
<i>budget earphones</i>	<i>neutral</i>	<i>inaccurate</i>	<i>negative</i>	<i>sam halpern</i>	<i>neutral</i>
<i>buds</i>	<i>neutral</i>	<i>inane</i>	<i>negative</i>	<i>sarcasm</i>	<i>negative</i>
<i>burn the koran</i>	<i>neutral</i>	<i>incorrect</i>	<i>negative</i>	<i>satire</i>	<i>neutral</i>
<i>busted</i>	<i>neutral</i>	<i>in-ear</i>	<i>neutral</i>	<i>save america</i>	<i>neutral</i>
<i>but terrible language</i>	<i>negative</i>	<i>in-ear earphones</i>	<i>neutral</i>	<i>save your money</i>	<i>negative</i>
<i>butchery</i>	<i>negative</i>	<i>in-ear headphones</i>	<i>neutral</i>	<i>scammer</i>	<i>neutral</i>
<i>campaign promises</i>	<i>neutral</i>	<i>in-ear monitors</i>	<i>neutral</i>	<i>scandanavian noir</i>	<i>neutral</i>
<i>can it be nathiocool</i>	<i>neutral</i>	<i>inexpensive</i>	<i>positive</i>	<i>scandinavian literature</i>	<i>neutral</i>
<i>can t wait to read it</i>	<i>positive</i>	<i>innacurate</i>	<i>negative</i>	<i>scarpetta</i>	<i>neutral</i>
<i>capital punishment</i>	<i>neutral</i>	<i>inner</i>	<i>neutral</i>	<i>schlock</i>	<i>negative</i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>career choices</i>	<i>neutral</i>	<i>inner overachiever</i>	<i>neutral</i>	<i>science fiction</i>	<i>neutral</i>
<i>career ending</i>	<i>neutral</i>	<i>inner peace</i>	<i>neutral</i>	<i>science thriller</i>	<i>neutral</i>
<i>catching</i>	<i>neutral</i>	<i>insightful</i>	<i>positive</i>	<i>scientology</i>	<i>neutral</i>
<i>catching fire</i>	<i>neutral</i>	<i>inspirational</i>	<i>positive</i>	<i>scifi</i>	<i>neutral</i>
<i>catching fire suzanne collins</i>	<i>neutral</i>	<i>integrity</i>	<i>positive</i>	<i>sci-fi action</i>	<i>neutral</i>
<i>challenges</i>	<i>neutral</i>	<i>interesting</i>	<i>positive</i>	<i>sci-fi adventure</i>	<i>neutral</i>
<i>challenging</i>	<i>positive</i>	<i>international best seller</i>	<i>neutral</i>	<i>sci-fi romance</i>	<i>neutral</i>
<i>change</i>	<i>neutral</i>	<i>internets</i>	<i>neutral</i>	<i>search and rescue</i>	<i>neutral</i>
<i>cheap</i>	<i>neutral</i>	<i>intertament</i>	<i>neutral</i>	<i>second</i>	<i>neutral</i>
<i>cheap ear buds</i>	<i>neutral</i>	<i>intolerant</i>	<i>negative</i>	<i>secret service</i>	<i>neutral</i>
<i>cheap earbuds</i>	<i>neutral</i>	<i>intrigue</i>	<i>neutral</i>	<i>secrets of a lab rat</i>	<i>neutral</i>
<i>cheap earphones</i>	<i>neutral</i>	<i>intriguing</i>	<i>positive</i>	<i>self defense</i>	<i>neutral</i>
<i>cheap headphones</i>	<i>neutral</i>	<i>investigation</i>	<i>neutral</i>	<i>self-absorbed</i>	<i>neutral</i>
<i>cheating</i>	<i>negative</i>	<i>ipad</i>	<i>neutral</i>	<i>self-help</i>	<i>neutral</i>
<i>check it out</i>	<i>positive</i>	<i>iphone</i>	<i>neutral</i>	<i>self-published</i>	<i>neutral</i>
<i>chick book</i>	<i>neutral</i>	<i>iphone earbuds</i>	<i>neutral</i>	<i>sennheiser hd-280 pro headphones</i>	<i>neutral</i>
<i>child rearing</i>	<i>neutral</i>	<i>iphone earphones</i>	<i>neutral</i>	<i>sennheiser</i>	<i>neutral</i>
<i>childhood</i>	<i>neutral</i>	<i>iphone headphones</i>	<i>neutral</i>	<i>sennheiser 280</i>	<i>neutral</i>
<i>childhood memories</i>	<i>neutral</i>	<i>iphone microphone</i>	<i>neutral</i>	<i>sennheiser 280pro headphones</i>	<i>neutral</i>
<i>childrens books</i>	<i>neutral</i>	<i>iphone replacement earphones</i>	<i>neutral</i>	<i>sennheiser broken</i>	<i>neutral</i>
<i>choices</i>	<i>neutral</i>	<i>ipod</i>	<i>neutral</i>	<i>sennheiser hd-280</i>	<i>neutral</i>
<i>christian crap</i>	<i>neutral</i>	<i>ipod accessories</i>	<i>neutral</i>	<i>sennheiser hd-280 professional headphones</i>	<i>neutral</i>
<i>christian fiction</i>	<i>neutral</i>	<i>ipod ear pain</i>	<i>negative</i>	<i>sennheiser headphones</i>	<i>neutral</i>
<i>christian romance</i>	<i>neutral</i>	<i>ipod earbuds</i>	<i>neutral</i>	<i>sennheiser_hd-280_headphones</i>	<i>neutral</i>
<i>christianity</i>	<i>neutral</i>	<i>ipod earphones</i>	<i>neutral</i>	<i>sentimental</i>	<i>neutral</i>
<i>circumaural</i>	<i>neutral</i>	<i>ipod headphones</i>	<i>neutral</i>	<i>sequels</i>	<i>neutral</i>
<i>circum-aural</i>	<i>neutral</i>	<i>ipod headset</i>	<i>neutral</i>	<i>serial killer</i>	<i>neutral</i>
<i>circumaural headphones</i>	<i>neutral</i>	<i>ipod mp3</i>	<i>neutral</i>	<i>series</i>	<i>neutral</i>
<i>citizenship at birth</i>	<i>neutral</i>	<i>ipod nano</i>	<i>neutral</i>	<i>sex</i>	<i>neutral</i>
<i>classic literature</i>	<i>neutral</i>	<i>ipod nano headset</i>	<i>neutral</i>	<i>sham</i>	<i>negative</i>
<i>clean romance</i>	<i>positive</i>	<i>ipod stereo headset</i>	<i>neutral</i>	<i>shatner</i>	<i>neutral</i>
<i>cleaving</i>	<i>neutral</i>	<i>ipod touch</i>	<i>neutral</i>	<i>shipping</i>	<i>neutral</i>
<i>cliche</i>	<i>negative</i>	<i>irresponsible</i>	<i>negative</i>	<i>shocking</i>	<i>undefined</i>
<i>clint eastwood</i>	<i>neutral</i>	<i>irresponsible journalism</i>	<i>negative</i>	<i>signed john grisham book</i>	<i>neutral</i>
<i>closed</i>	<i>neutral</i>	<i>isolating</i>	<i>neutral</i>	<i>simpleton</i>	<i>negative</i>
<i>closed back headphones</i>	<i>neutral</i>	<i>isolation</i>	<i>neutral</i>	<i>simplistic drivell</i>	<i>negative</i>
<i>closed headphone</i>	<i>neutral</i>	<i>james patterson</i>	<i>neutral</i>	<i>size</i>	<i>neutral</i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>closed headphones</i>	<i>neutral</i>	<i>janet evanovich</i>	<i>neutral</i>	<i>skye the troll</i>	<i>neutral</i>
<i>closed-ear</i>	<i>neutral</i>	<i>jbuds</i>	<i>neutral</i>	<i>skye the troll and other fairy tales for children</i>	<i>neutral</i>
<i>coached</i>	<i>neutral</i>	<i>jeff kinney</i>	<i>neutral</i>	<i>small ears</i>	<i>neutral</i>
<i>coby</i>	<i>neutral</i>	<i>jim crow</i>	<i>neutral</i>	<i>smear campaign</i>	<i>negative</i>
<i>coby cve92 isolation stereo earphones earbuds</i>	<i>neutral</i>	<i>jlal</i>	<i>neutral</i>	<i>so bad it sporks itself</i>	<i>negative</i>
<i>coby cve92 isolation stereo earphones review</i>	<i>neutral</i>	<i>job search</i>	<i>neutral</i>	<i>so sweet</i>	<i>positive</i>
<i>coby cve92 isolation stereo earphones xboxkrazy</i>	<i>neutral</i>	<i>john grisham</i>	<i>neutral</i>	<i>social unrest</i>	<i>neutral</i>
<i>coby earbuds</i>	<i>neutral</i>	<i>journey</i>	<i>neutral</i>	<i>socialism</i>	<i>neutral</i>
<i>coby earphones</i>	<i>neutral</i>	<i>judgemental</i>	<i>negative</i>	<i>socialists and other anti-american extremists</i>	<i>neutral</i>
<i>coby headphones</i>	<i>neutral</i>	<i>judicial system</i>	<i>neutral</i>	<i>solid bass</i>	<i>neutral</i>
<i>coby headphones inears</i>	<i>neutral</i>	<i>julie klassen</i>	<i>neutral</i>	<i>something very diferrent</i>	<i>neutral</i>
<i>coby iso bud headset</i>	<i>neutral</i>	<i>julie powell</i>	<i>neutral</i>	<i>son</i>	<i>neutral</i>
<i>coby isolation earphones</i>	<i>neutral</i>	<i>junior audiophile</i>	<i>neutral</i>	<i>sons</i>	<i>neutral</i>
<i>coby super bass</i>	<i>positive</i>	<i>junk</i>	<i>negative</i>	<i>soul</i>	<i>neutral</i>
<i>cold war</i>	<i>neutral</i>	<i>junk ear buds</i>	<i>negative</i>	<i>sound</i>	<i>neutral</i>
<i>collins</i>	<i>neutral</i>	<i>just plain dumb</i>	<i>negative</i>	<i>sound blocking</i>	<i>neutral</i>
<i>collins catching fire</i>	<i>neutral</i>	<i>justice</i>	<i>neutral</i>	<i>sound enriching</i>	<i>positive</i>
<i>collins hunger games</i>	<i>neutral</i>	<i>justin halpern</i>	<i>neutral</i>	<i>sound isolating</i>	<i>neutral</i>
<i>comedy</i>	<i>neutral</i>	<i>juvenile humor</i>	<i>neutral</i>	<i>sound isolating headphones</i>	<i>neutral</i>
<i>comfort</i>	<i>positive</i>	<i>kathy</i>	<i>neutral</i>	<i>sound isolation</i>	<i>neutral</i>
<i>comfortable</i>	<i>positive</i>	<i>kathy anne pippig harris</i>	<i>neutral</i>	<i>specifics</i>	<i>neutral</i>
<i>comfortable ear</i>	<i>positive</i>	<i>katniss</i>	<i>neutral</i>	<i>speculative fiction</i>	<i>neutral</i>
<i>comical</i>	<i>neutral</i>	<i>katniss everdeen</i>	<i>neutral</i>	<i>spies</i>	<i>neutral</i>
<i>coming of age</i>	<i>neutral</i>	<i>kay scarpetta</i>	<i>neutral</i>	<i>spirituality</i>	<i>neutral</i>
<i>commitment</i>	<i>neutral</i>	<i>kenyan</i>	<i>neutral</i>	<i>splinter skills</i>	<i>neutral</i>
<i>communist sympathizer</i>	<i>neutral</i>	<i>keychain</i>	<i>neutral</i>	<i>spy</i>	<i>neutral</i>
<i>compassion</i>	<i>neutral</i>	<i>kids</i>	<i>neutral</i>	<i>steig larrson thriller</i>	<i>neutral</i>
<i>competition</i>	<i>neutral</i>	<i>kids book</i>	<i>neutral</i>	<i>steig larsson</i>	<i>neutral</i>
<i>complete crap</i>	<i>negative</i>	<i>kids love books</i>	<i>neutral</i>	<i>steig larsson</i>	<i>neutral</i>
<i>computer savvy</i>	<i>neutral</i>	<i>kindle</i>	<i>neutral</i>	<i>stephanie plum</i>	<i>neutral</i>
<i>conclusion of the trilogy</i>	<i>neutral</i>	<i>kindle 3</i>	<i>neutral</i>	<i>stereo earphones</i>	<i>neutral</i>
<i>confession</i>	<i>neutral</i>	<i>kindle accessory</i>	<i>neutral</i>	<i>stereo headphones</i>	<i>neutral</i>
<i>conservativism</i>	<i>neutral</i>	<i>kindle book</i>	<i>neutral</i>	<i>stereo nano headset</i>	<i>neutral</i>
<i>conspiracy</i>	<i>neutral</i>	<i>kindle book club</i>	<i>neutral</i>	<i>stieg larssen</i>	<i>neutral</i>
<i>constitution</i>	<i>neutral</i>	<i>kindle earphones</i>	<i>neutral</i>	<i>stieg larsson</i>	<i>neutral</i>
<i>contemporary fiction</i>	<i>neutral</i>	<i>kindle free books</i>	<i>neutral</i>	<i>stockholm</i>	<i>neutral</i>
<i>contrived</i>	<i>negative</i>	<i>kindle freebie</i>	<i>neutral</i>	<i>strange</i>	<i>negative</i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>cooch smells like rotting intestines</i>	<i>neutral</i>	<i>kindle headsets</i>	<i>neutral</i>	<i>strong cord</i>	<i>positive</i>
<i>cooking</i>	<i>neutral</i>	<i>kindle humor pop culture</i>	<i>neutral</i>	<i>strong women</i>	<i>neutral</i>
<i>cooking schmooking</i>	<i>neutral</i>	<i>kindle more than hardcover</i>	<i>neutral</i>	<i>struggling</i>	<i>neutral</i>
<i>cooper lawremce</i>	<i>neutral</i>	<i>kindle version costs more than the physical book</i>	<i>neutral</i>	<i>stuck-up</i>	<i>negative</i>
<i>cooper lawrence</i>	<i>neutral</i>	<i>kindling</i>	<i>negative</i>	<i>studio headphones</i>	<i>neutral</i>
<i>cooper lawrence sucks</i>	<i>neutral</i>	<i>kingston</i>	<i>neutral</i>	<i>studio monitoring headphone</i>	<i>neutral</i>
<i>cornwell</i>	<i>neutral</i>	<i>kinney</i>	<i>neutral</i>	<i>stupid</i>	<i>negative</i>
<i>corruption</i>	<i>neutral</i>	<i>kiss the girls</i>	<i>neutral</i>	<i>stupidity</i>	<i>negative</i>
<i>courtroom</i>	<i>neutral</i>	<i>kiss the girls review</i>	<i>neutral</i>	<i>sucky</i>	<i>negative</i>
<i>covered bridge</i>	<i>neutral</i>	<i>knock-off</i>	<i>undefined</i>	<i>sudokids</i>	<i>neutral</i>
<i>coverup</i>	<i>neutral</i>	<i>koss</i>	<i>neutral</i>	<i>summer read</i>	<i>neutral</i>
<i>cracked hearts</i>	<i>neutral</i>	<i>kristyim50</i>	<i>neutral</i>	<i>super bass</i>	<i>positive</i>
<i>crappy headphones</i>	<i>negative</i>	<i>lacks credibility</i>	<i>negative</i>	<i>superb ending for the series</i>	<i>positive</i>
<i>crapstacular</i>	<i>undefined</i>	<i>lacks originality</i>	<i>negative</i>	<i>superbass</i>	<i>positive</i>
<i>credible</i>	<i>positive</i>	<i>lacks originality</i>	<i>negative</i>	<i>survival</i>	<i>neutral</i>
<i>crime drama</i>	<i>neutral</i>	<i>lame</i>	<i>negative</i>	<i>survival fiction</i>	<i>neutral</i>
<i>crime fiction</i>	<i>neutral</i>	<i>lame conservative tripe</i>	<i>negative</i>	<i>susense</i>	<i>neutral</i>
<i>crime novel</i>	<i>neutral</i>	<i>laptop accessories</i>	<i>neutral</i>	<i>suspense</i>	<i>neutral</i>
<i>crime thriller</i>	<i>neutral</i>	<i>larson would turn over in this grave</i>	<i>neutral</i>	<i>suspense thriller</i>	<i>neutral</i>
<i>crimes</i>	<i>neutral</i>	<i>larsson</i>	<i>neutral</i>	<i>suspenseful</i>	<i>neutral</i>
<i>criminal justice system</i>	<i>neutral</i>	<i>larsson thriller</i>	<i>neutral</i>	<i>suspensful</i>	<i>neutral</i>
<i>cross</i>	<i>neutral</i>	<i>late</i>	<i>neutral</i>	<i>suzanne</i>	<i>neutral</i>
<i>cuisine</i>	<i>neutral</i>	<i>laugh</i>	<i>positive</i>	<i>suzanne collins</i>	<i>neutral</i>
<i>cult</i>	<i>neutral</i>	<i>laugh out loud</i>	<i>positive</i>	<i>suzette standring</i>	<i>neutral</i>
<i>cult of personality</i>	<i>neutral</i>	<i>laugh out loud funny</i>	<i>positive</i>	<i>swearing</i>	<i>negative</i>
<i>current events</i>	<i>neutral</i>	<i>laughs</i>	<i>positive</i>	<i>sweden</i>	<i>neutral</i>
<i>current events republican lies</i>	<i>neutral</i>	<i>laughter</i>	<i>positive</i>	<i>swedish</i>	<i>neutral</i>
<i>cve92</i>	<i>neutral</i>	<i>law</i>	<i>neutral</i>	<i>swedish mysteries</i>	<i>neutral</i>
<i>dad</i>	<i>neutral</i>	<i>lawyer</i>	<i>neutral</i>	<i>swedish mystery king</i>	<i>neutral</i>
<i>dad gifts</i>	<i>neutral</i>	<i>lazy</i>	<i>negative</i>	<i>swedish thriller</i>	<i>neutral</i>
<i>dads</i>	<i>neutral</i>	<i>legal lynching</i>	<i>neutral</i>	<i>t a smith</i>	<i>neutral</i>
<i>dangerous</i>	<i>neutral</i>	<i>legal quagmire of arcaic death penalty systems revealed</i>	<i>neutral</i>	<i>tacky</i>	<i>negative</i>
<i>darien gee</i>	<i>neutral</i>	<i>legal thriller</i>	<i>neutral</i>	<i>tag</i>	<i>neutral</i>
<i>david limbaugh</i>	<i>neutral</i>	<i>lesa trapp</i>	<i>neutral</i>	<i>tap</i>	<i>neutral</i>
<i>death</i>	<i>neutral</i>	<i>lethal injection</i>	<i>neutral</i>	<i>taxed to death</i>	<i>neutral</i>
<i>death penalty</i>	<i>neutral</i>	<i>liar</i>	<i>negative</i>	<i>tea party</i>	<i>neutral</i>
<i>decent headphones</i>	<i>positive</i>	<i>liar lies fraud</i>	<i>negative</i>	<i>ted smith</i>	<i>neutral</i>
<i>deception</i>	<i>negative</i>	<i>liberal</i>	<i>neutral</i>	<i>teen</i>	<i>neutral</i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>deep bass</i>	<i>neutral</i>	<i>liberal threat</i>	<i>neutral</i>	<i>teen adventure</i>	<i>neutral</i>
<i>defectors</i>	<i>neutral</i>	<i>liberty</i>	<i>neutral</i>	<i>teen dystopia</i>	<i>neutral</i>
<i>delight</i>	<i>positive</i>	<i>liberty and tyranny</i>	<i>neutral</i>	<i>teen survival</i>	<i>neutral</i>
<i>delightful</i>	<i>positive</i>	<i>lie</i>	<i>negative</i>	<i>television</i>	<i>neutral</i>
<i>delusional</i>	<i>negative</i>	<i>lies</i>	<i>negative</i>	<i>temporary replacement</i>	<i>neutral</i>
<i>democrat</i>	<i>neutral</i>	<i>life</i>	<i>neutral</i>	<i>terrible</i>	<i>negative</i>
<i>democrats</i>	<i>neutral</i>	<i>life lessons</i>	<i>neutral</i>	<i>terrible books</i>	<i>negative</i>
<i>desperate</i>	<i>neutral</i>	<i>lift</i>	<i>positive</i>	<i>terror</i>	<i>neutral</i>
<i>detective</i>	<i>neutral</i>	<i>lift audio</i>	<i>neutral</i>	<i>test preparation</i>	<i>neutral</i>
<i>detective series</i>	<i>neutral</i>	<i>lift audio groove</i>	<i>neutral</i>	<i>texas</i>	<i>neutral</i>
<i>detrimental</i>	<i>negative</i>	<i>light</i>	<i>neutral</i>	<i>texas justice</i>	<i>neutral</i>
<i>dfgdf</i>	<i>neutral</i>	<i>lightweight</i>	<i>positive</i>	<i>the apothecary</i>	<i>neutral</i>
<i>diaray of a wimpy kid</i>	<i>neutral</i>	<i>lightworkers</i>	<i>neutral</i>	<i>the best</i>	<i>positive</i>
<i>diary</i>	<i>neutral</i>	<i>limbaugh</i>	<i>neutral</i>	<i>the book cursader</i>	<i>neutral</i>
<i>diary of a wimpy kid</i>	<i>neutral</i>	<i>limbaugh and divisiveness</i>	<i>neutral</i>	<i>the clayborne brides</i>	<i>neutral</i>
<i>diary of wimpy kid series</i>	<i>neutral</i>	<i>lisabeth salander</i>	<i>neutral</i>	<i>the confession</i>	<i>neutral</i>
<i>dickies</i>	<i>neutral</i>	<i>lisbeth</i>	<i>neutral</i>	<i>the cutting edge</i>	<i>positive</i>
<i>digital stereo</i>	<i>neutral</i>	<i>lisbeth salader</i>	<i>neutral</i>	<i>the end and the best</i>	<i>positive</i>
<i>disappointing</i>	<i>negative</i>	<i>lisbeth salander</i>	<i>neutral</i>	<i>the fixer upper</i>	<i>neutral</i>
<i>discrimination</i>	<i>neutral</i>	<i>little wild flower</i>	<i>neutral</i>	<i>the gateway</i>	<i>neutral</i>
<i>disposable</i>	<i>negative</i>	<i>logic games</i>	<i>neutral</i>	<i>the hollow man by g patrick huskins</i>	<i>neutral</i>
<i>district 12</i>	<i>neutral</i>	<i>love</i>	<i>neutral</i>	<i>the hunger games</i>	<i>neutral</i>
<i>dj</i>	<i>neutral</i>	<i>love books</i>	<i>neutral</i>	<i>the narrow gate</i>	<i>neutral</i>
<i>dog book</i>	<i>neutral</i>	<i>love dogs</i>	<i>neutral</i>	<i>the passage</i>	<i>neutral</i>
<i>dog fiction</i>	<i>neutral</i>	<i>love story</i>	<i>neutral</i>	<i>the secret trilogy</i>	<i>neutral</i>
<i>dog lovers</i>	<i>neutral</i>	<i>love triangle</i>	<i>neutral</i>	<i>the surgeon</i>	<i>neutral</i>
<i>dog lovers book</i>	<i>neutral</i>	<i>love your pet</i>	<i>neutral</i>	<i>the ugly truth</i>	<i>neutral</i>
<i>dog psychology</i>	<i>neutral</i>	<i>loving</i>	<i>neutral</i>	<i>they suck</i>	<i>negative</i>
<i>dogs</i>	<i>neutral</i>	<i>low cost</i>	<i>positive</i>	<i>this was an awesome book after the first</i>	<i>positive</i>
<i>dogs and humans</i>	<i>neutral</i>	<i>low fi sound</i>	<i>negative</i>	<i>thriller</i>	<i>neutral</i>
<i>dogs and love</i>	<i>neutral</i>	<i>low quality</i>	<i>negative</i>	<i>thriller fiction</i>	<i>neutral</i>
<i>dogs reincarnation</i>	<i>neutral</i>	<i>low volume</i>	<i>negative</i>	<i>thrilling</i>	<i>positive</i>
<i>done</i>	<i>neutral</i>	<i>loyal</i>	<i>positive</i>	<i>tight</i>	<i>negative</i>
<i>dont waste your money</i>	<i>negative</i>	<i>lsat</i>	<i>neutral</i>	<i>tiresome</i>	<i>negative</i>
<i>door stopper</i>	<i>neutral</i>	<i>lunatic fringe</i>	<i>neutral</i>	<i>titillating mystery</i>	<i>neutral</i>
<i>drivel</i>	<i>negative</i>	<i>many lives</i>	<i>neutral</i>	<i>to check</i>	<i>positive</i>
<i>dumb</i>	<i>negative</i>	<i>maradonia</i>	<i>neutral</i>	<i>to read list</i>	<i>positive</i>
<i>dumb broad</i>	<i>negative</i>	<i>marital relations</i>	<i>neutral</i>	<i>toilet paper</i>	<i>negative</i>
<i>dumbed down</i>	<i>negative</i>	<i>marriage</i>	<i>neutral</i>	<i>tom dulaney</i>	<i>neutral</i>
<i>durban</i>	<i>neutral</i>	<i>marxism</i>	<i>neutral</i>	<i>too much fluff</i>	<i>negative</i>
<i>dysfunctional family</i>	<i>neutral</i>	<i>mass</i>	<i>neutral</i>	<i>top dog</i>	<i>positive</i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i> durable </i>	<i> positive </i>	<i> marvelous reader </i>	<i> positive </i>	<i> too many long named swedes </i>	<i> neutral </i>
<i> dyst </i>	<i> neutral </i>	<i> mass effect rules </i>	<i> neutral </i>	<i> touching </i>	<i> positive </i>
<i> dystopia </i>	<i> neutral </i>	<i> mawkishness </i>	<i> negative </i>	<i> tough females </i>	<i> neutral </i>
<i> dystopia- adventure </i>	<i> neutral </i>	<i> maybe this time </i>	<i> neutral </i>	<i> traditions </i>	<i> neutral </i>
<i> dystopian </i>	<i> neutral </i>	<i> mccarthyistic </i>	<i> neutral </i>	<i> transcription headset </i>	<i> neutral </i>
<i> dystopian fiction </i>	<i> neutral </i>	<i> meaningless </i>	<i> negative </i>	<i> trash </i>	<i> negative </i>
<i> dystopian society </i>	<i> neutral </i>	<i> meat </i>	<i> neutral </i>	<i> trashy book </i>	<i> negative </i>
<i> ear </i>	<i> neutral </i>	<i> media </i>	<i> neutral </i>	<i> treble </i>	<i> neutral </i>
<i> ear buds </i>	<i> neutral </i>	<i> medical </i>	<i> neutral </i>	<i> trendy </i>	<i> positive </i>
<i> ear pain </i>	<i> negative </i>	<i> medical examiners </i>	<i> neutral </i>	<i> trilogy </i>	<i> neutral </i>
<i> earbud </i>	<i> neutral </i>	<i> medical fiction </i>	<i> neutral </i>	<i> tripe </i>	<i> negative </i>
<i> earbud headphones </i>	<i> neutral </i>	<i> medical mystery </i>	<i> neutral </i>	<i> triple exposure </i>	<i> neutral </i>
<i> earbuds </i>	<i> neutral </i>	<i> mediocre </i>	<i> negative </i>	<i> true crime </i>	<i> neutral </i>
<i> earbuds coby </i>	<i> neutral </i>	<i> memior </i>	<i> neutral </i>	<i> true future </i>	<i> neutral </i>
<i> earcup </i>	<i> neutral </i>	<i> memoir </i>	<i> neutral </i>	<i> trust </i>	<i> neutral </i>
<i> earpeace </i>	<i> positive </i>	<i> memories </i>	<i> neutral </i>	<i> truth </i>	<i> neutral </i>
<i> earphone </i>	<i> neutral </i>	<i> mentally ill </i>	<i> neutral </i>	<i> truth hurts </i>	<i> neutral </i>
<i> earphones </i>	<i> neutral </i>	<i> meryl streep </i>	<i> neutral </i>	<i> tts enabled </i>	<i> neutral </i>
<i> earphones coby </i>	<i> neutral </i>	<i> mia king </i>	<i> neutral </i>	<i> twb </i>	<i> neutral </i>
<i> earphones earbuds headphones ipod coby bass stereo </i>	<i> neutral </i>	<i> miami snow </i>	<i> neutral </i>	<i> twit </i>	<i> negative </i>
<i> earphones for iphone </i>	<i> neutral </i>	<i> mic </i>	<i> neutral </i>	<i> twitter </i>	<i> neutral </i>
<i> earphones with bass </i>	<i> neutral </i>	<i> middle grade boy book </i>	<i> neutral </i>	<i> tyranny </i>	<i> neutral </i>
<i> ebook </i>	<i> neutral </i>	<i> middle school </i>	<i> neutral </i>	<i> tyranny by default </i>	<i> neutral </i>
<i> ebook reader </i>	<i> neutral </i>	<i> mikael blomkvist </i>	<i> neutral </i>	<i> ultimateears </i>	<i> neutral </i>
<i> economic collapse </i>	<i> neutral </i>	<i> milk maid in heaven </i>	<i> neutral </i>	<i> unclear </i>	<i> negative </i>
<i> economical </i>	<i> positive </i>	<i> millenium </i>	<i> neutral </i>	<i> uncomfortable </i>	<i> negative </i>
<i> effect </i>	<i> neutral </i>	<i> millenium trilogy </i>	<i> neutral </i>	<i> unedited </i>	<i> neutral </i>
<i> ego </i>	<i> neutral </i>	<i> millennium </i>	<i> neutral </i>	<i> uninformed </i>	<i> negative </i>
<i> ejm kindle 9 99 tag </i>	<i> neutral </i>	<i> mind-opening </i>	<i> neutral </i>	<i> unlicensed pyschologist </i>	<i> negative </i>
<i> election 2008 </i>	<i> neutral </i>	<i> misconduct </i>	<i> neutral </i>	<i> unoriginal </i>	<i> negative </i>
<i> electronics </i>	<i> neutral </i>	<i> misinformative </i>	<i> negative </i>	<i> unreadable </i>	<i> negative </i>
<i> embarrassing moments </i>	<i> neutral </i>	<i> misinformed </i>	<i> negative </i>	<i> unresearched </i>	<i> negative </i>
<i> emotional </i>	<i> neutral </i>	<i> mistake </i>	<i> negative </i>	<i> us constitution </i>	<i> neutral </i>
<i> enemies and playmates </i>	<i> neutral </i>	<i> mistakes </i>	<i> negative </i>	<i> useless </i>	<i> negative </i>
<i> entertaining </i>	<i> positive </i>	<i> mockingjay </i>	<i> neutral </i>	<i> useless crap </i>	<i> negative </i>
<i> epic adventure </i>	<i> neutral </i>	<i> modern family </i>	<i> neutral </i>	<i> utopia </i>	<i> neutral </i>
<i> epic fail </i>	<i> negative </i>	<i> moronic </i>	<i> negative </i>	<i> value </i>	<i> positive </i>
<i> equalizer </i>	<i> undefined </i>	<i> mothers day </i>	<i> neutral </i>	<i> values </i>	<i> neutral </i>
<i> estrus </i>	<i> neutral </i>	<i> motivating readers </i>	<i> positive </i>	<i> vanity </i>	<i> negative </i>
<i> evanovich </i>	<i> neutral </i>	<i> motorcycle gang </i>	<i> neutral </i>	<i> vanity publishing </i>	<i> neutral </i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>evil</i>	<i>negative</i>	<i>mouthpiece</i>	<i>neutral</i>	<i>very funny</i>	<i>positive</i>
<i>excellent</i>	<i>positive</i>	<i>mp3 accessories</i>	<i>neutral</i>	<i>video games</i>	<i>neutral</i>
<i>excellent book</i>	<i>positive</i>	<i>mp3 headphones</i>	<i>neutral</i>	<i>video production</i>	<i>neutral</i>
<i>excellent sound</i>	<i>positive</i>	<i>mp3 player</i>	<i>neutral</i>	<i>vincent zandri</i>	<i>neutral</i>
<i>exciting</i>	<i>positive</i>	<i>muffins</i>	<i>neutral</i>	<i>violence</i>	<i>neutral</i>
<i>executive branch</i>	<i>neutral</i>	<i>murder</i>	<i>neutral</i>	<i>vision</i>	<i>neutral</i>
<i>explosion in paris</i>	<i>neutral</i>	<i>murder mystery</i>	<i>neutral</i>	<i>wackbuds</i>	<i>negative</i>
<i>exposing leftist ideology</i>	<i>neutral</i>	<i>murder trial</i>	<i>neutral</i>	<i>wait for mass market paperback</i>	<i>neutral</i>
<i>extraordinary</i>	<i>positive</i>	<i>music</i>	<i>neutral</i>	<i>wait for paperback</i>	<i>neutral</i>
<i>facebook</i>	<i>neutral</i>	<i>musim</i>	<i>neutral</i>	<i>waiting for kindle edition</i>	<i>neutral</i>
<i>fail</i>	<i>negative</i>	<i>muslim</i>	<i>neutral</i>	<i>waitress</i>	<i>neutral</i>
<i>failure</i>	<i>negative</i>	<i>muslim liar</i>	<i>neutral</i>	<i>waller</i>	<i>neutral</i>
<i>fair</i>	<i>positive</i>	<i>my dad says</i>	<i>neutral</i>	<i>wandas evil scheme</i>	<i>neutral</i>
<i>fair price</i>	<i>positive</i>	<i>mysteries and thrillers</i>	<i>neutral</i>	<i>war</i>	<i>neutral</i>
<i>faith</i>	<i>neutral</i>	<i>mysterious</i>	<i>neutral</i>	<i>war on reality</i>	<i>neutral</i>
<i>fake outrage</i>	<i>negative</i>	<i>mystery</i>	<i>neutral</i>	<i>waste</i>	<i>negative</i>
<i>false</i>	<i>negative</i>	<i>mystery book</i>	<i>neutral</i>	<i>waste of a good tree</i>	<i>negative</i>
<i>false advertising</i>	<i>negative</i>	<i>mystery series</i>	<i>neutral</i>	<i>waste of money</i>	<i>negative</i>
<i>false information</i>	<i>negative</i>	<i>mystery suspense thriller</i>	<i>neutral</i>	<i>waste of paper</i>	<i>negative</i>
<i>families</i>	<i>neutral</i>	<i>mystery thriller</i>	<i>neutral</i>	<i>waste of time and money</i>	<i>negative</i>
<i>family</i>	<i>neutral</i>	<i>naive</i>	<i>neutral</i>	<i>waste of trees</i>	<i>negative</i>
<i>family relationships</i>	<i>neutral</i>	<i>nano</i>	<i>neutral</i>	<i>weak love triangle</i>	<i>neutral</i>
<i>fantasy</i>	<i>neutral</i>	<i>narcissist in the white house</i>	<i>neutral</i>	<i>weapons</i>	<i>neutral</i>
<i>fantasy adventure</i>	<i>neutral</i>	<i>narcissistic personality disorder</i>	<i>neutral</i>	<i>well plotted</i>	<i>positive</i>
<i>fantasy series</i>	<i>neutral</i>	<i>natural born citizen</i>	<i>neutral</i>	<i>well written book</i>	<i>positive</i>
<i>fascist government</i>	<i>neutral</i>	<i>nero wolfe</i>	<i>neutral</i>	<i>what are facts</i>	<i>neutral</i>
<i>fatal encryption</i>	<i>neutral</i>	<i>new</i>	<i>neutral</i>	<i>white</i>	<i>neutral</i>
<i>father and son</i>	<i>neutral</i>	<i>new age</i>	<i>neutral</i>	<i>white earphones</i>	<i>neutral</i>
<i>father and sons</i>	<i>neutral</i>	<i>new york times bestselling author</i>	<i>positive</i>	<i>william shatner</i>	<i>neutral</i>
<i>fatherhood</i>	<i>neutral</i>	<i>nice</i>	<i>positive</i>	<i>wimpy</i>	<i>neutral</i>
<i>fathers and sons</i>	<i>neutral</i>	<i>nicely build</i>	<i>positive</i>	<i>wimpy hero</i>	<i>neutral</i>
<i>fathers day</i>	<i>neutral</i>	<i>no credibility</i>	<i>negative</i>	<i>wimpy kid</i>	<i>neutral</i>
<i>father-son relationship</i>	<i>neutral</i>	<i>no good</i>	<i>negative</i>	<i>wimpy kid series</i>	<i>neutral</i>
<i>faux news puppet</i>	<i>negative</i>	<i>no justice</i>	<i>neutral</i>	<i>wisdom</i>	<i>neutral</i>
<i>favorite</i>	<i>positive</i>	<i>no that funny</i>	<i>negative</i>	<i>women sleuths</i>	<i>neutral</i>
<i>favorite books</i>	<i>positive</i>	<i>noise</i>	<i>negative</i>	<i>women warriors</i>	<i>neutral</i>
<i>fear</i>	<i>neutral</i>	<i>noise attenuation</i>	<i>neutral</i>	<i>womens role</i>	<i>neutral</i>
<i>fear-mongering</i>	<i>negative</i>	<i>noise canceling</i>	<i>neutral</i>	<i>wonderful</i>	<i>positive</i>
<i>female protagonist</i>	<i>neutral</i>	<i>noise cancelling</i>	<i>neutral</i>	<i>wonderful book</i>	<i>positive</i>

<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>	<i>Tag</i>	<i>ClassificationTag</i>
<i>feminist</i>	<i>neutral</i>	<i>noise isolating</i>	<i>neutral</i>	<i>wondrous</i>	<i>positive</i>
<i>fiction</i>	<i>neutral</i>	<i>noise reducing</i>	<i>neutral</i>	<i>worst</i>	<i>negative</i>
<i>fiction novel</i>	<i>neutral</i>	<i>noise reduction</i>	<i>neutral</i>	<i>worst person of the year</i>	<i>negative</i>
<i>fiction read in 2010</i>	<i>positive</i>	<i>noise-canceling</i>	<i>neutral</i>	<i>worst president ever</i>	<i>neutral</i>
<i>fictional novel</i>	<i>neutral</i>	<i>noise-canceling headphones</i>	<i>neutral</i>	<i>worthless</i>	<i>negative</i>
<i>fighting</i>	<i>neutral</i>	<i>noise-isolating</i>	<i>neutral</i>	<i>write one leaf</i>	<i>neutral</i>
<i>fire</i>	<i>neutral</i>	<i>noise-reducing headphones</i>	<i>neutral</i>	<i>wrong</i>	<i>negative</i>
<i>fit</i>	<i>positive</i>	<i>nonsense</i>	<i>negative</i>	<i>ya fiction</i>	<i>neutral</i>
<i>flat responce</i>	<i>undefined</i>	<i>not durable</i>	<i>negative</i>	<i>ya issues</i>	<i>neutral</i>
<i>fleishers</i>	<i>neutral</i>	<i>not that great</i>	<i>negative</i>	<i>ya sci-fi</i>	<i>neutral</i>
<i>foam bud</i>	<i>neutral</i>	<i>not worth it</i>	<i>negative</i>	<i>ya survival</i>	<i>neutral</i>
<i>foley work</i>	<i>undefined</i>	<i>novel</i>	<i>neutral</i>	<i>ya survival fiction</i>	<i>neutral</i>
<i>for a good laugh and a good cry</i>	<i>positive</i>	<i>novels</i>	<i>neutral</i>	<i>young adult</i>	<i>neutral</i>
<i>forbidden love</i>	<i>neutral</i>	<i>novice</i>	<i>neutral</i>	<i>young adult action</i>	<i>neutral</i>
<i>forensic</i>	<i>neutral</i>	<i>obama hater</i>	<i>neutral</i>	<i>young girls</i>	<i>neutral</i>
<i>forensic mystery</i>	<i>neutral</i>	<i>obama in wonderland</i>	<i>neutral</i>	<i>young-adult</i>	<i>neutral</i>
<i>forensic thriller</i>	<i>neutral</i>	<i>obama is awesome</i>	<i>neutral</i>	<i>youth politics</i>	<i>neutral</i>
<i>forensics</i>	<i>neutral</i>	<i>obama liberty limbaugh truth conservative</i>	<i>neutral</i>	<i>zen</i>	<i>positive</i>
<i>forget this one</i>	<i>negative</i>	<i>obama lies</i>	<i>negative</i>	<i>zune accesories</i>	<i>neutral</i>
<i>fox news</i>	<i>neutral</i>	<i>obamanation</i>	<i>neutral</i>		