



# Desenvolvimento de um sistema de recomendação de planos nutricionais e de atividade física para doentes diabéticos do tipo 2

NAIR GOMES D'ARAÚJO

novembro de 2022

# **Desenvolvimento de um sistema de recomendação de planos nutricionais e de atividade física para doentes diabéticos do tipo 2**

**Nair Gomes d'Araújo**

Engenheira Biomédica pelo Instituto Superior de Engenharia do Porto

Dissertação apresentada no Instituto Superior de Engenharia do Porto para a obtenção de grau de Mestre em Engenharia Biomédica

Orientador: Luís Conceição e Diogo Martinho

**Novembro 2022**



## **Agradecimentos**

Gostaria de aproveitar esta secção para dedicar os meus agradecimentos às pessoas que me ajudaram durante todo o processo de pesquisa, desenvolvimento e escrita do presente documento.

Em primeiro lugar gostaria de agradecer aos engenheiros Diogo Martinho e Luís Conceição orientadores da presente dissertação. Por todo o apoio, orientação e paciência que me deram durante todo este processo, principalmente durante a finalização do presente documento, um sincero obrigada!

Um agradecimento especial à minha família, que sempre me deram todo o apoio e ajuda principalmente durante os últimos meses no desenvolvimento da tese. E a todos os colegas que me acompanharam durante o meu percurso académico sem vocês este percurso não seria o mesmo, muito obrigada.



## Resumo

Em 2019, o estudo realizado pela APAH, revelou que 66 % dos portugueses levavam um estilo de vida sedentário sem prática regular de atividade física, para além disso, mais de metade da população apresenta excesso de peso. O mesmo estudo colocou, em 2021, Portugal como segundo país da União Europeia com maior prevalência de diabetes. A diabetes do tipo II representa 90 % dos casos diagnosticados. É uma doença crónica cujo desenvolvimento pode levar semanas ou meses e os sintomas podem passar despercebidos. Estas características tornam-na uma doença diagnosticada de forma tardia, o que aumenta o risco de desenvolvimento de comorbilidades como a perda de visão. O principal meio de prevenção da doença é a adoção de estilos de vida ativos e saudáveis, sendo crucial promover mecanismos de incentivo a tal na população portuguesa.

Tecnologias como os sistemas de recomendação, são uma das ferramentas utilizadas na promoção de novos hábitos, pois são capazes de recolher de forma eficiente grandes quantidades de informação acerca de um dado domínio e condensá-la antes de serem apresentada ao utilizador. Ao utilizar informações acerca de alimentos, exercícios físicos, tipos de dietas e regimes fitness são capazes de apresentar sugestões que vão ao encontro das preferências do utilizador e o direcionam a alcançar o objetivo desejado, neste caso um estilo de vida ativo e uma alimentação saudável e equilibrada.

Como tal é proposto um sistema de recomendação de receitas utilizando a recomendação baseada em conteúdo e por filtragem colaborativa para sugerir receitas que vão ao encontro dos gostos do utilizador. Para o desenvolvimento e teste do sistema foi utilizado um *dataset* com informações reais relativas a 14.099 receitas e 1.000 utilizadores e respetivas interações entre ambos. Através dos casos de estudo desenvolvidos, percebeu-se que a recomendação baseada em conteúdo apresenta limitações na identificação das preferências dos utilizadores, sendo estas colmatadas pela filtragem colaborativa.

### **Palavras-chave:**

Sistema de recomendação nutricional, diabetes tipo 2, sistema de recomendação baseado em conteúdo, sistema de recomendação por filtragem colaborativa



## **Abstract**

In 2019, the study carried out by APAH revealed that 66% of Portuguese people led a sedentary lifestyle without regular practice of physical activity, in addition, more than half of the population is overweight. The same study ranked Portugal as the second country in the European Union with the highest prevalence of diabetes in 2021. Type II diabetes accounts for 90% of diagnosed cases. It is a chronic disease whose development can take weeks or months and the symptoms can go unnoticed. These characteristics make it a late diagnosis disease, which increases the risk of developing comorbidities such as vision loss. The main means of disease prevention is the adoption of active and healthy lifestyles, and it is crucial to promote mechanisms to encourage this in the Portuguese population.

Technologies such as recommendation systems are one of the tools used to promote new habits, as they are capable of efficiently collecting large amounts of information about a given domain and condensing it before being presented to the user. By using information about food, physical exercises, types of diets and fitness regimes, they can present suggestions that meet the user's preferences and direct them to achieve the desired goal, in this case an active lifestyle and a healthy and balanced diet.

As such, a recipe recommendation system is proposed using content-based recommendation and collaborative filtering to suggest recipes that meet the user's tastes. For the development and testing of the system, a dataset with information regarding 14099 recipes, 1000 users, and their respective interactions, was used. Through the developed case studies, it was noticed that the content-based recommendation has limitations in identifying user preferences, which are overcome by collaborative filtering.

### **Key words:**

Nutrition recommendation system, type 2 diabetes, content-based recommendation system, collaborative filtering recommendation system



# Índice

AGRADECIMENTOS.....	III
RESUMO .....	V
ABSTRACT .....	VII
ÍNDICE.....	IX
LISTA DE FIGURAS.....	XI
LISTA DE TABELAS.....	XIII
LISTA DE ABREVIATURAS .....	XV
1. INTRODUÇÃO.....	3
1.1. ENQUADRAMENTO .....	3
1.2. MOTIVAÇÃO .....	4
1.3. OBJETIVOS.....	5
1.4. ORGANIZAÇÃO DO DOCUMENTO .....	5
2. DIABETES MELLITUS.....	9
2.1. GESTÃO E ACOMPANHAMENTO DA DIABETES .....	11
2.2. TERAPIA NUTRICIONAL .....	13
2.3. ATIVIDADE FÍSICA .....	15
3. ESTADO DA ARTE .....	21
3.1. ABORDAGENS UTILIZADAS NOS SISTEMAS DE RECOMENDAÇÃO .....	21
3.1.1. SISTEMAS DE RECOMENDAÇÃO BASEADO EM CONTEÚDO .....	22
3.1.2. SISTEMA DE RECOMENDAÇÃO POR FILTRAGEM COLABORATIVA.....	23
3.1.3. SISTEMA DE RECOMENDAÇÃO BASEADO NA DEMOGRAFIA .....	24
3.1.4. SISTEMA DE RECOMENDAÇÃO BASEADO EM CONHECIMENTO.....	25
3.1.4.1. SISTEMA BASEADO EM CASO.....	26
3.1.4.2. SISTEMA BASEADO EM RESTRIÇÕES .....	26
3.1.5. SISTEMA DE RECOMENDAÇÃO HÍBRIDO.....	27
3.2. SISTEMAS DE RECOMENDAÇÃO PARA A DIABETES .....	28
3.2.1. SISTEMAS DE RECOMENDAÇÃO NUTRICIONAL .....	29
3.2.1.1. SISTEMA DE RECOMENDAÇÃO DE SNACKS DO IRANIANO .....	29
3.2.1.2. FOOD RECOMMENDATION SYSTEM USING CLUSTERING ANALYSIS FOR DIABETIC PATIENTS .....	29
3.2.2. SISTEMAS DE RECOMENDAÇÃO DE EXERCÍCIO FÍSICO E DIETA.....	30
3.2.2.1. SISTEMA DE RECOMENDAÇÃO HÍBRIDO PARA SUGESTÃO DE EXERCÍCIOS E DIETAS.....	30
3.2.2.2. SHADE .....	31

3.2.2.3.	SISTEMA DE RECOMENDAÇÃO PERSONALIZADO DE DIETA E EXERCÍCIO PARA A DIABETES TIPO I .....	31
3.3.	DISCUSSÃO: .....	32
4.	PROPOSTA DE SOLUÇÃO.....	39
4.1.	TECNOLOGIAS.....	39
4.2.	DATASET .....	39
4.3.	MODELO DE DOMÍNIO.....	40
4.4.	ARQUITETURA .....	40
4.5.	IMPLEMENTAÇÃO DO SISTEMA DE RECOMENDAÇÃO .....	42
4.5.1.	<i>SISTEMA DE RECOMENDAÇÃO BASEADO EM CONTEÚDO:</i> .....	42
4.5.2.	<i>SISTEMA DE RECOMENDAÇÃO POR FILTRAGEM COLABORATIVA</i> .....	45
5.	CASOS DE ESTUDO E RESULTADOS .....	51
5.1.	CASOS DE ESTUDO .....	52
5.2.	RESULTADOS .....	54
6.	CONCLUSÕES .....	63
6.1.	SÍNTESE E CONCLUSÃO .....	63
6.2.	TRABALHOS FUTUROS .....	65
	REFERÊNCIAS .....	67

## Lista de Figuras

<b>Figura 4.1</b> – Modelo de Domínio .....	40
<b>Figura 4.2</b> – Criação da matriz A. ....	43
<b>Figura 4.3</b> – Criação da matriz B.....	43
<b>Figura 4.4</b> – Criação da matriz peso. ....	44
<b>Figura 4.5</b> – Cosine similarity. ....	44
<b>Figura 4.6</b> – Calculo da similaridade entre receitas.....	45
<b>Figura 4.7</b> – Criação da matriz M.....	46
<b>Figura 4.8</b> – Previsão de classificação.....	47
<b>Figura 5.1</b> – Média de classificações por utilizador .....	51



## Lista de Tabelas

<b>Tabela 2.1</b> – Objetivos para a gestão da DM tipo II. ....	12
<b>Tabela 2.2</b> – Intervalos do nível glicémico desejável. ....	12
<b>Tabela 2.3</b> – Modalidade de exercício e impacto sobre o indivíduo. ....	17
<b>Tabela 2.4</b> – Modalidade de exercícios e atividades físicas. ....	18
<b>Tabela 3.1</b> – Comparação entre as funcionalidades dos SR destacados. ....	32
<b>Tabela 3.2</b> – Técnicas utilizadas para a concessão dos SR destacados. ....	34
<b>Tabela 4.1</b> – Dados de entrada do utilizador com ID 12. ....	41
<b>Tabela 4.2</b> – Dados de entrada das receitas com Id: 60532, 9369 e 5. ....	41
<b>Tabela 5.1</b> – Resumo dos dados de entrada para os casos de teste. ....	52
<b>Tabela 5.2</b> – Preferências dos utilizadores 39, 309 e 444. ....	54
<b>Tabela 5.3</b> – Recomendações para o active user 39. ....	55
<b>Tabela 5.4</b> – Recomendações para o active user 309. ....	55
<b>Tabela 5.5</b> – Recomendações para o <i>active user</i> 39 utilizando K=10. ....	56
<b>Tabela 5.6</b> – Recomendações para o <i>active user</i> 309 utilizando K=10. ....	56
<b>Tabela 5.7</b> – Recomendações para o <i>active user</i> 444 utilizando K=10. ....	56
<b>Tabela 5.8</b> – Recomendações para o active user 39 utilizando K=20. ....	57
<b>Tabela 5.9</b> – Recomendações para o <i>active user</i> 309 utilizando K=20. ....	57
<b>Tabela 5.10</b> – Recomendações para o <i>active user</i> 444 utilizando K=20. ....	57
<b>Tabela 5.11</b> – Recomendações para o <i>active user</i> 39 utilizando K=30. ....	58
<b>Tabela 5.12</b> – Recomendações para o <i>active user</i> 309 utilizando K=30. ....	58
<b>Tabela 5.13</b> – Recomendações para o <i>active user</i> 444 utilizando K=30. ....	58
<b>Tabela 6.1</b> – Listagem dos objetivos propostos. ....	65



## Lista de Abreviaturas

- ADA – Associação de Diabetes Americana
- APAH – Associação Portuguesa de Administradores Hospitalares
- IMC – Índice de Massa Corporal
- DM – Diabetes Mellitus
- FMUP – Faculdade de Medicina da Universidade do Porto
- ISEP – Instituto Superior de Engenharia do Porto
- KNN – *K-nearest neighbors*
- NI – *Nutrition Informatics*
- SR – Sistema de Recomendação
- SRN – Sistemas de recomendação nutricional
- TF-IDF – *term-frequency inverse document frequency*



## CAPÍTULO 1 – INTRODUÇÃO



## **1. Introdução**

O presente documento apresenta e descreve o desenvolvimento de um algoritmo de recomendação de receitas composto pela abordagens filtragem colaborativa e recomendação baseada em conteúdo. O objetivo do protótipo projetado é perceber e identificar preferências alimentares em pacientes diagnosticados com a Diabetes Mellitus tipo II (DM). Este capítulo encontra-se dividido em 4 seções. Será iniciado pelo enquadramento da DM tipo II, seguido pela motivação prendida ao desenvolvimento do protótipo e respetiva apresentação dos objetivos a que este se propõe a concretizar. O capítulo termina com uma breve revisão dos temas abordados nos capítulos que se seguem.

### **1.1. Enquadramento**

A presente dissertação está enquadrada no FoodFriend, um projeto direcionado para: prevenção da desnutrição de pacientes que requerem alimentação por sonda; cuidado nutricional para doenças crónicas, como a diabetes. O projeto conta com um consórcio de 4 parceiros nacionais: Faculdade de Medicina da Universidade do Porto (FMUP); Instituto Superior de Engenharia do Porto (ISEP), através do GECAD; FYI Digital Innovation; e CLOO Behavioral Insights Unit.

A DM tipo II é descrita como uma patologia onde as células do organismo perdem ou diminuem a capacidade de produzir a hormona insulina, que está relacionado com a regulação de glucose no sangue entre outros macronutrientes, como os lípidos e as proteínas. Quando descontrolada a diabetes pode levar a episódios de hiperglicemia ou hipoglicemia, excesso ou défice de glucose no sangue respetivamente, devendo estes ser evitados. É uma doença crónica controlada através da adoção de hábitos alimentares saudáveis e um estilo de vida ativo, podendo o seu tratamento incluir o uso de medicação, para auxiliar no controlo glicémico, e a utilização de injeções de insulina. Na grande maioria dos casos a diabetes pode ser gerida unicamente através da adoção de um estilo de vida saudável: manter uma boa alimentação, realizar exercício físico e manter um IMC (Índice de Massa Corporal) saudável. No entanto uma má gestão da doença pode levar ao surgimento de complicações como a retinopatia diabética (perda de visão) e nefropatia (insuficiência renal) [1].

Em Portugal, durante 2021 foram registados 2,7 milhões de portugueses com diabetes ou em risco de desenvolver a doença. É o segundo país da União Europeia com maior prevalência. Através de uma análise aos dados de 2019 pode-se perceber que o crescente número de casos prende-se com os elevados níveis de sedentarismo e obesidade, cerca de 66 % da população não pratica qualquer tipo de atividade desportiva diariamente e mais de metade sofre de excesso de peso. Ambos os cenários são fatores de risco ao desenvolvimento da doença. Dentro das complicações geradas pela diabetes as mais usuais são a cegueira, afeta 3.000 portugueses anualmente; amputação, afeta 1.300 portugueses anualmente; doenças cardiovasculares, 1 em cada 3 paciente diabéticos sofre AVC ou EAM; e insuficiência renal crónica, afeta 1 em cada 3 portugueses com diabetes [2].

É, no entanto, importante partilhar que durante 2020 verificou-se uma diminuição de 23 % no número de novos casos com diabetes e a diminuição de 62 % no número de indivíduos diagnosticados com obesidade. Além disso, verificou-se uma diminuição de 15 % na quantidade de casos de diabetes tratados sob contexto hospitalar. Segundo o jornal Público, estes valores podem não ser o espelho de uma realidade em que a incidência de ambas as condições está a diminuir, mas estarem associados ao nível de acesso disponibilizado aos portugueses no que toca a cuidados de saúde primários [3].

## **1.2.Motivação**

Segundo o estudo realizado pela Associação Portuguesa de Administradores Hospitalares (APAH), durante 2020 e 2021, anos fortemente marcados pela pandemia provocada pela Covid-19, 60 % dos utentes com diabetes receberam o apoio através de consultas virtuais. Após uma análise de recetividade a ferramentas de telemedicina onde participaram doentes diabéticos dos 32 países europeus, os doentes portugueses demonstraram-se os mais recetivos. É de realçar que o uso de tecnologias de telemedicina na gestão de doente com DM demonstrou ter um impacto significativo no controlo glicémico levando a uma redução de 0,5% do HbA1c [2].

Para além de consultas virtuais existem outras formas de aliar a tecnologia ao tratamento da diabetes, nomeadamente *nutrition informatics* (NI) uma área da informática

direcionada à saúde que conjuga tecnologia da informação, ciências da informação e nutrição. Uma das principais tecnologias aplicadas no âmbito da NI são os sistemas de recomendação nutricional (SRN), estes são explorados de maneira a auxiliarem os utilizadores a alterarem os seus hábitos alimentares para escolhas mais saudáveis [4].

A sua aplicação no tratamento da DM tipo II tem sido enumerada em diferentes artigos direcionados, por exemplo, à recomendação de receitas que respeitam as restrições médicas dos utilizadores. A utilização de SRN é apelativa no tratamento da diabetes, uma vez que esta é uma patologia crónica e a gestão ou recessão da mesma é determinada pelos hábitos alimentares dos indivíduos com a doença. Neste ponto, a bibliografia avaliada até ao momento tem demonstrado a eficiência dos SRN em alterar os hábitos alimentares para opções mais saudáveis, bem como promover a adoção dos mesmos por longos períodos [5].

### **1.3. Objetivos**

Sendo o objetivo final do projeto a concessão de um algoritmo que empregue abordagens de recomendação e a sua testagem em indivíduos, é importante destacar os objetivos do trabalho proposto:

- Desenvolver um algoritmo capaz de analisar e perceber os gostos do utilizador e realizar sugestões de receitas que vão ao encontro dos mesmos.
- Fazer a filtragem de receitas que possam provocar uma má gestão do controlo glicémico, aumento do colesterol, pressão arterial ou do peso do indivíduo.
- Combinar as abordagens selecionadas e analisar os benefícios e desvantagens que os algoritmos híbridos trazem à recomendação de receitas.

### **1.4. Organização do documento**

O presente documento encontra-se dividido em 6 capítulos. É iniciado pelo capítulo 1, uma introdução ao projeto onde é feito um levantamento sobre o contexto da

diabetes em Portugal. Seguido da apresentação dos motivos que se prendem à concessão do projetos bem como os objetivos propostos.

Segue-se o capítulo 2, onde é feita uma análise bibliográfica referente aos vários tipos de diabetes com foco na DM tipo II. Compreensão dos modos de tratamento e qual o papel desempenhado pela alimentação e pratica regular de exercício físico sobre o tratamento da doença.

No capítulo 3 encontra-se uma síntese referente aos sistema de recomendação, quais as abordagens que existem para a sua concessão, sob que princípio atuam e quais o algoritmos empregues. No final é feito um levantamento de sistemas desenvolvidos para a gestão da DM tipo II direcionados para recomendação de planos nutricionais e exercícios físicos. O capítulo termina com a comparação entre as funcionalidades presentes nos diferentes sistemas mencionados.

No capítulo 4 é feita a apresentação da proposta de solução. É definido que abordagens de recomendação serão empregues, qual o método utilizado para a concessão do protótipo e quais as funcionalidades que o protótipo deverá desempenhar.

No capítulo 5 o protótipo desenvolvido é testado através da concessão de casos de estudo, a descrição dos mesmos e os resultados alcançados serão apresentados neste capítulo. Através dos resultados, é feita uma análise quanto às vantagens e desvantagens que cada abordagem apresenta quando empregue individualmente.

A dissertação é concluída no capítulo 6, onde são apresentadas as conclusões retiradas durante a concessão do mesmo bem como a apresentação de uma proposta para melhorar o trabalho feito até a data.

## CAPÍTULO 2 – DIABETES MELLITUS



## 2. Diabetes Mellitus

A insulina é uma hormona segregada pelas células beta do pâncreas, a sua função é não só transportar glucose para as células, como também outros nutrientes com gordura e proteínas. Após as refeições, a quantidade de glucose presente no sangue aumenta o que faz com que o pâncreas aumente a produção de insulina. Esta viaja pela corrente sanguínea e sinaliza às células que podem começar a absorver a substância. O que resulta numa progressiva diminuição dos níveis de glucose no sangue [6].

Grandes quantidades de açúcar no sangue, levam a consequências negativas para o organismo, caso esta situação se perpetue e não seja tratada pode levar à morte. Existem, no entanto, condições clínicas em que as células deixam de responder corretamente à hormona, este cenário é denominado por resistência à insulina. Como resposta aos elevados níveis de glucose no sangue, o pâncreas segrega cada vez mais insulina para a corrente sanguínea, cenário denominado por hiperinsulinemia, se estes valores continuarem altos, com o tempo, induz nas células resistência à hormona. Resultado, tanto os níveis de glucose como de insulina aumentam. Se esta condição se prolongar, levará à danificação do pâncreas e à diminuição de produção de insulina. A DM tipo II é diagnosticada quando os níveis de glucose ultrapassam um dado limite [7].

No caso da DM tipo II as principais fatores que levam à resistência de insulina estão relacionados com o estilo de vida, obesidade, hábitos alimentares pouco saudáveis, sedentarismo e idade, no entanto a diabetes tem diferentes formas de se manifestar e com elas diferentes causas.

Na DM tipo I, as suas causas são ainda desconhecidas, no entanto, a literatura avaliada até ao momento atribui parte das responsabilidades à condição genética do individuo. Neste cenário o sistema imunitário ataca e destrói as células responsáveis pela produção da insulina no pâncreas, levando a que este não seja capaz de a produzir. Nestes casos os indivíduos estão dependentes de doses diárias de insulina para repor os níveis de glucose no sangue [8] [9].

A diabetes pode ainda desenvolver-se durante a gravidez, esta condição é denominada por diabetes gestacional, e usualmente, após a gravidez, os níveis de glucose no sangue voltam ao normal. A diabetes gestacional pode surgir em qualquer estagio da gravidez, mas é mais comum no segundo e terceiro trimestre. Durante a gravidez pode acontecer que o pâncreas não consegue produzir a quantidade de insulina extra que é

necessária, o que pode causar problemas tanto à mãe como ao bebê, no entanto, os riscos ao seu desenvolvimento são minimizados se esta for detetada cedo e bem gerida [10].

A diabetes pode desenvolver-se por outras causas como drogas ou químicos, por exemplo glucocorticoides utilizados no tratamento da HIV/AIDS ou após transplantes de órgão, ou ainda consequentes de outras doenças [11].

Na DM tipo II, se não tratada os níveis de glucose no sangue aumentam, quando atingem ou ultrapassam os 7 mmol/L depois de uma refeição ou 8,5 mmol/L passadas 2h é denominada hiperglicemia, ou excesso de glucose no sangue. O seu inverso, ou baixa concentração de glucose no sangue, valores abaixo dos 4 mmol/L é denominada por hipoglicemia [12] [13].

O tratamento da DM tipo II é composto por [14]:

- **Medicação:** Existem diferentes medicamentos prescritos para a diabetes que auxiliam no controlo glicémico, sendo o mais comum, e usualmente, o inicialmente prescrito, Metmorfin. As medicações prescritas atuam de diferentes formas, por exemplo como estimulante à produção de insulina, é o caso dos *Sulphonylures*, ou podem auxiliar na perda de peso.
- **Cirurgias para a perda de peso:** A bibliografia analisada até ao momento aponta que a diminuição do peso tem impacto significativo no controlo glicémico. Por isso, pode ser recomendado em certos casos, procedimentos cirúrgicos direcionados ao estômago ou intestinos que auxiliem à perda de peso.
- **Dieta e exercício:** A grande maioria dos indivíduos diagnosticados com DM tipo II conseguem fazer uma boa gestão da doença através da adoção de hábitos de vida mais saudáveis, como priorizar uma dieta saudável composta por alimentos ricos em nutrientes e combinar com um estilo de vida mais ativo que incorpora mais exercício físico no quotidiano
- **Insulina:** Embora nem todos os casos utilizem insulina regularmente, pode ser necessário em alguma etapa da gestão da diabetes conjugar o tratamento com injeções de insulina. Por exemplo, existem casos em que os pacientes quando são inicialmente diagnosticados apresentam elevados valores de glucose no sangue, como medida a curto prazo, pode ser prescrita injeções de insulina para auxiliar no tratamento. Outras razões

podem estar ligadas à medicação, por exemplo, o facto desta não ser suficiente para controlar o índice glicémico ou não ser apropriada ao paciente.

- Suporte emocional: A diabetes não afeta só fisicamente os pacientes também atua negativamente no estado emocional deles. Quer seja no início ou após um longo período a lidar com a doença esta pode levar a stress, depressão ou esgotamento. É importante fazer um acompanhamento psicológico quer por parte de profissionais de saúde, ou simplesmente ter acesso ao apoio psicológico necessário através da família ou amigos.

A diabetes está entre uma das doenças crônicas com crescimento mais rápido no mundo e está associada a complicações como a perda de visão, doenças renais, problemas cardíacos, problemas ósseos e danos nos nervos. Mesmo assim, a taxa de aumento de pacientes com diabetes mundialmente é de 70.000 por ano. O tratamento e controlo das complicações envolve uma dieta saudável, prática regular de exercício, controlo do peso e testagem sanguínea regular para monitorizar os níveis de glucose no sangue. Nos pacientes jovens, é aconselhado o foque numa dieta controlada, prática regular de exercício e evitar, tanto quanto possível, a prescrição de medicamentos [15].

## **2.1. Gestão e acompanhamento da diabetes**

Depois de diagnosticada, e uma vez que a diabetes é uma doença crónica, é necessário gerir os seus efeitos adversos como a hipoglicemia e hiperglicemia e certificar que condições consequentes não se manifestem ou piorem (retinopatia, doenças renais, etc.). A gestão diária da doença fica ao cargo do diabético, sendo este ocasionalmente acompanhado por profissionais de saúde para perceber a evolução da doença [16]. O diabético tem por isso alguns objetivos que deverá cumprir. Na tabela 2.1 estão inumeradas as mensuradas que deveram ser monitorizadas, bem como os valores limites a não ultrapassar [17],[18].

**Tabela 2.1** – Objetivos para a gestão da DM tipo II.

	<b>OBJETIVOS</b>
<i>PRESSÃO SANGUÍNEA</i>	<140/80 mm Hg
<i>COLESTEROL LDL</i>	<100 mg/dL
<i>COLESTEROL HDL</i>	<ul style="list-style-type: none"> <li>▪ &gt; 40 mg/dL para os homens</li> <li>▪ &gt; 50 mg/dL para as mulheres</li> </ul>
<i>NÍVEL DE TRIGLICERÍDEOS</i>	<150 mg/dL
<i>HEMOGLOBINA GLICADA (HBA1C)<sup>1</sup></i>	<7 %
<i>RITMO CARDÍACO</i>	<ul style="list-style-type: none"> <li>▪ 70bpm em repouso;</li> <li>▪ 200bpm máximo para indivíduos com 20 anos.</li> </ul>

A DM tipo II é a consequência da perda de segregação de insulina por parte das células beta e a consequente resistência à mesma. Por consequência, o nível glicémico aumenta. Por isso é necessário que o diabético esteja atento à subida e descida de glucose no sangue que acontece antes, durante e após as refeições. Na tabela 2.2 estão apresentados os intervalos de nível glicémico para indivíduos saudáveis e os níveis glicémicos a respeitar pelos indivíduos com DM tipo II [13].

**Tabela 2.2** – Intervalos do nível glicémico desejável.

	<b>NÍVEL GLICÉMICO PARA INDÍVIDUOS SAUDÁVEIS</b>	<b>NÍVEL GLICÉMICO PARA INDÍVIDUOS COM DM TIPO II</b>
<b>ANTES DO PEQUENO-ALMOÇO (JEJUM)</b>	< 100 mg/dL	70 – 130 mg/dL
<b>ANTES O ALMOÇO, JANTAR OU LANCHE</b>	< 110 mg/dL	70 – 130 mg/dL
<b>2 HORAS APÓS A REFEIÇÃO</b>	< 140 mg/dL	< 180 mg/dL
<b>HORA DE DORMIR</b>	< 120 mg/dL	90- 150 mg/dL
<b>HBA1C</b>	< 5,5 %	< 7 %

Fatores como a alimentação e a prática regular de exercício físico tem um impacto significativo no controlo do nível glicémico, ritmo cardíaco, pressão arterial e as demais

---

<sup>1</sup> O teste de hemoglobina glicada (ou HbA1c) mede a quantidade de glucose no sangue. Até janeiro de 2011, em Portugal, o teste era apenas utilizado para monitorizar o controlo metabólico de pacientes diagnosticados com diabetes. Atualmente, este é também recomendado como método de diagnóstico [60] [61].

mensuradas apresentadas. Nos próximos capítulos será feito um levantamento bibliográfico referente a este tópico.

## **2.2.Terapia nutricional**

A terapia nutricional é um conjunto de diretrizes que ajudam o paciente a adotar escolhas alimentares mais saudáveis, no final, estas escolhas terão um impacto significativo na gestão da diabetes [19].

Para muitos indivíduos, este é um dos pontos mais desafiantes, pois envolve a substituição de hábitos alimentares à muito adquiridos, por outros que não são tão intuitivos [20]. Além de que lhes é exigido algum conhecimento sobre as escolhas alimentares mais acertadas, o que pressupõe que o indivíduo deva ser educado previamente sobre estes assuntos, e também, a contínua opção pelas mesmas, ou seja as diretrizes nutricionais devem ser seguidas ao longo da vida do paciente diabético [21].

Por esta razão, a todo o paciente diabético deve ser oferecido o contínuo aconselhamento nutricional por parte de profissionais da área, quer a partir de diretrizes sobre que alimentos evitar, quer a partir do aconselhamento sobre as porções ideais que cada alimento deve ter nas refeições, ou ainda através do desenvolvimento de planos alimentares. Isto é claro, tendo sempre em atenção às necessidades nutricionais do indivíduo e as limitações impostas pela sua condição médica [22].

É necessário, no entanto notar, que em certos casos a terapia nutricional pode não ser suficiente para a gestão eficaz dos níveis glicémicos, nesses casos é associada medicação [23]. Segundo a Associação Americana de Diabetes (ADA), quando a terapia nutricional é acompanhada por nutricionistas registrados contribui para a diminuição de 0,4 a 2 % do HbA1c em pacientes com DM tipo II [24].

Em suma, a terapia nutricional tem como objetivos ajudar o indivíduo no controlo glicémico, pressão arterial, controlo do LDL e HDL, controlo do peso e foca em retardar ou prevenir complicações da diabetes; ajudar o indivíduo a perceber como pode atender às suas necessidades nutricionais sem comprometer as suas preferências alimentares pessoais e culturais; respeitar o prazer de comer e transmitir valores positivos em torno das escolhas alimentares tendo sempre em atenção que quaisquer restrições alimentares impostas devem ter uma base científica; dar ao indivíduo as ferramentas necessárias para

que este seja responsável pelo seu próprio planeamento de refeições e desconstruir o dogma que muitas vezes os pacientes caem em optar por estratégias que privilegiam o registo constante de calorias, macro e micronutrientes [25].

A literatura analisada até à data, sugere que não existe uma distribuição percentual ideal de macro e micronutrientes e reforça o foco em respeitar as preferências pessoais, culturais e demográficas dos indivíduos com DM tipo II. Além disso é importante ter em conta os objetivos metabólicos, como o nível glicémico e de colesterol, no que toca ao planeamento de refeições [16].

Esta abordagem tem ainda um conjunto de concelhos direcionados ao quotidiano dos indivíduos com DM tipo II:

- No consumo de carboidratos deve existir uma preferência por fontes de carboidratos densas em nutriente e ricas em fibras, é o caso das frutas, vegetais, grãos integrais e laticínios.
- Para indivíduos que intercalam o tratamento da diabetes com insulina, é aconselhada a educação na contagem de carboidratos e em alguns casos, considerar a quantidade de gordura e proteínas para determinar a dosagem de insulina na hora das refeições. Desta forma podem alcançar um melhor controlo glicémico.
- Evitar a ingestão de carboidratos simples, é o caso de bebidas açucaradas e os doces, e a sua adição nos alimentos, controlando assim o nível glicémico, o peso e previne o desenvolvimento de doenças cardiovasculares.
- Privilegiar uma alimentação rica e saudável.
- No que toca à ingestão de proteínas, tem se vindo a notar que existe uma relação entre a ingestão das mesmas e o aumento da resposta da insulina, sem que haja aumento da concentração de glucose no sangue. Por isso, fontes de carboidratos ricas em proteínas devem ser evitadas no tratamento e prevenção da hipoglicemia.
- Ingerir de forma moderada peixe e nozes uma vez que são ricas em ácidos graxos n-3 de cadeia longa e benéficas para a prevenção e tratamento de doenças cardiovasculares. No entanto o consumo regular das mesmas não é aconselhado.

- Evitar o consumo de gorduras trans e à medida que as gorduras saturadas vão progressivamente desaparecendo da dieta é importante que sejam substituídas por gorduras insaturadas e não carboidratos refinados.
- Uma vez que o consumo de álcool está relacionado com o aumento de risco de hipoglicemia, deve ser ingerido moderadamente. Uma dose diária para as mulheres adultas e duas para os homens.
- A quantidade de sódio ingerida diariamente não deve ultrapassar as 2,300 mg.
- Indivíduos que estão habituados à ingestão regular de produtos açucarados é aceitável a substituição dos adoçantes nutritivos (aqueles que tem calorias tais como o açúcar o mel e o xarope de agave) por adoçantes não nutritivos (baixo ou nenhum conteúdo de calorias). O uso de adoçantes não nutritivos não tem impacto significativo no controlo glicémico.
- Para indivíduos que não acompanham o tratamento da diabetes com insulina, uma maneira simples e eficaz para controlar o índice glicémico e o peso é introduzindo o controlo de porções. Ou seja, explicar definindo porções que quantidade de cada alimento deve estar presente nas refeições.

### **2.3. Atividade física**

A perda excessiva de tecido muscular esquelético e o aumento do tecido adiposo é uma das principais alterações que levam ao aumento do risco de desenvolvimento da DM tipo II, este cenário é fortemente predominante na população idosa e um dos fatores que levam a faixa etária a ser atacada pela doença [25]. Para além disso, o músculo esquelético é um dos principais pontos de alocação de glucose pós-prandial e a sua perda está diretamente relacionada com a diminuição da sensibilidade à insulina, efeito precedente ao desenvolvimento da DM tipo II [26].

Para contrariar o desenvolvimento da doença e evitar que esta dê aso ao aparecimento das complicações que lhe estão associadas, é aconselhado tanto à população geral, e principalmente ao indivíduos com diabetes que pratiquem regularmente exercício físico (entende-se por exercício físico toda a pratica desportiva ou sessões de fitness) ou atividade física (ações do quotidiano como caminhar ou subir escadas) [27]. A pratica de exercício físico está associada ao aumento da sensibilidade à insulina, diminuição da

concentração de glucose no sangue e ao aprimoramento da condição cardiovascular dos indivíduos com DM tipo II [28].

A obesidade e o excesso de peso são fatores associados ao desenvolvimento da DM tipo II esta associação deve-se ao acréscimo de tecido adiposo que impacta negativamente a sensibilidade à insulina. Em casos de obesidade ou excesso de peso, a pratica de exercício é fortemente aconselhada e serve um papel não só de contribuição para o aumento da massa muscular, como também para o controlo e diminuição do peso e tecido adiposo [26].

Em suma, a pratica regular de exercício físico resulta não só num melhor controlo glicémico como também minimiza o risco de desenvolvimento de doenças cardiovasculares e auxilia à manutenção ou redução do peso.

A literatura analisada até ao momento demonstra que sem diminuição do IMC sessões regulares de exercício estruturado, com pelo menos 8 semanas de duração em indivíduos com DM tipo II, resultam numa média de diminuição de HbA1c de 0,66 % [16]. Quando associada a uma perda de peso entre 2 e 5 %, a HbA1c sofre uma diminuição entre 0,2 % e 0,3 %. Com perdas de peso superiores, entre os 5 % e os 10 %, é observada uma diminuição entre 0,6 % to 1,0 % de HbA1c [29].

Nos casos em que a perda progressiva de peso ronda os 15 % é possível observar a remissão da patologia bem como o aprimoramento das condições metabólicas numa grande fração da população com DM tipo II [28]. Para além disso, os benefícios no controlo metabólico mantém-se mesmo que o peso seja adquirido passados 4 anos [29].

É recomendado aos indivíduos com DM tipo II que [30]:

- Procurem realizar semanalmente 150 minutos de exercício aeróbico moderado ou vigoroso.
- Realizar 2 a 3 sessões de exercício de resistência para promover o aumento da massa muscular em sessões com duração de meia hora.
- Praticar 2 a 3 vezes por semana exercícios de flexibilidade e equilíbrio, embora existam menos evidencias da sua contribuição para o tratamento da diabetes trazem benefícios na gestão de comorbilidades, por exemplo nas populações mais idosas o aprimoramento do equilíbrio contribui para a prevenção de quedas.

- É importante respeitar o descanso e intercalar as sessões de exercício para que não sejam executadas em dias consecutivos, ou seja dia-sim-dia-não. Este regime otimiza os benefícios induzidos pelo exercício e aumenta a sensibilidade à insulina.
- No que toca à intensidade dos exercícios, esta é uma questão subjetiva ao indivíduo, uma vez que cada um tem o seu nível de aptidão física. Por isso, qualquer regime proposto deve ser personalizado ao mesmo.
- É importante também iniciar a prática de exercício físico com intensidade moderada e à medida que o paciente demonstra ter mais aptidão e resistência aos exercícios, aumentar a carga de esforço necessária.

Na tabela 2.3 é possível observar o impacto que as modalidades de exercício têm sobre o controlo glicémico, diminuição do risco de doenças cardiovasculares, saúde mental, equilíbrio, força muscular e saúde óssea [30].

**Tabela 2.3** – Modalidade de exercício e impacto sobre o indivíduo.

	<b>EXERCÍCIO AERÓBICO</b>	<b>EXERCÍCIO DE RESISTÊNCIA</b>	<b>EXERCÍCIO DE FLEXIBILIDADE</b>	<b>EXERCÍCIO DE EQUILÍBRIO</b>
<b>CONTROLO GLICÉMICO</b>	***	**	—	—
<b>REDUÇÃO DO RISCO CARDIOVASCULAR</b>	***	*	—	—
<b>SAÚDE MENTAL/HUMOR</b>	***	*	—	—
<b>EQUILÍBRIO</b>	*	*	**	***
<b>FORÇA MUSCULAR</b>	*	***	—	*
<b>SAÚDE ÓSSEA</b>	*	***	—	—

No que toca à atividade física é importante encorajar os indivíduos com DM tipo II a aumentar a sua ocorrência no quotidiano, por exemplo favorecer as escadas ao elevador e participar em caminhadas. Na tabela 2.4 estão apresentados exercícios físicos e a modalidade de exercício que incorporam: aeróbico, resistência, flexibilidade ou equilíbrio [30].

**Tabela 2.4** – Modalidade de exercícios e atividades físicas.

	<b>COMPONENTE AERÓBICO</b>	<b>COMPONENTE DE RESISTÊNCIA</b>	<b>COMPONENTE DE FLEXIBILIDADE</b>	<b>COMPONENTE DE EQUILÍBRIO</b>
<b>YOGA</b>	Fraco	Medio	Elevado	Elevado
<b>TREINO DE CIRCUITO</b>	Medio	Medio	Fraco	Fraco
<b>TAI CHI</b>	–	Fraco	Fraco	Medio
<b>ANDAR</b>	Medio	Fraco	–	Fraco
<b>JOGGING/ CORRER</b>	Elevado	Fraco	–	Fraco
<b>PILATES</b>	–	Medio	Fraco	Fraco
<b>ANDAR DE BICICLETA</b>	Elevado	Fraco	–	–
<b>DANÇAR</b>	Medio	–	Fraco	Fraco
<b>JARDINAGEM</b>	Fraco	Fraco	Fraco	Fraco
<b>TREINO COM PESOS</b>	Fraco	Elevado	Fraco	Fraco

## CAPÍTULO 3 – ESTADO DA ARTE



### **3. Estado da Arte**

Sistemas de Recomendação (SR) são tecnologias que preveem e sugerem itens relevantes a utilizadores com base nas suas preferências tentando minimizar os esforços do utilizador na procura de informação [31]. Podemos encontra-los nas mais variadas áreas, seja na recomendação de produtos para compras, musica, pesquisa de web sites, artigos de pesquisa e até planos nutricionais, ponto que será abordado ao longo do documento [32].

Os SR evoluíram como uma área independente de pesquisa em meados da década de 70, na Universidade Duke, Estados Unidos da América. O primeiro sistema apresentado foi Tapestry1, desenvolvido em 1992 pelo centro de investigação Xerox Pablo Alto. A motivação que originou o projeto prende-se ao crescente número de emails recebidos e à necessidade de uma filtragem em função das preferências do utilizador. Para a concessão do projeto foi utilizado um sistema de recomendação por filtragem colaborativa [33], [34].

Desde a sua introdução nos anos 90, cada vez mais SR têm sido desenvolvidos com o propósito de auxiliar a população na seleção de produtos. Atualmente centenas de estudos são lançados com propostas de utilização de SR baseados nas mais diversas abordagens para as mais diversas aplicações, no entanto, todavia existem limitações nestas poderosas tecnologias e por isso a comunidade científica continua a desenvolver novas e mais arrojadas soluções [34], [35].

De um modo geral, podemos admitir que o processo de recomendação de itens está dividido em duas etapas. A primeira, prever a classificação e a adequação de itens desconhecidos pelo utilizador, é baseada no histórico de interesses demonstrado pelo utilizador. A segunda, fornecer os principais itens de interesse sob a forma de recomendações. O sucesso de ambas as etapas é assegurado pela escolha correta da abordagem pela qual é definido o SR [36], [37]

#### **3.1. Abordagens utilizadas nos sistemas de recomendação**

Nesta secção será feito o levantamento bibliográfico referente aos SR: baseado em conteúdo, filtragem colaborativa, baseado em demografia e baseado em conhecimento. As diferentes abordagens são descritas em subsecções em que para além

da descrição do método sob a qual atuam, também são descritos alguns dos algoritmos empregues na sua concessão.

### 3.1.1. Sistemas de recomendação baseado em conteúdo

O sistema de recomendação baseado em conteúdo sugere itens ao utilizador com base nas características do item e no perfil de interesses do utilizador. Isto é, um SR de filmes (item), que utilize esta abordagem, recomendaria filmes cujos géneros cinematográficos (característica) fossem apreciados pelo utilizador.

A abordagem segue a seguinte metodologia, através de um set de interações entre o utilizador e um dado item, o algoritmo recolhe as características dos itens e por via de cálculos atribuí pesos a cada uma. Pesos mais elevados estão associados a um maior interesse por parte do utilizador. O sistema faz corresponder este set de características com a dos demais itens, e os itens que sejam compostos por características com pesos mais elevados, serão mais prováveis de serem recomendados ao utilizador. Em suma, a principal tarefa da recomendação baseada em conteúdo é fazer corresponder as características que interessam ao utilizador com as características dos itens [38], [39], [40].

Alguns dos métodos utilizados no desenvolvimento são o método *Bayen, cosine similarity* (Eq. 3.1) e Redes Neurais Artificiais [33], [43].

$$sim(i, j) = \frac{r_i r_j}{\|r_i\| \|r_j\|} \quad (\text{Eq. 3.1})$$

Seja  $i$  e  $j$  dois vetores,  $r_i r_j$  representa a soma da multiplicação dos elementos de cada vetor e  $\|r_i\|, \|r_j\|$  representam a magnitude do vetor.

Esta abordagem descarta a necessidade de relacionar os interesses do utilizador com os de uma vizinhança, característica que os torna vantajosos em cenários onde o perfil do utilizador é único e não coincide com os interesses dos demais e onde a capacidade computacional é limitada. No entanto as recomendações baseadas em conteúdo dependem da qualidade e complexidade da descrição do item [43].

### 3.1.2. Sistema de recomendação por filtragem colaborativa

O sistema de recomendação por filtragem colaborativa utiliza as opiniões de utilizadores vizinhos para prever escolhas de um determinado utilizador. A técnica atua no sentido de recolher informações comportamentais sobre o utilizador, como classificações atribuídas a itens, feedback, preferências e atividades.

Seguindo o exemplo do SR de filmes, os filmes classificados pelo utilizado (ou *active user*) são comparados com os dos restantes utilizadores, com base nisto é calculada a similaridade entre as classificações do *active user* e os dos restantes utilizadores. Os  $n$  utilizadores com maior similaridade são agrupados, denominamo-los por vizinhança. Todos os itens classificados pela vizinhança são inseridos numa lista, da qual são removidos quaisquer itens classificados pelo *active user*, os itens que compõem esta lista podem ou não ser recomendados ao *active user*, sendo o fator de serialização a previsão de classificação. Esta é calculada com base nas classificações atribuídas pela vizinhança. Por fim são seleccionados os itens com previsões mais elevadas e apresentados ao *active user* [44].

A diferença entre esta abordagem e a anterior é que as preferências do utilizador são determinadas pela sua vizinhança, não pelo gosto de uma dada característica do item [38].

Uma vez que a filtragem colaborativa se baseia na determinação da similaridade entre utilizadores e na previsão de classificação de itens, é usual empregar algoritmos como *Pearson correlation* (Eq. 3.2), *Cosine similarity measure* (Eq. 3.1) e *K-nearest neighbors* (Eq. 3.3) [43].

$$\text{sim}(a, b) = \frac{\sum_{p \in P} (r_{a,p} - \bar{r}_\alpha)(r_{b,p} - \bar{r}_\beta)}{\sqrt{\sum_{p \in P} (r_{a,p} - \bar{r}_\alpha)^2 \sum_{p \in P} (r_{b,p} - \bar{r}_\beta)^2}} \quad (\text{Eq. 3.2})$$

Onde  $a, b$  são utilizadores,  $r_{a,p}$  é a classificação atribuída pelo utilizador  $a$  ao item  $p$ ,  $\bar{r}_\alpha$  e  $\bar{r}_\beta$  são a média de classificações e  $P$  representa um conjunto de produtos classificados pelo utilizador  $a$  e  $b$ .

$$\sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (\text{Eq. 3.3})$$

A Eq. 3.3 apresenta o calculo da distancia euclidiana entre os vetores  $x$  e  $y$ .

Uma das grandes limitações enfrentadas pelo sistema, é a escalabilidade. Conforme o aumento de utilizadores e de itens é imposto ao sistema o aumento dos

recursos devido à necessidade de determinar similaridade entre todos os utilizadores do sistema [43].

### 3.1.3. Sistema de recomendação baseado na demografia

O sistema de recomendação baseado na demográfica recomenda itens com base no perfil demográfico dos utilizadores. Os três atributos mais utilizados são: idade, sexo e ocupação, outros atributos incluem: língua materna e país. A recolha dos dados demográficos costuma ser efetuada no registo do sistema, sendo usual a utilização de formulários neste processo. Após a recolha de dados é solicitado ao utilizador que classifique um conjunto de itens. O formulário de registo e as classificações formam o perfil demográfico do utilizador. A próxima etapa assenta no calculo da semelhança entre o perfil demográfico do *active user* e o dos restantes utilizadores do sistema, utilizando métodos como *cosine similarity* (Eq. 3.1). É definida uma vizinhança, isto é, um conjunto de  $n$  utilizadores com perfis demográficos mais semelhantes ao do *active user*. São recolhidos os itens classificados pela vizinhança, quaisquer itens classificados pelo *active user* são removidos e prevê-se a classificação que o *active user* atribuirá a estes (Eq. 3.4 e 3.5). Por fim recomendam-se os itens com previsões mais altas [45] [42].

$$pr_{x,k} = \frac{\sum_{u_y \in N_x} sim(u_x, u_y) \times r_{y,k}}{\sum_{u_y \in N_x} |sim(u_x, u_y)|} \quad (\text{Eq. 3.4})$$

Sendo  $pr_{x,k}$  a previsão de classificação que o *active user*,  $u_x$ , atribuirá ao item  $k$ .  $u_y$ , é um dado utilizador da vizinhança  $N_x$ .  $r_{y,k}$  é a previsão que o  $u_y$  atribuiu ao item  $k$ . E  $sim(u_x, u_y)$  é a similaridade entre os dois utilizadores.

$$pr_{x,k} = \frac{1}{N} \sum_{c' \in \hat{C}} r_{c',k} \quad (\text{Eq. 3.5})$$

Sendo  $\hat{C}$  o conjunto de utilizadores da vizinhança que classificou o item  $k$ .  $N$  é o tamanho do conjunto  $\hat{C}$ , e  $c'$  um dado utilizador do conjunto.  $r_{c',k}$  é a classificação do utilizador  $c'$  ao item  $k$ .

A descrição do funcionamento do sistema deixa claro que a abordagem demográfica utiliza uma metodologia bastante semelhante à filtragem colaborativa, sendo a principal diferença os atributos utilizados no calculo da similaridade entre utilizadores. No caso da filtragem colaborativa os utilizadores são comparados em termos de classificações, já a abordagem demográfica utiliza o perfil demográfico.

Uma vez que esta abordagem não necessita de uma lista de classificações do *active user*, ponto crucial na recomendação utilizando abordagens como recomendação baseada em conteúdo ou pela filtragem colaborativa, são uteis para colmatar o problema do “novo utilizador”, isto é, um utilizador que acaba de chegar ao sistema sem ter quaisquer classificações prévias [45].

A principal desvantagem da abordagem é o uso de dados demográficos. Uma vez que são dados de natureza sensível estão fortemente ligados a questões de segurança. Para além disso, é observado que os um conjunto substancial de utilizadores não se sente confortável em partilhar informações demográficas, devido ao seu caráter sensível, e necessitam de um forte sentido de segurança por parte do sistema para as partilharem. Para além disso, é reportado por alguns utilizadores que a partilha de dados demográficos poem em causa a anonimidade do sistema [45].

#### **3.1.4. Sistema de recomendação baseado em conhecimento**

O conceito geral do sistema de recomendação baseado em conhecimento assenta na recomendação de itens cujas características vão ao encontro das especificações do utilizador [46].

Imaginemos, um dado utilizador pretende comprar uma máquina fotográfica. Para isso vai ao catálogo online de uma dada empresa e procura modelos de máquinas que vão ao encontro dos seus requisitos, como o preço, tipo de lentes, etc. Como retorno, serão apresentadas máquinas fotográficas que vão ao encontro das características indicadas pelo utilizador. Este sistema não depende de classificações para recomendar itens [47], [48]. Existem, no entanto, casos onde o domínio é complexo, o que torna o processo de especificação dos requisitos mais confuso para os utilizados, isto porque podem não ser completamente capazes de enumerar todas os atributos de interesse. Mas podem ser simplificados através de um ciclo de feedback onde o utilizador expõem a sua opinião quanto às recomendações geradas quer estas sejam positivas ou não. Este é um processo iterativo que promoverá a eficiência do sistema [46].

Esta abordagem é muito popular na recomendação de dietas para paciente diabéticos e oncológicos. Sendo as condições médicas mais sensíveis, é exigido ao sistema que cumpra rigorosos requisitos de forma a ir ao encontro das necessidades dos utilizadores [49].

Existem dois tipos de sistemas baseados em conhecimento [34]:

#### **3.1.4.1. Sistema baseado em caso**

O sistema baseado em caso procura adaptar soluções anteriormente utilizadas na resolução de problemas e emprega-as na resolução de problemas novos e semelhantes. É uma técnica onde o algoritmo procura casos antigos resolvidos e semelhantes ao problema presente e adapta as soluções anteriores levando em consideração as diferenças entre ambos [49], [50].

Nesta abordagem o utilizador começa por definir um caso ou problema para a qual precisa de uma recomendação. Este é comparado através de medições de similaridade com os restantes itens do sistema e os que partilharem maior similaridade serão recomendados ao utilizador. É importante utilizar medições de similaridade adequadas ao domínio do sistema. É usual utilizar os itens recomendados como novos casos, isto é utilizar os seus atributos para criar novos casos e registra-los no sistema [46].

#### **3.1.4.2. Sistema baseado em restrições**

Em sistemas baseados em restrições, o utilizador especifica os requisitos para as quais um produto ou item é lhe recomendado, recebendo uma justificação para a recomendação. Para tal, utiliza um *dataset* predefinido de conhecimento com regras específicas para relacionar os requisitos do utilizador com as características dos itens [49]. Os requisitos ou restrições podem ser direcionados ao item, por exemplo “carros com matrícula anterior a 1970, mas combustível tem de ser gasolina”; ou ainda relacionar informações do utilizador com características do item, por exemplo “Investidores mais velhos não investem em itens de alto risco” [46].

O utilizador tem ainda oportunidade de atualização os requisitos nos casos em que poucos ou nenhuns itens possam ser recomendados, ou o contrário, aumentar o número de restrições no casos das recomendações serem excessivas [46].

### 3.1.5. Sistema de recomendação híbrido

A principal vantagem em combinar uma ou mais abordagens é o facto destes sistemas permitirem combinar duas ou mais abordagens e superar as limitações e desvantagens que a técnica, no singular, de outra forma não conseguiria. A este facto, acrescenta-se ainda que a qualidade das recomendações do sistema é melhorada [49].

Dos métodos existentes para combinar um ou mais tipos de abordagens de recomendação foram seleccionados sete: *Weighted*, *Mixed*, *Switching*, *Feature Combination*, *Augmented Features*, *Cascade*, *Meta-level* [51] [52]:

- *Weighted*: as recomendações providenciadas por cada um dos sistemas são recolhidas. A cada sistema é atribuído um peso estático que é multiplicado à previsão de classificação do respetivo sistema. Para cada recomendação é somada a multiplicação ente a previsão de classificação e o peso do sistema. Este novo valor será denominado previsão de classificação combinada. O sistema irá apresentar por ordem decrescente de previsão de classificação combinada as recomendações
- *Mixed*: As recomendações das diferentes abordagens são apresentadas lado a lado ao utilizador.
- *Switching*: Cada uma das abordagens de recomendação irá recomendar um conjunto de itens. Segundo uma ordem pré-definida é determinado se as classificações vão ao encontro das preferências do utilizador, ou não. Caso não coincidam, utilizam-se as recomendações da abordagem seguinte e verifica-se se estas correspondem ou não às preferências. Este processo é repetido até uma das abordagens corresponder às preferências.
- *Feature Combination*: São injetadas características de uma abordagem (exemplo: filtragem colaborativa) num algoritmo desenhado para recomendar itens segundo outra abordagem (exemplo: recomendação baseada em conteúdo).
- *Augmented Features*: Este método emprega dois tipos de SR: um que contribui com dados e o principal. O primeiro gerará previsões de classificações para itens ou perfil do utilizador, que serão adicionadas ao input do SR principal. Este utiliza as informações para gerar recomendações finais.

- *Cascade*: Este método utiliza uma estrutura hierárquica de algoritmos onde as recomendações do sistema secundário não podem anular as recomendações geradas pelo sistema principal. São, no entanto, utilizadas para refinar as recomendações principais. Este método é especialmente vantajoso em situações em que as recomendações geradas estão empatadas em termos de previsão de classificação. Nestes cenários utiliza-se o SR secundário para “desempatar” as recomendações.
- *Meta-level*: Semelhante à *feature augmentation*, este método emprega um SR que adicionará novas informações ao *dataset* original. A diferença entre os métodos recai sobre a quantidade de informações que são adicionadas ou alteradas ao original, neste método, o *dataset* original é substituído por um gerado pelo SR contributivo, sendo o substituto utilizado como input no SR principal.

### **3.2. Sistemas de recomendação para a diabetes**

Durante o levantamento bibliográfico foram utilizadas as seguintes plataformas web: Science Direct, Reaserch Gate, Springer e Google Scholar para a recolha de artigos referentes ao desenvolvimento de sistemas de recomendação de planos nutricionais e de exercício físico. Foi tido em conta o ano de publicação do mesmos, selecionaram-se publicações entre 2015 e 2022. As palavra-chave utilizadas foram: “sistema de recomendação nutricional”, o nome da abordagem de recomendação empregue e “sistema de recomendação de exercício físico”.

Durante a fase inicial priorizou-se artigos de revisão sobre sistemas de recomendação direcionados á diabetes. Dos artigos enumerados destacaram-se aqueles que apresentavam melhor detalhe na descrição da abordagem utilizada, bem como os métodos empregues, e teve-se em atenção às funcionalidades inerentes ao sistema.

Dos artigos analisados selecionaram-se 5 que serão apresentados na presente secção. Esta encontrasse dividida entre sistemas de recomendação nutricional, onde são apresentados os artigos direcionadas a este tema, seguida da subsecção referente à recomendação de exercício físico.

### **3.2.1. Sistemas de recomendação nutricional**

Os SRN são uma das principais tecnologias aplicadas no âmbito da NI, tendo como objetivo apoiar os utilizadores durante a mudança de comportamentos alimentares. São tecnologias eficientes para a extração de informações dos hábitos alimentares e preferência de alimentos do utilizador [53]. Nesta secção será feito um levantamento literário de sistemas de recomendação nutricionais.

#### **3.2.1.1. Sistema de recomendação de snacks do iraniano**

Norouzi et al. (2018) descrevem uma plataforma para smartphones de recomendação de snack para pacientes diabéticos. A preferência de snack advém dos hábitos alimentares da população iraniana. Na cultura iraniana é usual a realização de 3 refeições diárias, no entanto, para um paciente diabético esta dieta não é aconselhável uma vez que é propícia a episódios de hipoglicemia e por isso recomenda-se aos pacientes que realizem pequenas refeições entre as refeições principais, como snacks. Com vista a remediar este problema os autores desenvolveram um SR de snack.

O estudo descreve o desenvolvimento e avaliação de um SR utilizando um sistema de recomendação baseado em conhecimento, baseado em restrições com a aplicação do método *roulette wheel selection*, As recomendações tem em atenção as diretrizes da ADA, as preferências do utilizador e as suas restrições alimentares e a sazonalidade dos alimentos. Após a sua precisão ter sido testada por três nutricionistas e os seus planos nutricionais por 30 pacientes diabéticos foi possível comprovar que o sistema dá prioridade às restrições alimentares do individuo, fator que é realçado em indivíduos com índices de massa corporal maiores, e que as recomendações vão ao encontro da disponibilidade sazonal dos alimentos [54].

#### **3.2.1.2. Food recommendation system using clustering analysis for diabetic patients**

O estudo desenvolvido por Phanich et al. (2010) no Departamento de matemática da universidade de Bangkok, Tailândia, descreve desenvolvimento de um SR de alimentos com o intuito de substituir alimentos prejudiciais por alimentos benéficos que auxiliem o controlo da diabetes. O sistema é projetado utilizando *Self-Organizing Map* e

*K-mean clustering* para o agrupamento de alimentos. A similaridade entre alimentos é calculada a partir dos oito nutrientes significativos para os diabéticos: Carboidratos, energia, gordura, proteína, fibras, vitamina E, tiamina e vitamina C. O sistema foi avaliado por nutricionistas demonstrando grande utilidade na gestão da doença [55].

### **3.2.2. Sistemas de recomendação de exercício físico e dieta**

Sistemas de recomendação de dieta e exercícios são frequentemente desenvolvidos para um público geral, como consequência, fatores como o estilo de vida, preferências de exercícios e dieta são mitigados. Este tipo de sistema não responde totalmente às necessidades dos pacientes com DM tipo II que necessitam de recomendações personalizadas e direcionadas ao caso de cada paciente. Uma vez que a dieta e o exercício estão fortemente relacionados na gestão da diabetes e especial atenção deve ser tomada no balanço entre ambas as atividades para controlar de forma eficaz o nível de glucose é comum na literatura a escolha de SR que permitem relacionar conceitos e requisitos de ambos os domínios assim como o perfil de saúde do indivíduo e a doença [56]. Nesta secção será feito um levantamento literário de sistemas de recomendação de exercício e dieta.

#### **3.2.2.1. Sistema de recomendação híbrido para sugestão de exercícios e dietas.**

Ali et al. (2018) apresentam uma plataforma de sugestão de exercício físico e dieta direcionado para a população de meia-idade. O trabalho foi desenvolvido pelo departamento de ciência computacional da universidade de Seoul, Coreia do Sul, e foca-se também na educação dos seus utilizadores através de notificações sobre a importância da prática física e da dieta na qualidade de vida do indivíduo. As recomendações de prática física têm como propósito alcançar objetivos para a queima de calorias. A recomendação dietética tem em consideração os requisitos nutricionais do indivíduo, nomeadamente a quantidade diária de calorias a ingerir. O trabalho explora superficialmente condições médicas como a diabetes à qual recomenda dietas com alimentos de baixo teor glicémico [57].

### 3.2.2.2. SHADE

Fraiz et al. (2017), uma equipa da Universidade Nacional das Ciências e tecnologias e Islamabad, Paquistão, desenvolveu um SR de dietas e atividade física para o auxílio na gestão da DM tipo II. O sistema intitulado como SHADE, um acrónimo para *Semantic Healthcare Assistant for Diet and Exercise*, foi implementado segundo uma abordagem integrada baseada em ontologia que combina conhecimentos de vários domínios para originar sugestões de dietas e exercícios para diabéticos.

As recomendações do SHADE são uma reflexão das preferências do indivíduo, fatores do perfil da saúde pessoal, o nível de glucose no sangue e restrições alimentares e físicas impostas pela doença. As sugestões de dieta são geradas sob a forma de menu de refeição onde a dimensão da porção é especificada. Relativamente às sugestões de exercício são uma combinação das preferências do utilizador e das limitações impostas pela diabetes, tendo o acréscimo de estarem especificadas quanto à duração e intensidade do exercício. É um sistema de recomendação baseado em conhecimento implementado segundo *Web Ontology Language* [15].

### 3.2.2.3. Sistema de recomendação personalizado de dieta e exercício para a diabetes tipo I

Xie et al. (2017) apresentam o desenvolvimento de um SR de dieta e exercício físico para pacientes com DM tipo I. É um sistema de recomendação baseado em modelo, um subgrupo da recomendação por filtragem colaborativa. Faz 5 recomendações diárias da quantidade de carboidratos a ingerir por refeição, determina o tamanho e porções ideais de cada refeição, determina a frequência cardíaca alvo durante os exercícios, faz a recomendação de quantidades de carboidratos a ingerir antes do exercício e recomenda lanches antes de dormir para minimizar o risco clínico. As recomendações são feitas segundo um horário que pode ser personalizado pelo utilizador. O sistema foi testado em dois grupos classificados segundo a experiência na autogestão da diabetes: *Starter* e *Skilled*. Segundo os autores, a utilização da plataforma permitiu a diminuição do risco clínico médio em 49 % no grupo *Starter* e uma diminuição de hipoglicemia em eventos noturnos e pós-exercício [18].

### 3.3. Discussão:

Nesta secção será feita uma breve discussão acerca dos SR, quais as vantagens e funcionalidades. Na tabela 3.1 são apresentadas as funcionalidades dos sistemas.

**Tabela 3.1** – Comparação entre as funcionalidades dos SR destacados

<i>FUNCIONALIDADE</i>	[54]	[55]	[57]	[15]	[18]
<i>SUBSTITUIÇÃO DE ALIMENTOS</i>		X			
<i>RECOMENDAÇÃO DE REFEIÇÕES</i>	X		X	X	X
<i>RECOMENDAÇÃO DE EXERCÍCIO FÍSICO</i>			X	X	X
<i>RECOMENDAÇÃO DE EXERCÍCIOS DIRECIONADA AO CONTEXTO ONDE SE ENCONTRA O UTILIZADOR</i>			X	X	
<i>RECOMENDAÇÃO DE EXERCÍCIO FÍSICO DIRECIONADO ÀS PREFERÊNCIAS DO UTILIZADOR</i>				X	X
<i>RECOMENDAÇÃO DE ALIMENTOS DE ACORDO COM A ÉPOCA</i>	X		X		
<i>RECOMENDAÇÃO DE ALIMENTOS DE ACORDO COM AS PREFERÊNCIAS DO UTILIZADOR</i>	X	X			X
<i>RECOMENDAÇÃO DE REFEIÇÕES COM BASE NO NÍVEL GLICÉMICO DO UTILIZADOR</i>				X	
<i>RECOMENDAÇÃO DE REFEIÇÕES QUE RESPEITAM A DOSE DIÁRIA DE INGESTÃO DE NUTRIENTES DO UTILIZADOR</i>	X		X		X
<i>IDENTIFICAÇÃO DE FATORES ASSOCIADOS AO GASTO ENERGÉTICO POR DIA E NOITE</i>	X		X	X	X
<i>UTILIZAÇÃO DE SENSORES PARA ESTUDAR O COMPORTAMENTO DO UTILIZADOR</i>			X		
<i>MONITORIZAÇÃO DO NÍVEL GLICÉMICO</i>				X	
<i>PREVÊ O NÍVEL GLICÉMICO DO UTILIZADOR</i>					X
<i>PREVÊ O RITMO CARDÍACO DO UTILIZADOR</i>					X
<i>ESTIMAÇÃO DO GASTO DE ENERGIA DIÁRIO</i>	X		X	X	
<i>ESTIMAÇÃO DA ENERGIA A SER INGERIDA POR REFEIÇÃO</i>	X		X	X	

<b>ESTIMAÇÃO DOS NUTRIENTES A SEREM INGERIDOS POR REFEIÇÃO</b>				X
<b>CRIAÇÃO DE REGRAS PARA A RECOMENDAÇÃO</b>	X		X	X
<b>INTERFACE GRÁFICA</b>	X		X	X
<b>RANKING DE ALIMENTOS</b>	X			X
<b>HORÁRIO PERSONALIZÁVEL DE REFEIÇÕES E EXERCÍCIOS FÍSICOS</b>				X
<b>REGISTO DAS PREFERÊNCIAS DO UTILIZADOR</b>	X		X	X
<b>NOTIFICAÇÕES/ALARMES/ADVERTÊNCIAS NUTRICIONAIS E FÍSICAS</b>	X		X	X
<b>PERFIL DE UTILIZADOR</b>	X		X	X
<b>RELATÓRIOS PERIÓDICOS</b>	X			
<b>REGISTO DA ATIVIDADE FÍSICA EXECUTADA DURANTE O DIA</b>	X		X	X
<b>REGISTO DA ENERGIA CONSUMIDA DURANTE AS REFEIÇÕES</b>	X		X	X
<b>PERMITE A ADIÇÃO DE NOVOS ALIMENTOS</b>			X	
<b>PERMITE A ADIÇÃO DE NOVOS EXERCÍCIOS FÍSICOS</b>			X	
<b>TIPO DE APLICAÇÃO</b>	Móvel	Apenas sistema de recomendação	Movel	Desktop

Após uma análise à tabela 3.1 é possível destacar 10 funcionalidades que são partilhadas por 3 SR. São exemplos o cálculo das calorias gastas diariamente, cálculo da quantidade de calorias a ingerir por refeição, recomendação de exercício físico, utilização e criação de regras de recomendação e uso de interface gráfica. Para além existem 4 funcionalidades partilhadas por 4 dos sistemas de recomendação, são exemplo a recomendação de refeições, a identificação de fatores associados ao gasto energético, a criação de um perfil do utilizador onde são especificadas preferências e a condição médica do mesmo e a utilização de um sistema de notificações com o intuito de instruir o utilizador para a adoção de hábitos de vida mais saudáveis.

Por fim, é possível perceber que 2 dos 5 sistemas destacados combinam o sistema de recomendação com uma aplicação móvel.

De seguida será feito um levantamento das abordagens utilizadas no desenvolvimento dos sistemas, bem como as técnicas utilizadas para o desenvolver, as limitações e os testes realizados ao sistema, tabela 3.2.

**Tabela 3.2** – Técnicas utilizadas para a concessão dos SR destacados

	[54]	[55]	[57]	[15]	[18]
<b>SISTEMA DE RECOMENDAÇÃO</b>	Recomendação baseada em conhecimento	Filtragem Colaborativa	Sistema de recomendação híbrido ( <i>context aware</i> + recomendação baseada em restrições)	Recomendação baseada em conhecimento	Filtragem Colaborativa
<b>TÉCNICA UTILIZADA</b>	<i>constraint-based reasoning and roulette wheel algorithm</i>	<i>Clustering Analysis; Self-Organizing Map; K-mean clustering</i>	-	Ontologia	<i>Nonlinear Autoregressive Moving Average with Exogenous Inputs Model (NARMAX)</i>
<b>LIMITAÇÕES</b>	-	-	Não permite o feedback do utilizador. Não regista a ação do utilizador, isto é, se comeu uma das recomendações ou praticou uma das atividades	Até à data não foi testado por utilizadores reais. Não monitoriza o ritmo cardíaco do utilizador	Não tem em conta o contexto do utilizador, ou seja, estado psicológico, risco clínico para a hipoglicemia, estilo de vida

				físicas sugeridas.  Qualidade das notificações e conselhos nutricionais deve ser aumentada	
<b>TESTES</b>	Testado por nutricionistas	-	-	Ainda não foi testado por utilizadores reais	Simulação da utilização da aplicação por diabéticos com pouca/muita experiência na gestão da diabetes

Analisando a tabela 3.2 é possível perceber uma preferência pela recomendação baseada em conhecimento e por filtragem colaborativa. Das limitações analisadas estas prendem-se à monitorização do utilizador, através do registo de receitas experimentadas ou da sua condição fisiológica, e são feitas menções à necessidade de perceber qual o contexto do utilizador quando solicita as recomendações. Quanto ao método de teste, 2 dos artigos destacados mencionam a testagem do algoritmo com utilizadores reais: nutricionistas ou pacientes com diabetes.



## CAPÍTULO 4 – PROPOSTA DE SOLUÇÃO



## 4. Proposta de Solução

No capítulo 2 foram enumerados os métodos utilizados na gestão da diabetes, no entanto, a sua concretização, e os resultados alcançados pelos pacientes, mostram-se dispersos. Isto é, embora a estrutura do plano seja a mesma, cada paciente irá alcançar resultados diferentes, que podem ser benéficos ou não ao mesmo. Por isso, neste trabalho, é proposto a concretização de um SR que identifique o perfil de preferências do utilizador e faça a recomendação de receitas que melhor se alinhem com o mesmo.

Para o concretizar serão empregues duas abordagens: sistemas de recomendação baseado em conteúdo e filtragem colaborativa. No final, o sistema será submetido a testes para perceber sob que condições atua melhor. Será a abordagem da filtragem colaborativa melhor ou terá a recomendação baseada em conteúdo melhores resultados, ou ainda, será preferível utilizar uma abordagem híbrida que conjugue os benefícios das duas?

Este capítulo está dividido em cinco secções, cada uma dedicada às várias etapas que levaram à concessão da solução proposta. O capítulo é iniciado pela enumeração das tecnologias utilizadas no desenvolvimento do algoritmo, seguido da apresentação do *dataset*, modelo de domínio e arquitetura do sistema. Termina com a descrição da implementação do algoritmo.

### 4.1. Tecnologias

Para desenvolver a solução utilizou-se Java como linguagem de programação. É uma linguagem orientada a objetos auxiliada pelo IntelliJ, um compilador de código, onde foram definidas as diferentes classes [58].

### 4.2. Dataset

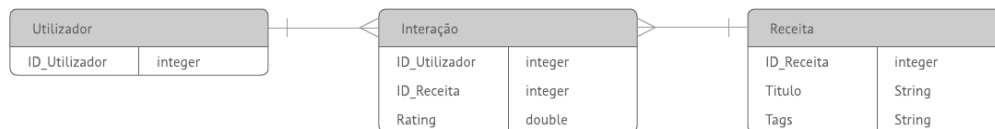
Antes de desenvolver o SR foi necessário ter previamente preparados dados referentes a receitas, utilizadores e interação entre os mesmos. Das opções encontradas destacou-se um conjunto de ficheiros Excel com dados referentes ao site **Food.com**. Nos documentos encontram-se apresentados dados relativos a receitas: nome, id, *tags* associadas, ingredientes, modo de preparação e comentários; e interações entre utilizadores com receitas: Id do utilizador, id da receita experimentada, cotação atribuída

à receita, data de interação com a receita. O dataset contém informações relativas a 231.215 receitas e 25.083 utilizadores

Este conjunto de documentos provém da plataforma digital *Kaggle* e foram utilizados para desenvolver tanto o sistema de recomendação baseado em conteúdo como por filtragem colaborativa [59].

### 4.3. Modelo de domínio

O modelo de domínio é uma representação visual das entidades que compõem um sistema e a relação entre as mesmas. Na Figura 4.1 encontra-se esquematizado as entidades e as relações que compõem o sistema desenvolvido.



**Figura 4.1** – Modelo de Domínio

Uma vez que se trata de um SR, simplificou-se as entidades e as associações de forma a serem armazenados apenas os dados necessários para a concessão do sistema:

- O utilizador, é representado pela classe Utilizador, composta apenas pelo ID, como forma de identificação.
- As receitas são representadas pela entidade Receita composta pelos atributos ID, Título e *tags*, sendo o último uma lista *de* todas as *tags* associadas à receita.
- A entidade auxiliar Interação armazena o ID do utilizador e o ID da receita com que interagiu, bem como a cotação, ou rating atribuída à mesma.

### 4.4. Arquitetura

Nesta secção, será apresentada a arquitetura do sistema. O algoritmo desenvolvido é composto por dois sistemas de recomendação. Optou-se pela utilização de algoritmos de recomendação baseado em conteúdo e por filtragem colaborativa sendo estes os mais utilizados para a identificação de preferências. Uma vez que as receitas são direcionadas

às principais refeições como o almoço e o jantar, as recomendações são direcionadas para estas refeições.

**Tabela 4.1** – Dados de entrada do utilizador com ID 12

ID UTILIZADOR	RECEITAS CLASSIFICADAS	CLASSIFICAÇÕES ATRIBUÍDAS
12	60532	5
	9369	5
	5	5

**Tabela 4.2** – Dados de entrada das receitas com Id: 60532, 9369 e 5.

NOME DA RECEITA	TAGS	ID DA RECEITA
<i>ITALIAN BEEF N' BEER</i>	<i>time-to-make, course, main-ingredient, cuisine, preparation, occasion, lunch, easy, European, fall, holiday-event, spring, summer, winter, Italian, stove-top, dietary, seasonal, sandwiches, comfort-food, meat, Super Bowl, taste-mood, equipment, number-of-servings</i>	60532
<i>GRILLED CAULIFLOWER</i>	<i>60-minutes-or-less, time-to-make, course, main-ingredient, preparation, 5-ingredients-or-less, side-dishes, vegetables, barbecue, easy, dietary, cauliflower, equipment, grilling</i>	9369
<i>LAMB STEW WITH TOMATOES CHICKPEAS AND SPICES</i>	<i>time-to-make, course, main-ingredient, cuisine, preparation, occasion, main-dish, lamb-sheep, African, oven, Moroccan, stove-top, dietary, spicy, low-carb, low-in-something, meat, taste-mood, equipment, 4-hours-or-less</i>	5

Nas tabelas 4.1 e 4.2 são apresentados os dados de entrada do utilizador com ID 12 e as informações referentes às receitas classificadas. Na primeira tabela podemos observar o ID das receitas classificadas e a respetiva classificação. Na segunda tabela estão apresentadas as seguintes informações: título, ID e *tags* associadas.

Nos próximos parágrafos será feita uma descrição dos sistema, mas antes de a iniciar é necessário notar que o utilizador que solicita uma recomendação será denominado por *active user*, do inglês, “utilizador ativo”.

O sistema começa por receber informações do *active user* referentes ao histórico de classificações. Como modo de descrição das receitas são utilizadas as *tags*. A partir de um levantamento destas são calculados pesos que representam o nível de preferência pela *tag* por parte do *active user*. Pesos elevados são identificados como mais apreciados. Com base neste conhecimento, são feitas sugestões de receitas compostas por *tags* com pesos maiores e cuja similaridade entre as *tags* que as compõem e as *tags* identificadas como preferências seja mais elevada. Este é o método utilizado pelo algoritmo de recomendação baseado em conteúdo.

O sistema desenvolvido utilizando a filtragem colaborativa faz uso das receitas que o *active user* já classificou e respetivas classificações, e procura o top 20 de utilizadores cujas receitas e classificações mais se assemelhem ao *active user*. Para tal, faz uso do *K nearest neighbors* (KNN). Dentro dos 20 utilizadores mais semelhantes faz se um levantamento das receitas que estes já classificaram e removem-se quaisquer receitas que o *active user* já tenha classificado. É feita uma média das classificações atribuídas e este valor é utilizado como previsão. No final, a lista de receitas é ordenada de forma descendente segundo o valor de previsão de classificações e apresentada.

## **4.5. Implementação do sistema de recomendação**

Nesta secção será descrita a implementação do sistema de recomendação baseado em conteúdo e por filtragem colaborativa. Ambos os algoritmos desenvolvidos seguem a metodologia descrita por Noratiqah Mohd Ariff et all [38]. A secção será iniciada pela apresentação da metodologia empregue pelo sistema de recomendação baseado em conteúdo e finaliza com a metodologia da recomendação por filtragem colaborativa.

### **4.5.1. Sistema de recomendação baseado em conteúdo:**

Começando a descrição pelo algoritmo de recomendação baseada em conteúdo, tem como input todas as receitas do sistema bem como as interações com utilizadores. Durante a análise dos dados, constroem-se duas matrizes. A primeira, matriz A, guarda a informação relativa às *tags* presentes em todas as receitas do sistema. Tem a dimensão  $m \times n$ , sendo  $m$ , o número *tags*, e  $n$  o número de receitas. Se a receita  $n_i$  tiver a *tag*  $m_i$

então nessa posição é colocado um 1. Caso contrario será colocado um 0. Na figura 4.2 é possível observar um excerto do código desenvolvido.

```

for (String tag :UniqueTagValues){
    ArrayList<Integer> containsRecipeTag = new ArrayList<>();
    for (Recipe recipe : activeUserRecipes){
        boolean contains = recipe.getRecipeTag().contains(tag);
        if (contains){
            containsRecipeTag.add(1);
        }
        else{
            containsRecipeTag.add(0);
        }
    }
    RecipesMatrix.add(containsRecipeTag);
}

```

**Figura 4.2** – Criação da matriz A.

A segunda matriz, matriz B, guarda todas as receitas do sistema e a respetiva classificação dada pelo *active user*, formando uma matriz com dimensões  $n \times 1$ , sendo  $n$  o número de receitas. Neste caso, se o *active user* tiver dado à receita  $n_i$  uma classificação igual ou superior a 4 é atribuído um 1 nessa posição. No caso de a classificação ser inferior, atribui-se -1. Nos casos em que o *active user* não classificou a receita atribui-se um 0, figura 4.3.

```

HashMap<Integer,Double> Rating = user.getInteractions();
//Recipe ID, Rating 1-1
ArrayList<Integer> FinalMatrix = new ArrayList<Integer>();
Rating.forEach((key, rate) -> {if (rate>=4.0){
    FinalMatrix.add(1);
} else if (rate<4.0 & rate>0.0){
    FinalMatrix.add(-1);
} else {
    FinalMatrix.add(0);
}
});

```

**Figura 4.3** – Criação da matriz B.

As duas matrizes são multiplicadas, gerando a matriz de pesos, tem dimensões  $n \times 1$  e guarda o peso que cada *tag* tem sobre o gosto do *active user*, figura 4.4.

```
ArrayList<Integer> userLikes = new ArrayList<Integer>();
int sum=0;
for (ArrayList<Integer> elemArrayMatrixA: MatrixA){
    for (int index=0; index<elemArrayMatrixA.size();index++){
        double a= elemArrayMatrixA.get(index);
        double b= MatrixB.get(index);
        sum+=elemArrayMatrixA.get(index)*MatrixB.get(index);
    }
    userLikes.add(sum);
}
```

**Figura 4.4** – Criação da matriz peso.

O algoritmo irá calcular a semelhança que cada receita do sistema tem com a matriz de peso e para tal utilizará a função *cosine similarity*, (E.q 3.1) e figura 4.5. Esta função calcula a similaridade entre a matriz peso e uma dada linha da matriz A, que representa uma receita.

```
for (int index=0; index<recipe.size();index++){
    sum+= recipe.get(index) * userLikes.get(index);
    toSqrtRecipe += Math.pow(recipe.get(index),2);
    toSqrtLikes += Math.pow(userLikes.get(index),2);
}

quotient=Math.sqrt(toSqrtRecipe)* Math.sqrt(toSqrtLikes);
if (quotient==0){
    alfa=0.0;
}else {
    alfa = sum / quotient;
}
return alfa;
```

**Figura 4.5** – Cosine similarity.

É calculada a similaridade para todas as receitas do sistema e apresentado ao *active user* o top 20 de receitas com maior similaridade, figura 4.6.

```
HashMap<Recipe,Double> similarity =new HashMap<>();

Recipe obj;

for (int index=0; index<arrayOfRecipes.size(); index++){
    double alfa = cosineSimilarity(arrayOfRecipes.get(index),userLikes);
    obj=recipeArray.get(index);
    similarity.put(obj,alfa);
}
```

**Figura 4.6** – Calculo da similaridade entre receitas.

#### 4.5.2. Sistema de recomendação por filtragem colaborativa

O algoritmo de filtragem colaborativa organiza, numa matriz, todas as classificações atribuídas pelos utilizadores do sistema. Durante organização dos dados, o algoritmo altera as classificações iguais a 0 (zero) para 0,5 e qualquer receita que não tenha sido classificada por qualquer utilizador, receberá um 0 (zero) na posição correspondente. O resultado é uma matriz, denominada matriz M, com classificações entre 0,5 e 5,0, figura 4.7.

```

for (User user: UserArray){
    ArrayList<Recipe> userRatedRecipes= user.getRecipesRated();
    ArrayList<Double> row_rating_per_user= new ArrayList<>();
    double magnitude=0.0;

    for (Recipe recipe: RecipeArray){
        boolean contains = userRatedRecipes.contains(recipe);
        if (contains){
            int index = userRatedRecipes.indexOf(recipe);
            double rating = user.getRating().get(index);
            magnitude+= Math.pow(rating,2);
            if (rating==0.0){
                row_rating_per_user.add(0.5);
            }
            else {
                row_rating_per_user.add(rating);
            }
        }
        else {
            row_rating_per_user.add(0.0); //User did not rated this recipe
        }
    }
    normalizeArray(row_rating_per_user, magnitude);
    matrixM.add(row_rating_per_user);
}
return matrixM;

```

**Figura 4.7** – Criação da matriz M.

Aplicando a equação 4.2, normaliza-se a matriz M.

$$\hat{u} = \frac{\vec{u}}{\|\vec{u}\|} \quad (\text{Eq. 4.1})$$

Sendo que  $\vec{u}$  um dado vetor;  $\|\vec{u}\|$  a sua magnitude; e  $\hat{u}$  o mesmo normalizado.

De seguida encontram-se os 20 utilizadores mais semelhantes ao *active user*. Como anteriormente mencionado, faz-se uso do algoritmo KNN. Este algoritmo fez uso da fórmula *cosine similarity*, (Eq. 3.1), para calcular a similaridade entre todos os utilizadores do sistema e o *active user*. Por fim a lista é ordenada de forma decrescente de similaridade e são recolhidos os 10/20/30 utilizadores mais semelhantes.

A partir deste conjunto, guarda-se as receitas classificadas pela amostra e eliminam-se quaisquer receitas classificadas pelo *active user*. Utilizando a equação 4.3 e a figura 4.8 faz-se uma previsão de classificação para o conjunto de receitas selecionado.

No final as receitas são ordenadas por ordem decrescente de previsão de classificação e apresentadas ao *active user*.

$$\hat{r}_{aj} = \frac{1}{N} \sum_i r_{ij} \quad (\text{Eq. 4.2})$$

Onde  $\hat{r}_{aj}$  representa a previsão de classificação que o *active user* atribuíra à receita  $j$ . A variável  $N$  representa a quantidade de vezes que a receita  $j$  foi classificada. Por fim  $r_{ij}$  representa a classificação que o utilizador  $i$  atribuiu à receita  $j$ .

```
public double calc_rating_prediction(ArrayList<Double> ratings_for_recipe, int freq){
    double prediction=0.0;
    for (double rating: ratings_for_recipe){
        prediction +=rating;
    }
    return prediction /freq;
}
```

**Figura 4.8** – Previsão de classificação



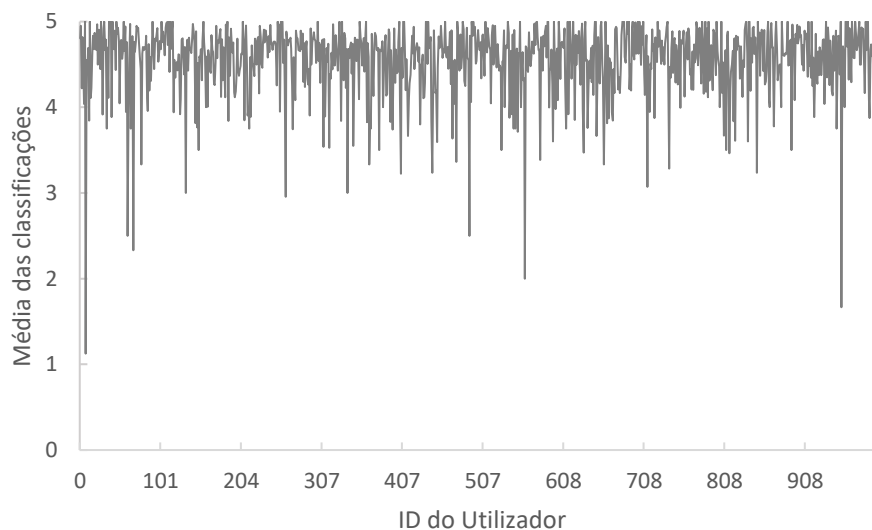
CAPITULO 5 – CASOS DE ESTUDO E  
RESULTADOS



## 5. Casos de Estudo e Resultados

Neste capítulo serão apresentados os casos de teste utilizados para avaliar a solução desenvolvida. Será iniciado pelo subcapítulo 5.1 onde serão enunciados os casos de teste desenvolvidos para o sistema de recomendação baseado em conteúdo e por filtragem colaborativa. No subcapítulo 5.2, serão apresentados os resultados obtidos.

Antes de iniciar a apresentação dos casos de estudo será feita uma breve chamada de atenção ao *dataset* utilizado. Este é composto por 231.215 receitas e 25.083 utilizadores [59]. Devido à sua extensão, foi necessário diminuir o número de utilizadores para 1000 e o de receitas para 14099. Das receitas seleccionadas foram identificadas 648 *tags*.



**Figura 5.1** – Média de classificações por utilizador

Na figura 5.1 é apresentada a média de classificações dos 1000 utilizadores. No eixo das ordenadas encontra-se a média das classificações, no eixo das abcissas o ID dos utilizadores. Como é possível observar, a grande maioria da média de classificação encontra-se na região dos 5 valores, o que remete para cenários onde os utilizadores tem uma grande tendência para classificar receitas com classificação máxima. De maneira a testar o algoritmo fora destas condições foram seleccionados utilizadores com média de classificação entre os 3 e os 4 valores. Dos obtidos, serão apresentados os dados de entrada e os resultados obtidos para os utilizadores com ID 39, 309 e 444. Como estes têm uma grande quantidade de classificação atribuídas, não é conveniente apresentar

todas as receitas classificadas nem as respectivas classificações, por isso, na tabela 5.1 será apresentado um resumo dos dados de entrada. Faço ainda um chamada de atenção referente aos mesmos, em ambos os sistemas foram utilizados as seguintes informações: ID do utilizador, lista de receitas classificadas, lista de classificações atribuídas.

**Tabela 5.1** – Resumo dos dados de entrada para os casos de teste

ID UTILIZADOR	QUANTIDADE DE CLASSIFICAÇÕES	MÉDIA DE CLASSIFICAÇÕES
39	540	3,89
309	203	3,54
444	17	3,24

Os três utilizadores foram escolhidos não só pela média de classificações, mas também pela quantidade de classificações atribuídas.

## 5.1. Casos de estudo

Neste subcapítulo serão apresentados os casos de estudo utilizados para testar a solução desenvolvida.

**Caso de Estudo 1: A partir das classificações que *tags* irá o sistema identificar como preferências do *active user*?**

### Objetivo:

- Identificar e quantificar as *tags* com a qual o *active user* tem maior preferência.

### Pré-requisitos:

- O *active user* deverá ter classificado receitas previamente.
- Utilização do sistema de recomendação baseado em conteúdo.

### Etapas:

1. Identificar as *tags* associadas às receitas classificadas pelo *active user*.
2. Através das classificações identificar quais as *tags* associadas a receitas com classificações mais altas.

**Caso de Estudo 2: A partir das classificações atribuídas recomendar receitas que vão ao encontro das preferências do *active user*:**

**Objetivo:**

- Recomendar 9 receitas.

**Pré-requisitos:**

- O *active user* deverá ter classificado receitas previamente.
- Utilizar o algoritmo de recomendação baseada em conteúdo.

**Etapas:**

1. Identificar as *tags* preferidas do utilizador.
2. Calcular a similaridade entre as *tags* preferidas pelo *active user* e as *tags* associadas às receitas.
3. Recomendar as 9 receitas com maior similaridade.

**Caso de Estudo 3: A partir de um conjunto de utilizadores, identificar os utilizadores mais semelhantes ao *active user*, quais as receitas que ainda não classificou e prever qual a classificação que dará.**

**Objetivos:**

- Perceber qual o impacto na previsão de classificações quando o tamanho da vizinhança do *active user* aumenta ou diminui.
- Recomendar 9 receitas.

**Pré-requisitos:**

- O *active user* deverá ter classificado receitas previamente.
- Utilizar o algoritmo por filtragem colaborativa.

**Etapas:**

1. Identificar a vizinhança do *active user*, deverá ser um conjunto de 10 utilizadores.
2. Identificar quais as receitas que foram classificadas pela vizinhança, mas que ainda não foram classificadas pelo *active user*.
3. Prever a classificação que o *active user* dará a estas receitas.

4. Recomendar as 9 receitas com previsões de classificação mais altas.
5. Repetir as etapas 1-4 para uma vizinhança composta por 20 e 30 utilizadores.

## 5.2. Resultados

Nesta secção será feito o levantamento dos resultados obtidos. Iniciando com o **caso de estudo 1**, podemos observar nas tabelas 5.2 as *tags* identificadas como preferências:

**Tabela 5.2** – Preferências dos utilizadores 39, 309 e 444

ID	PREFERÊNCIAS	QUANTIDADE DE TAGS
39	<i>equipment, main-dish, taste-mood, comfort-food, number-of-servings, bacon, dietary, meat, oven, inexpensive, main-ingredient, eggs-dairy, occasion, potluck, beginner-cook, time-to-make, European, to-go, German, vegetables, picnic, lunch, preparation, heirloom-historical, easy, cuisine, course, 60-minutes-or-less, holiday-event, pork, stove-top</i>	31
309	<i>equipment, main-dish, occasion, beginner-cook, time-to-make, poultry, soups-stews, North American, one-dish-meal, preparation, taste-mood, comfort-food, easy, number-of-servings, chicken, cuisine, course, 30-minutes-or-less, meat, inexpensive, stove-top, main-ingredient</i>	23
404	-	0

As *tags* das receitas são um modo de identificação das mesmas e podem conter informações quanto à quantidade de tempo que levam a preparar, qual os ingredientes que utilizam, a que estação do ano se direcionam entre outras informações. Podemos observar que os utilizadores 39 e 309 tem uma grande quantidade de *tags* identificadas, 31 e 23 respetivamente, no entanto para o utilizador 444 o sistema não foi capaz de identificar nenhuma.

Para o utilizador com ID 39 podemos destacar 3 *tags* direcionadas a ocasiões, este utilizador parece preferir receitas que possam ser confeccionadas de forma a partilhar com outros, destacam-se *occasion, holiday-event, potluck, picnic*. Também parece ter em

atenção ao número de porções, o que corrobora com a hipótese apresentada. É possível identificar uma preferência por carne, *meat*, *pork*, *bacon* e pratos assados *oven*.

O utilizador com ID 309 demonstra preferência por galinha ou carne de aves: *poultry*, *meat* e *chicken*. Também demonstra afinidade por receitas que sejam rápidas ou de fácil de preparação *30-minutes-or-less*, *time-to-make*, *easy*. *Stove-top* remete para a preferência por pratos que sejam cozinhados utilizando o fogão.

No **caso de estudo 2** as seguintes receitas foram recomendadas tabela 5.3 e tabela 5.4.

**Tabela 5.3** – Recomendações para o active user 39.

ID	TÍTULO	SIMILARIDADE
55132	Sausage and bell peppers	0,781
2775	Haluski	0,754
63048	My favorite beef paprika	0,740
131023	Hamburgers with brown gravy total comfort food A.K.A meat cakes	0,726
25455	Bacon tomato rotini	0,713
97849	Easy kielbasa casserole	0,713
71812	Italian casserole dinner	0,710
91341	Hearty bacon potato chowder	0,710
135192	Pennsylvania dutch potato bake	0,710

**Tabela 5.4** – Recomendações para o active user 309.

ID	TÍTULO	SIMILARIDADE
149606	One dish chicken and bows	0,81
89129	Buffalo chicken mac n cheese	0,77
47433	Chicken turnovers	0,76
120398	Italian parmesan baked chicken	0,75
59702	7 vegetable chicken stew with dumplings	0,75
127080	Oven fried chicken chimichangas	0,75
107810	Barbecued oven baked chicken thighs	0,74
102463	Turkey tacos	0,74
96346	Apricot glazed chicken and couscous	0,74

Uma vez que o utilizador com ID 444 não tem nenhuma *tag* nas preferências, não foi possível calcular a similaridade entre as preferências e as receitas detendo o sistema de fazer recomendações.

Das recomendações direcionadas ao utilizador com ID 309, podemos identificar 8 receitas com galinha como um dos ingredientes, *chicken* é uma das *tags* identificada como preferências.

No **caso de estudo 3** foram selecionadas 9 receitas para cada *active user*. Nas tabelas 5.5 a 5.13 encontramos as recomendações para uma vizinhança de 10, 20 ou 30 utilizadores, é necessário notar que o *k* remete para o tamanho da vizinhança.

**Tabela 5.5** – Recomendações para o *active user* 39 utilizando K=10.

ID	TÍTULO	PREVISÃO DE CLASSIFICAÇÃO
56425	Banana bread	5,0
151465	Fried egg cheese sandwich	5,0
146382	Cajun style oven fries	5,0
164837	Sweet salsa dump chicken oamc	5,0
30076	Grilled ham with glaze	5,0
44551	Tortured chicken beer can	5,0
65543	Easy stove top macaroni cheese	5,0
159543	Sour cream chive bread machine	5,0
55226	Strawberry julius	5,0

**Tabela 5.6** – Recomendações para o *active user* 309 utilizando K=10.

ID	TÍTULO	PREVISÃO DE CLASSIFICAÇÃO
139822	Kittenball's 5-minute cinnamon flop brunch cake	5,0
46490	The sweetest blueberry muffins	5,0
67898	Baked ham and cheese omelet roll	5,0
19812	Simply sour cream chicken enchiladas	5,0
28469	Parmesan crusted tilapia	5,0
130480	Copycat green giant nibbles corn in butter sauce	5,0
122140	Chickpea curry garbanzos	5,0
45824	Ednas apple crumble aka apple crisp	5,0
135961	Whatever floats your boat brownies	5,0

**Tabela 5.7** – Recomendações para o *active user* 444 utilizando K=10.

ID	TÍTULO	PREVISÃO DE CLASSIFICAÇÃO
117899	Creamy cajun chicken pasta	4.0

**Tabela 5.8** – Recomendações para o active user 39 utilizando K=20.

ID	TÍTULO	PREVISÃO DE CLASSIFICAÇÃO
129505	Kittenball's chicken crescent roll casserole	5,0
132493	Anthony's favorite stuffed shells	5,0
174036	Bev's sauteed yellow squash	5,0
168906	Chinese barbecue pork with garlic sauce	5,0
159203	Oatmeal raisin walnut chocolate chip cookies	5,0
147765	Christmas rum balls or bourbon balls	5,0
279	Savory crescent chicken squares	5,0
10566	Fresh blueberry cobbler	5,0
114478	The best moist sweet cornbread	5,0

**Tabela 5.9** – Recomendações para o active user 309 utilizando K=20.

ID	TÍTULO	PREVISÃO DE CLASSIFICAÇÃO
14230	Big batch spaghetti sauce	5,0
129505	Kittenball's chicken crescent roll casserole	5,0
89385	Southwest spaghetti squash	5,0
67106	Delicious fried chicken breast	5,0
51841	Best blueberry crisp	5,0
114480	No bake granola bars	5,0
4091	Thai grilled chicken thighs	5,0
111119	Carla's basic meatballs	5,0
81738	White trash candy snack	5,0

**Tabela 5.10** – Recomendações para o active user 444 utilizando K=20.

ID	TÍTULO	PREVISÃO DE CLASSIFICAÇÃO
88742	Everyday vinaigrette with lots of variations	5,0
56077	Pete's scratch pancakes	5,0
101819	Jo mama's world-famous spaghetti	5,0
2949	Oreo balls	5,0
28552	Crock pot whole chicken	4,5
117899	Creamy Cajun chicken pasta	4,0
155867	Ultimate pizza sauce	4,0

**Tabela 5.11** – Recomendações para o *active user* 39 utilizando K=30.

ID	TÍTULO	PREVISÃO DE CLASSIFICAÇÃO
11324	Best grilled pork chops	5,0
121075	Wendys frosty copycat 4 ww points	5,0
54644	Uncle bills deep fried zucchini strips	5,0
134452	Treebeards red beans and rice	5,0
44551	Tortured chicken beer can	5,0
99885	Cooked chicken for recipes barefoot Contessa style	5,0
3813	Inside out chocolate strawberries	5,0
6682	Boursin cheese homemade	5,0
155664	P. F. Chang's chicken lettuce wraps by Todd Wilbur	5,0

**Tabela 5.12** – Recomendações para o *active user* 309 utilizando K=30.

ID	TÍTULO	PREVISÃO DE CLASSIFICAÇÃO
28469	Parmesan crusted tilapia	5,0
45824	Ednas apple crumble A.K.A apple crisp	5,0
6486	Panera broccoli cheese soup	5,0
99885	Cooked chicken for recipes barefoot Contessa style	5,0
44551	Tortured chicken beer can	5,0
29746	Roasted cauliflower 16 roasted cloves of garlic	5,0
29638	Wholesome homemade honey whole wheat bread	5,0
6682	Boursin cheese homemade	5,0
61366	So, its bekahs pulled pork now	5,0

**Tabela 5.13** – Recomendações para o *active user* 444 utilizando K=30.

ID	TÍTULO	PREVISÃO DE CLASSIFICAÇÃO
10393	Vegetarian meatballs	5,0
147374	Kittenball's Italian melt in your mouth meatballs	5,0
32114	Kittenball's easy and delicious ranch parmesan chicken	5,0
99787	To die for crock pot roast	5,0
145359	Cauliflower cheddar soup	5,0
55033	Maple dijon brussels sprouts	5,0
88742	Everyday vinaigrette with lots of variations	5,0
15173	My family's favorite sloppy joes' pizza joes	5,0
81126	Kittenball's brown sugar caramel sauce or ice cream topping	5,0

Dos resultados obtidos, podem-se destacar duas ocorrências. Primeiro, a grande maioria das previsões de classificação são de 5 valores. Como explicado anteriormente,

a previsão é feita através da média de classificações atribuída pela vizinhança, ou seja, as receitas recomendadas foram avaliadas com a classificação máxima.

A segunda observação diz respeito ao utilizador com ID 444. Como observado na tabela 5.2 utilizando a recomendação baseada em conteúdo não foi possível identificar as preferências do *active user*, o que levou a não existirem recomendações. No entanto, utilizando a segunda abordagem, já foi possível recomendar receitas, para além disso, a quantidade de recomendações aumentou conforme o tamanho da vizinhança. O que leva a concluir que embora não seja possível identificar as *tags* preferidas do *active user*, é possível identificar outros utilizadores que partilham das mesmas preferências e a partir daí selecionar as receitas avaliadas apenas pela vizinhança e gerar recomendações.



## CAPITULO 6 – CONCLUSÕES



## 6. Conclusões

Neste capítulo serão apresentadas as conclusões retiradas do protótipo desenvolvido. Será iniciado pela secção 6.1. onde se fará um levantamento do trabalho desenvolvido até a data, quais os objetivos propostos e quais os realizados, dos que não foram concretizados qual a razão que levou a tal, e por fim a análise dos resultados obtidos através dos casos de estudo. O capítulo terminará com a secção 6.2 onde serão apresentadas melhorias para o presente trabalho bem como novas funcionalidades a desenvolver que o complementem.

### 6.1. Síntese e conclusão

A presente dissertação descreve o desenvolvimento de um protótipo responsável por identificar preferências alimentares de pacientes com DM tipo II. Sendo a DM tipo II uma doença crónica em que a uma das componentes do seu tratamento é a terapia nutricional. Esta foca-se em promover uma escolha saudável de alimentos, que não só são benéficas para o controlo metabólico do paciente como também respeitam a preferências alimentares do mesmo, para atingir este objetivo a terapia nutricional não só recorre ao desenvolvimento de planos nutricionais, com também à educação dos pacientes de forma a explicar como podem alcançar uma alimentação que promova a boa gestão da doença. É interessante aliar estes objetivos com tecnologias de recomendação, pois para além de serem empregadas em contextos em que a recomendação tem de ir ao encontro das preferências dos utilizadores, quer pessoais quer culturais, também tem a versatilidade de respeitar as suas necessidades, por exemplo restrições médicas.

Embora o foque da gestão da DM tipo II se centre em evitar alimentos que provoquem uma resposta negativa ao controlo glicémico, pressão sanguínea e perfil lipídico, o presente trabalho foca-se na identificação de preferências, para depois, posteriormente em trabalhos futuros, identificar uma relação entre as receitas sugeridas e a resposta fisiológica que provocará ao indivíduo, ponto que será abordado na secção seguinte.

Para concretizar este objetivo foram avaliadas várias abordagens de recomendação de receitas direcionadas para indivíduos com DM tipo II. Após uma análise bibliográfica sobre os cenários em que são empregues e as metodologias utilizadas para o desenvolvimento de protótipos, foi concluído que a utilização das abordagens de

recomendação baseada em conteúdo e por filtragem colaborativa são as que melhor se alinhavam com os objetivos do projeto.

A literatura avaliada até ao momento tem demonstrado uma preferência por aliar as duas abordagens no que toca à identificação de preferências, e com o presente artigo, é pretendido o mesmo. A filtragem colaborativa é uma abordagem que identifica utilizadores com perfis semelhantes ao *active user*, quer isto dizer que partilham um certo número de receitas avaliadas, identificam-se as receitas que ainda não foram classificadas pelo mesmo e prevê-se que classificação este atribuirá. As que tem maior previsão de classificação são recomendadas. A recomendação baseada em conteúdo identifica o histórico de classificações, que receitas já classificou utilizador? Quais as que tem classificações mais altas? Que *tags* lhes estão associadas? E a partir desta informação a abordagem procura por receitas semelhantes às já classificadas, recomendando as com maior similaridade.

O que torna vantajoso aliar as duas abordagens é o facto de permitir recomendar receitas que vão ao encontro do histórico de preferências do *active user* e conjugam com receitas que ainda não foram classificadas pelo mesmo, mas que já se demonstraram apreciadas por utilizadores que partilham das mesmas preferências que o *active user*.

Através dos casos de estudo foi possível perceber que a recomendação baseada em conteúdo tem algumas limitações, por exemplo no caso de estudo com o utilizador com ID 444 que não recebeu recomendações a partir desta abordagem, mesmo este tendo uma quantidade substancial de classificações prévias. Por outro lado, através da utilização da filtragem colaborativa foi possível identificar receitas que serão do interesse do mesmo. Ou seja, a filtragem colaborativa irá complementar as limitações da recomendação baseada em conteúdo. Para além disso, é possível observar nos resultados que as duas abordagens não recomendam os mesmos itens, o que indica que as recomendações serão abrangentes e variadas.

Por fim no que toca ao tamanho da vizinhança do *active user*, foi identificado com os melhores resultados  $K=30$ . Ou seja, determinar quais os 30 utilizadores mais semelhantes ao *active user*. Isto porque em situações como a do *active user* com ID 444, onde não é possível identificar as *tags* que este prefere e as suas preferências são um pouco distantes das dos restantes, a utilização de uma vizinhança composta por 30 utilizadores foi o único cenário que permitiu a sugestão de 9 receitas.

Quanto à performance das abordagens foi perceptível uma necessidade maior de capacidade computacional por parte da filtragem colaborativa. Isto porque é necessário calcular a similaridade do *active user* com todos os utilizadores do sistema e ordená-los segundo a ordem decrescente de similaridade, o que leva ao aumento tempo de sugestão. Para além, com o aumento de utilizadores o tempo de recomendação também irá aumentar. A recomendação baseada em conteúdo mostrou ser mais rápida, no entanto, não assegura que todos os utilizadores recebam recomendações.

No início do projeto foram propostos objetivos, no entanto, devido a razões relacionadas com limitações de tempo e falta de dados, não foi possível desenvolver todos os pontos propostos. Na tabela 6.1 estão identificados os objetivos cumpridos e não cumpridos.

**Tabela 6.1** – Listagem dos objetivos propostos.

<b>OBJETIVO PROPOSTO</b>	<b>CUMPRIDO/NÃO CUMPRIDO</b>
<b>Desenvolver um algoritmo capaz de analisar e perceber os gostos do utilizador e realizar sugestões de receitas que vão ao encontro dos mesmos.</b>	Cumprido
<b>Fazer a filtragem de receitas que possam provocar uma má gestão do controlo glicémico, aumento do colesterol, pressão arterial ou do peso do indivíduo.</b>	Não cumprido
<b>Combinar as abordagens selecionadas e analisar os benefícios e desvantagens que os algoritmos híbridos trazem à recomendação de receitas.</b>	Não cumprido

## 6.2. Trabalhos futuros

Neste trabalho foi possível identificar os gostos dos utilizador, no entanto, como mencionado anteriormente, o objetivo principal é contruir um sistema que possa ter um impacto direto na gestão da diabetes. Como tal, para trabalhos futuros, as seguintes funcionalidades são propostas:

- Combinar os algoritmos desenvolvidos para criar uma abordagem híbrida de recomendação de receitas.
- Filtrar as recomendações com base no impacto que terão sobre o controlo metabólico dos pacientes com diabetes tipo II.

- Permitir aos utilizadores registar os valores médicos, como índice glicémico, pressão arterial e valores lipídicos,
- Permitir aos utilizadores registar o histórico de refeições.
- Permitir aos utilizadores registar o histórico de exercício físico.

## Referências

- [1] J. Godinho, S. Batista, D. Martinho, and L. Conceição, “A Recommendation System of Nutrition and Physical Activity for Patients with Type 2 Diabetes Mellitus,” in *Lecture Notes in Computer Science*, 2020, p. 11.
- [2] Associação Portuguesa de Administradores Hospitalares and MOAI, “Um PRR para a diabetes: a oportunidade é agora.” p. 69, 2022.
- [3] “Retrato da diabetes em Portugal: um debate necessário,” *Público*, 2021.
- [4] J. C. N. Ms, K. C. Jordan, J. Metos, and J. F. Hurdle, “Nutrition Informatics Applications in Clinical Practice : a Systematic Review Division of Nutrition,” no. November 2015, pp. 963–972, 2016.
- [5] S. Abhari *et al.*, “A Systematic Review of Nutrition Recommendation Systems: With Focus on Technical Aspects,” no. September, pp. 591–602, 2019.
- [6] J. F. Ndisang, A. Vannacci, and S. Rastogi, “Insulin Resistance, Type 1 and Type 2 Diabetes, and Related Complications 2017,” *J. Diabetes Res.*, vol. 2017, pp. 10–12, 2017.
- [7] K. Gunnars, “Insulin and Insulin Resistance — The Ultimate Guide,” 2019. [Online]. Available: <https://www.healthline.com/nutrition/insulin-and-insulin-resistance#basics>. [Accessed: 15-Oct-2022].
- [8] Diabetes UK, “Differences between type 1 and type 2 diabetes.” [Online]. Available: [https://www.diabetes.org.uk/diabetes-the-basics/differences-between-type-1-and-type-2-diabetes#:~:text=The main difference between the,producing cells in your pancreas](https://www.diabetes.org.uk/diabetes-the-basics/differences-between-type-1-and-type-2-diabetes#:~:text=The main difference between the,producing cells in your pancreas.). [Accessed: 16-Oct-2022].
- [9] NHS, “What is type 1 diabetes?,” 2021. [Online]. Available: <https://www.nhs.uk/conditions/type-1-diabetes/about-type-1-diabetes/what-is-type-1-diabetes/>. [Accessed: 16-Oct-2022].
- [10] NHS, “Overview -Gestational diabetes,” 2019. [Online]. Available: <https://www.nhs.uk/conditions/gestational-diabetes/%0A>. [Accessed: 16-Oct-2022].
- [11] American Diabetes Association, “The path to understanding diabetes starts here.”

- [Online]. Available: <https://diabetes.org/diabetes>. [Accessed: 16-Oct-2022].
- [12] Diabetes UK, “What is a hypo?” [Online]. Available: <https://www.diabetes.org.uk/guide-to-diabetes/complications/hypos>. [Accessed: 16-Oct-2022].
- [13] J. Diabetes, “Joslin Diabetes Center ’ s Clinical Guidelines for Management of Adults with Diabetes,” 2020.
- [14] Diabetes UK, “Diabetes treatments.” [Online]. Available: <https://www.diabetes.org.uk/diabetes-the-basics/diabetes-treatments>. [Accessed: 16-Oct-2022].
- [15] I. Faiz, H. Mukhtar, and S. Khan, “An integrated approach of diet and exercise recommendations for diabetes patients,” *2014 IEEE 16th Int. Conf. e-Health Networking, Appl. Serv. Heal. 2014*, no. January 2015, pp. 537–542, 2014.
- [16] D. Care and S. S. Suppl, “Lifestyle management: Standards of medical care in diabetes 2019,” *Diabetes Care*, vol. 42, no. January, pp. S46–S60, 2019.
- [17] C. Seery, “Blood Sugar Level Ranges,” 2022. [Online]. Available: [https://www.diabetes.co.uk/diabetes\\_care/blood-sugar-level-ranges.html](https://www.diabetes.co.uk/diabetes_care/blood-sugar-level-ranges.html). [Accessed: 10-Jun-2022].
- [18] J. Xie and Q. Wang, “A personalized diet and exercise recommender system in minimizing clinical risk for type 1 diabetes: An in silico study,” *ASME 2017 Dyn. Syst. Control Conf. DSCC 2017*, vol. 1, no. October 2017, 2017.
- [19] “Nutrition recommendations and interventions for diabetes: A position statement of the American Diabetes Association,” *Diabetes Care*, vol. 31, no. SUPPL. 1, 2008.
- [20] M. J. Davies *et al.*, “Management of hyperglycemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the european association for the study of diabetes (EASD),” *Diabetes Care*, vol. 41, no. 12, pp. 2669–2701, 2018.
- [21] A. B. Evert *et al.*, “Nutrition therapy recommendations for the management of adults with diabetes,” *Diabetes Care*, vol. 37, no. SUPPL.1, pp. 120–143, 2014.
- [22] K. Briggs Early and K. Stanley, “Position of the Academy of Nutrition and

- Dietetics: The Role of Medical Nutrition Therapy and Registered Dietitian Nutritionists in the Prevention and Treatment of Prediabetes and Type 2 Diabetes,” *J. Acad. Nutr. Diet.*, vol. 118, no. 2, pp. 343–353, 2018.
- [23] F. D. Diabetes UK Know Diabetes, “Tablets and Medication - Diabetes and Medications.” [Online]. Available: <https://www.diabetes.org.uk/guide-to-diabetes/managing-your-diabetes/treating-your-diabetes/tablets-and-medication>. [Accessed: 09-Nov-2022].
- [24] M. J. Franz, “Diabetes Nutrition Therapy: Effectiveness, Macronutrients, Eating Patterns and Weight Management,” *Am. J. Med. Sci.*, vol. 351, no. 4, pp. 374–379, 2016.
- [25] C. Kamada, “Optimal Carbohydrate and Nutrient Intake for Japanese Elderly Patients with Type 2 Diabetes,” *Glucose Intake Util. Pre-Diabetes Diabetes Implic. Cardiovasc. Dis.*, pp. 315–325, 2015.
- [26] S. X. H. Mthembu *et al.*, “Impact of physical exercise and caloric restriction in patients with type 2 diabetes: Skeletal muscle insulin resistance and mitochondrial dysfunction as ideal therapeutic targets,” *Life Sci.*, vol. 297, no. February, p. 120467, 2022.
- [27] Physical Activity Guidelines Advisory Committee, “Physical Activity Guidelines Advisory Committee Scientific Report,” *Dep. Heal. Hum. Serv.*, p. 779, 2018.
- [28] I. Lingvay, P. Sumithran, R. V. Cohen, and C. W. le Roux, “Obesity management as a primary treatment goal for type 2 diabetes: time to reframe the conversation,” *Lancet*, vol. 399, no. 10322, pp. 394–405, 2022.
- [29] M. J. Franz, J. L. Boucher, S. Rutten-Ramos, and J. J. VanWormer, “Lifestyle Weight-Loss Intervention Outcomes in Overweight and Obese Adults with Type 2 Diabetes: A Systematic Review and Meta-Analysis of Randomized Clinical Trials,” *J. Acad. Nutr. Diet.*, vol. 115, no. 9, pp. 1447–1463, 2015.
- [30] A. Williams, J. Radford, J. O. Brien, and K. Davison, “Type 2 diabetes and the medicine of exercise,” *Aust. J. Gen. Pract.*, vol. 49, no. 4, pp. 189–193, 2020.
- [31] S. Kulkarni and S. F. Rodd, “Context Aware Recommendation Systems: A review of the state of the art techniques,” *Comput. Sci. Rev.*, vol. 37, p. 100255, 2020.
- [32] B. G. Patra *et al.*, “A content-based literature recommendation system for datasets

- to improve data reusability – A case study on Gene Expression Omnibus (GEO) datasets,” *J. Biomed. Inform.*, vol. 104, no. May 2019, p. 103399, 2020.
- [33] J. Huttner, “From Tapestry to SVD: A Survey of the Algorithms That Power Recommender Systems,” no. May 2009, p. 32, 2009.
- [34] R. Sharma and R. Singh, “Evolution of recommender systems from ancient times to modern era: A survey,” *Indian J. Sci. Technol.*, vol. 9, no. 20, 2016.
- [35] J. Beel, B. Gipp, S. Langer, and C. Breiting, “Research-paper recommender systems: a literature survey,” *Int. J. Digit. Libr.*, vol. 17, no. 4, pp. 305–338, 2016.
- [36] B. B. Sinha and R. Dhanalakshmi, “Evolution of recommender system over the time,” *Soft Comput.*, vol. 23, no. 23, pp. 12169–12188, 2019.
- [37] R. Sharma, D. Gopalani, and Y. Meena, “Collaborative filtering-based recommender system: Approaches and research challenges,” *3rd IEEE Int. Conf.*, pp. 1–6, 2017.
- [38] N. M. Ariff, M. A. A. Bakar, and N. F. Rahim, “Comparison between content-based and collaborative filtering recommendation system for movie suggestions,” *AIP Conf. Proc.*, vol. 2013, 2018.
- [39] M. J. Pazzani and D. Billsus, *Content-Based Recommender Systems*. 2007.
- [40] P. Lops, M. De Gemmis, and G. Semeraro, *Content-Based Recommender Systems: State of art and trends*, no. January. 2011.
- [41] J. Leskovec, *9.1 A Model for Recommendation Systems*, 2nd ed. Cambridge University Press, 2014.
- [42] G. Adomavicius and A. Tuzhilin, “Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions,” *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, 2005.
- [43] S. Jain, A. Grover, P. S. Thakur, and S. K. Choudhary, “Trends, problems and solutions of recommender system,” *Int. Conf. Comput. Commun. Autom. ICCCA 2015*, no. May, pp. 955–958, 2015.
- [44] R. Systems, “Collaborative Filtering.” [Online]. Available: <http://recommender-systems.org/collaborative-filtering/>. [Accessed: 10-Nov-2022].
- [45] M. Y. H. Al-Shamri, “User profiling approaches for demographic recommender

- systems,” *Knowledge-Based Syst.*, vol. 100, pp. 175–187, 2016.
- [46] C. C. Aggarwal, *Knowledge-Based Recommender Systems*, no. August. 2016.
- [47] B. Smyth, *Case-Based Recommendation*, vol. 4321, no. January 2007. 2007.
- [48] J. Wu, “Knowledge-Based Recommender Systems: An Overview,” 2019. .
- [49] A. Kumar, P. Tanwar, and S. Nigam, “Survey and evaluation of food recommendation systems and techniques,” *Proc. 10th INDIACom; 2016 3rd Int. Conf. Comput. Sustain. Glob. Dev. INDIACom 2016*, pp. 3592–3596, 2016.
- [50] J. E. Aronson, “Expert Systems,” in *Encyclopedia of Information Systems*, 2003, pp. 277–289.
- [51] R. Burke, “Hybrid Web Recommender Systems,” *Adapt. Web*, pp. 377–408, 2007.
- [52] J. Chiang, “7 Types of Hybrid Recommendation System,” 2021. [Online]. Available: <https://medium.com/analytics-vidhya/7-types-of-hybrid-recommendation-system-3e4f78266ad8>. [Accessed: 14-Nov-2022].
- [53] T. N. Trang Tran, M. Atas, A. Felfernig, and M. Stettinger, “An overview of recommender systems in the healthy food domain,” *J. Intell. Inf. Syst.*, vol. 50, no. 3, pp. 501–526, 2018.
- [54] S. Norouzi *et al.*, “A mobile application for managing diabetic patients’ nutrition: A food recommender system,” *Arch. Iran. Med.*, vol. 21, no. 10, pp. 466–472, 2018.
- [55] M. Phanich, P. Pholkul, and S. Phimoltares, “Food recommendation system using clustering analysis for diabetic patients,” *2010 Int. Conf. Inf. Sci. Appl. ICISA 2010*, 2010.
- [56] A. A. Aljumah, M. G. Ahamad, and M. K. Siddiqui, “Application of data mining: Diabetes health care in young and old patients,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 25, no. 2, pp. 127–136, 2013.
- [57] S. I. Ali, M. B. Amin, and S. Kim, *A Hybrid Framework for a Comprehensive Physical Activity and Diet Recommendation System*. Springer International Publishing, 2018.
- [58] JetBrains, “What is IntelliJ IDEA?” [Online]. Available: <https://www.jetbrains.com/idea/features/>. [Accessed: 17-Nov-2022].

- [59] S. Li and B. P. Majumder, “Food.com Recipes and Interactions,” 2019. [Online]. Available: [https://www.kaggle.com/datasets/shuyangli94/food-com-recipes-and-user-interactions?select=ingr\\_map.pkl](https://www.kaggle.com/datasets/shuyangli94/food-com-recipes-and-user-interactions?select=ingr_map.pkl). [Accessed: 12-Nov-2022].
- [60] I. L. Antunes, “Utilização da hemoglobina glicada no diagnostico da diabetes mellitus,” 2012.
- [61] C. C. medical Professional, “Glycated Hemoglobin Test (A1c),” 2021. [Online]. Available: <https://my.clevelandclinic.org/health/diagnostics/9731-glycated-hemoglobin-test-a1c#:~:text=What’s a glycated hemoglobin test,diabetes%2C heart disease and stroke>. [Accessed: 01-Oct-2022].