



Sistemas de recomendação para intervenções do consumo de álcool e tabaco utilizando técnicas de agrupamento

Thiago Costa Rizuti da Rocha

JUIZ DE FORA
DEZEMBRO, 2016

Sistemas de recomendação para intervenções do consumo de álcool e tabaco utilizando técnicas de agrupamento

THIAGO COSTA RIZUTI DA ROCHA

Universidade Federal de Juiz de Fora
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Bacharelado em Ciência da Computação

Orientador: Heder Soares Bernardino

JUIZ DE FORA
DEZEMBRO, 2016

SISTEMAS DE RECOMENDAÇÃO PARA INTERVENÇÕES DO
CONSUMO DE ÁLCOOL E TABACO UTILIZANDO TÉCNICAS
DE AGRUPAMENTO

Thiago Costa Rizuti da Rocha

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTEGRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Heder Soares Bernardino
Doutor em Modelagem Computacional

Jairo Francisco de Souza
Doutor em Informática

Victor Ströele de Andrade Menezes
Doutor em Engenharia de Sistemas e Computação

JUIZ DE FORA
15 DE DEZEMBRO, 2016

Aos meus pais, familiares e amigos.

Resumo

Sistemas de recomendação podem ser encontrados em diversas aplicações e serviços disponíveis na web, permitindo que a informação seja selecionada e exibida de acordo com o perfil do usuário. Usualmente são recomendados filmes, livros, músicas e produtos. Um dos diferenciais dos sistemas de *e-health*, como as intervenções destinadas às mudanças comportamentais, é a geração de conteúdo personalizado. Desta forma, sistemas de recomendação podem ser uma excelente solução para este fim. Neste trabalho foram desenvolvidos sistemas de recomendação colaborativos para uma intervenção para a cessação do consumo de tabaco e para uma intervenção para a redução do consumo de álcool. Foi proposta uma abordagem com a utilização de algoritmos de agrupamento como forma de definir vizinhanças de usuários similares e contornar o problema da partida fria. As recomendações geradas foram analisadas através de avaliações internas com uso de uma base de dados e métricas de qualidade.

Palavras-chave: sistema de recomendação; filtragem colaborativa; aprendizagem de máquina; mineração de dados; agrupamento; e-health; intervenção; álcool; tabaco

Abstract

Recommender systems can be seen in many systems and services available on the web, allowing information to be selected and displayed according to the user's profile. Usually, movies, books, music and products are recommended. One of the advantages of e-health systems, as interventions intended to cause behavioral change, is the availability of custom content. Thus, recommender systems can be an excellent solution to this purpose. In this work, we developed collaborative recommender systems for a smoking quit intervention and for a drinking cut down intervention. We proposed the use of clustering algorithms as a way of defining neighborhoods of similar users and to avoid the cold-start problem. The recommendations were analysed through internal evaluations with the support of a test data set and quality metrics.

Keywords: recommender systems; collaborative filtering; machine learning; data mining; clustering; e-health; intervention; alcohol; tobacco; drinking; smoking

Agradecimentos

Primeiramente, agradeço a meus pais pelo sustento e apoio incondicional . A meu irmão Túlio, pelo exemplo e incentivo a área científica e acadêmica. E a Camila, minha companheira e amiga, que faz-se sempre presente em meus dias.

Agradeço a Heder Soares Bernardino pela orientação, dedicação, paciência, e amizade. Sem o qual este trabalho não seria possível.

Agradecimentos especiais a Henrique Gomide, Leonardo Martins, Telmo Ronzani e a todos os pesquisadores do CREPEIA pela oportunidade e suporte.

Aos meus amigos e colegas que tornaram os dias de estudo e trabalho mais agradáveis. E finalmente, a todos que, direta ou indiretamente, me acompanharam e contribuíram para a conclusão deste trabalho.

“O Deep Thought computer, the task we have designed you to perform is this. We want you to tell us The Answer .

...

Forty-two - said Deep Thought - with infinite majesty and calm ”.

Douglas Adams (The Hitchhiker’s Guide to the Galaxy)

Conteúdo

Lista de Figuras	8
Lista de Tabelas	9
Lista de Abreviações	10
1 Introdução	11
1.1 Apresentação do Tema	11
1.2 Problema	11
1.3 Justificativa e Motivação	12
1.4 Hipóteses	13
1.5 Objetivos	13
1.6 Estrutura do trabalho	14
2 Revisão Bibliográfica	15
2.1 Sistemas de Recomendação	15
2.2 Aprendizagem de Máquina e Mineração de Dados	17
2.3 Agrupamento	18
2.4 Intervenções computadorizadas	19
2.5 Intervenções customizadas e sistemas de recomendação	20
3 Intervenções Para Tabaco e Álcool	22
3.1 Viva Sem Tabaco	22
3.2 Álcool e Saúde	24
3.3 Avaliação das páginas informativas	26
4 Métodos	28
4.1 Agrupamento de Usuários	28
4.1.1 Caracterização dos usuários	28
4.1.2 Métricas de similaridade	29
4.1.3 Algoritmos de agrupamento	30
4.1.4 Métrica de qualidade	32
4.2 Sistemas de Recomendação	32
4.2.1 Itens recomendados	32
4.2.2 Cálculo das predições e recomendações	33
4.2.3 Avaliação das recomendações	33
4.3 Validação Cruzada	34
5 Experimentos	36
5.1 Conjunto de dados utilizado	36
5.1.1 Validação cruzada estratificada	38
5.2 Experimento 1	38
5.3 Experimento 2	39
5.3.1 Viva Sem Tabaco	40
5.3.2 Álcool e Saúde	41

5.4	Experimento 3	43
5.4.1	Viva Sem Tabaco	43
5.4.2	Álcool e Saúde	45
5.5	Melhores Resultados	46
6	Conclusões	48
6.1	Considerações Finais	48
6.2	Trabalhos Futuros	49
	Bibliografia	50

Lista de Figuras

3.1	Tela principal do Viva Sem Tabaco.	23
3.2	Exemplo de um formulário durante a etapa “Quero parar”.	24
3.3	Tela principal do Álcool e Saúde.	25
3.4	Exemplo de um formulário durante a etapa “Avalie como você bebe”.	26
3.5	Botões de avaliação em uma página do Viva Sem Tabaco.	27
3.6	Botões de avaliação em uma página do Álcool e Saúde.	27
5.1	Distribuição de avaliações positivas e negativas para o VST.	37
5.2	Distribuição de avaliações positivas e negativas para o AES.	37

Lista de Tabelas

4.1	Atributos para os usuários do VST.	28
4.2	Atributos para os usuários do AES.	29
4.3	Matriz de confusão para o caso de recomendação	34
5.1	Resultados do experimento 1 para o VST - coeficiente de silhueta.	38
5.2	Resultados do experimento 1 para o AES - coeficiente de silhueta.	38
5.3	Resultados do experimento 2 para o VST - <i>precision</i>	40
5.4	Resultados do experimento 2 para o VST - <i>recall</i>	41
5.5	Resultados do experimento 2 para o VST - <i>f-measure</i>	41
5.6	Resultados do experimento 2 para o AES - <i>precision</i>	42
5.7	Resultados do experimento 2 para o AES - <i>recall</i>	42
5.8	Resultados do experimento 2 para o AES - <i>f-measure</i>	42
5.9	Resultados do experimento 3 para o VST - <i>precision</i>	44
5.10	Resultados do experimento 3 para o VST - <i>recall</i>	44
5.11	Resultados do experimento 3 para o VST - <i>f-measure</i>	45
5.12	Resultados do experimento 3 para o AES - <i>precision</i>	45
5.13	Resultados do experimento 3 para o AES - <i>recall</i>	46
5.14	Resultados do experimento 3 para o AES - <i>f-measure</i>	46
5.15	Melhores resultados encontrados para ambas as intervenções.	47

Lista de Abreviações

DCC Departamento de Ciência da Computação

UFJF Universidade Federal de Juiz de Fora

SR Sistema de recomendação

VST Viva Sem Tabaco

AES Álcool e Saúde

1 Introdução

1.1 Apresentação do Tema

Sistemas de recomendação podem ser definidos como programas com o objetivo de inferir as preferências e necessidades do usuário e indicar os itens mais adequados de acordo com seu perfil. Tais recomendações são feitas analisando-se dados dos usuários, dos itens e as relações entre ambos (LU et al., 2015). Porém recomendações arbitrárias não são interessantes e o principal desafio é recomendar conteúdo que realmente seja do útil ao usuário. Diferentes formas de se obter informação, técnicas e algoritmos podem ser explorados para se obter maior acurácia e qualidade nas recomendações (BOBADILLA et al., 2013).

Atualmente, os sistemas de recomendação são amplamente utilizados para filmes, músicas, livros, *e-commerce*, *e-learning*, mídias sociais e buscadores web. Porém, eles também podem ser estendidos a qualquer contexto em que se busque personalização de conteúdo (LU et al., 2015).

1.2 Problema

Tratamentos e intervenções assistidas por computador têm a vantagem de reduzir barreiras de tempo, custo, disponibilidade de profissionais e anonimidade, porém geralmente fornecem conteúdo padronizado. Sistemas de recomendação se mostram como uma excelente solução para gerar conteúdo adaptado a essas intervenções. À medida que os usuários interagem com o sistema e a quantidade de dados coletados aumenta, a personalização pode se tornar mais adequada e sofisticada (LIEBERMAN et al., 2008).

Desenvolvidas juntamente com o Centro de Referência em Pesquisa, Intervenção e Avaliação em Álcool e Outras Drogas (CREPEIA) da Universidade Federal de Juiz de Fora, as intervenções on-line **Viva Sem Tabaco**¹ e **Álcool e Saúde**² oferecem conteúdo gratuito e personalizado a pessoas interessadas em tratamento para a cessação do con-

¹<http://www.vivasemtabaco.com.br>

²<http://www.alcoolesaude.com.br>

sumo de tabaco e redução do consumo de álcool. O uso de sistemas de recomendação pode ser uma forma de aprimorar a personalização de ambas intervenções. A fim de aprimorar a personalização de ambas as intervenções, propõe-se aqui desenvolver e sistemas de recomendação para direcionar aos seus usuários informações relacionadas ao tratamento, saúde e qualidade de vida.

1.3 Justificativa e Motivação

Sistemas de recomendação é um tópico recente, suas aplicações são cada vez mais visíveis em diversos contextos e ainda há muito sendo explorado. Na literatura, encontram-se diferentes propostas para o problema de recomendação, entre elas a aplicação de diferentes técnicas de recuperação da informação, aprendizagem de máquinas e análise de redes complexas.

Entre as características do estudo de sistemas de recomendação, uma das mais evidentes é a alta aplicabilidade e multidisciplinaridade. Sistemas voltados para tratamentos médicos e de saúde são cada vez mais comuns. Uma das características mais positivas do *e-health* é a personalização de conteúdo. Sistemas de recomendação estão diretamente ligados a este assunto e podem ser utilizados para se obter conteúdos específicos a cada usuário individualmente.

É importante considerar que usuário que recebe sugestões de conteúdos que não são de seu interesse pode se desmotivar e até mesmo abandonar o tratamento. Assim, mesmo que a quantidade de conteúdos diferentes não seja grande, identificar e escolher aqueles que de fato são relevantes para o usuário é de suma importância.

Desta forma, este trabalho mostra-se em duas faces. Na primeira, do ponto de vista técnico, visa estudar diferentes técnicas para a implementação de sistemas de recomendação e propor uma abordagem que gere recomendações de qualidade. Na segunda, do ponto de vista multidisciplinar, propõe a aplicação em um contexto relacionado à saúde coletiva, no qual os usuários podem se beneficiar diretamente das recomendações geradas pelo sistema.

1.4 Hipóteses

A primeira hipótese a ser verificada é se a qualidade das recomendações geradas pelo sistema de recomendação é melhor do que recomendações escolhidas aleatoriamente.

Utilizando técnicas de agrupamento é possível extrair conhecimento da base dados e descobrir padrões em grupos de usuários similares (XU; WUNSCH, 2005). Assim, usuários podem receber recomendações geradas a partir de outros usuários agrupados com ele. A partir da análise comparativa de diferentes algoritmos de agrupamento e métricas de distância, a segunda hipótese a ser verificada é se o algoritmo com melhor desempenho na avaliação de agrupamento é aquele que vai gerar as melhores recomendações.

1.5 Objetivos

O objetivo geral é propor a adoção de uma abordagem via agrupamento de usuários em sistemas de recomendação para as intervenções *Viva Sem Tabaco* e *Álcool e Saúde*. Estes devem entregar aos usuários recomendações das páginas de conteúdo informativo e educativo de cada uma das intervenções. Como objetivos específicos propõe-se:

- Disponibilizar intervenções *web* para álcool e tabaco.
- Explorar um modelo de recomendação utilizando avaliações dos usuários e outros dados coletados durante os fluxos de utilização das intervenções;
- Analisar diferentes algoritmos de agrupamento para gerar grupos de usuários, nos quais as recomendações serão feitas;
- Avaliar a qualidade dos agrupamentos gerados;
- Avaliar as recomendações utilizando uma base de testes e métricas de qualidade;
- Investigar a relação entre o desempenho dos algoritmos de agrupamento e a qualidade das recomendações.

1.6 Estrutura do trabalho

O restante do conteúdo do trabalho está dividido em cinco capítulos, a seguir cada um deles é introduzido de forma breve.

O capítulo 2 envolve a revisão bibliográfica dos principais tópicos envolvidos: sistemas de recomendação, aprendizagem de máquina, agrupamento e intervenções customizadas. São apresentados leitor alguns trabalhos existentes e os principais conceitos necessários para a contextualização da leitura.

O capítulo 3 apresenta de forma mais detalhada as intervenções desenvolvidas: *Viva Sem Tabaco e Álcool e Saúde*. São descritas as principais funcionalidades do sistema, assim como suas fundamentações e objetivos.

O capítulo 4 tem como objetivo apresentar a forma como a pesquisa foi conduzida. São definidos os algoritmos, métricas e técnicas usados no trabalho.

No capítulo 5, são descritos os experimentos realizados. Cada um deles é descrito passo a passo, com seus resultados apresentados e discutidos.

Por fim, o capítulo 6 são as considerações finais, sintetizando a proposta e seus resultados de maneira geral. São indicados as principais conclusões e implicações, apontando os pontos fortes e fracos da pesquisa realizada. Também são propostos trabalhos futuros a fim de aprofundar e enriquecer a pesquisa iniciada por este trabalho

2 Revisão Bibliográfica

Neste capítulo, será apresentada a fundamentação teórica e revisão bibliográfica dos assuntos envolvidos neste trabalho. Primeiro será apresentado o conceito de sistemas de recomendação, suas principais aplicações e técnicas presentes na literatura. Em seguida, uma breve introdução sobre aprendizagem de máquina e mineração de dados, com foco nas técnicas de agrupamento. Por fim, serão apresentadas intervenções computadorizadas e como sistemas de recomendação podem ser usados neste contexto.

2.1 Sistemas de Recomendação

Com o crescimento da internet, sistemas de recomendação se tornaram cada vez mais comuns, principalmente em aplicações comerciais. Um dos principais exemplos é a *Amazon.com*, um dos primeiros serviços a adotarem sistemas de recomendação, recomendando produtos baseado em histórico, nas avaliações e visualizações de seus usuários. O uso dessa tecnologia levou a um aumento significativo nas vendas, de tal forma que é possível observar o uso de recomendações na maioria dos serviços de *e-commerce* existentes (EKSTRAND; RIEDL; KONSTAN, 2010).

Outro marco importante foi o Prêmio *Netflix* em 2006. A empresa ofereceu o prêmio de \$1.000.000 para o pesquisador ou grupo que desenvolvesse um algoritmo de recomendação que mostrasse desempenho melhor que o *Cinematch*, o algoritmo que era utilizado até então (BENNETT; LANNING, 2007).

A tarefa dos sistemas de recomendação é calcular recomendações de itens a seus usuários à partir da informação coletada sobre suas preferências. A coleta de dados pode ser feita explicitamente, através de formulários e avaliações do usuário, ou implicitamente, observando o seu comportamento (BOBADILLA et al., 2013).

Um dos primeiros trabalhos na área de recomendações foi o *GroupLens*, a proposta de uma arquitetura aberta para recomendações de notícias (RESNICK et al., 1994). Este foi um dos primeiros a introduzir a filtragem colaborativa. Esta abordagem utiliza

avaliações dos itens feitas pelos usuários e similaridades entre os usuários podem ser identificadas através de suas avaliações. Os valores de avaliações de itens ainda não avaliados são estimados com base nas avaliações dos usuários similares e os itens com maiores valores preditos são usados nas recomendações. Uma variação desta filtragem colaborativa baseada em usuários é a filtragem colaborativa baseada em itens. Neste caso, cada item é caracterizado pelas avaliações que recebeu de cada usuário, e são os itens que recebem usuários como recomendações, com base nas avaliações de itens similares a ele (EKSTRAND; RIEDL; KONSTAN, 2010).

Outra abordagem muito utilizada em sistemas de recomendação é a filtragem baseada em conteúdo. Esta identifica suas recomendações pelas similaridades entre itens, ou seja, são recomendados aos usuários itens similares aos que ele demonstrou interesse. Tal similaridade é definida a partir das características e descrições de cada um dos itens. Esta técnica originou-se a partir de pesquisas em recuperação da informação (FOLTZ; DUMAIS, 1992).

Essas diferentes formas de recomendação apresentam suas vantagens e desvantagens. O principal problema relacionado à filtragem colaborativa é a esparsidade dos dados e a baixa escalabilidade. Já a filtragem baseada em conteúdo tem sua desvantagem na superespecialização das recomendações (ADOMAVICIUS; TUZHILIN, 2005).

Novos usuários que ainda não interagiram com o sistema não podem receber recomendações, já que não há conhecimento sobre suas preferências. Além disso, novos itens não serão recomendados até que os usuários mostrem interesse sobre eles. Este é chamado de problema da partida fria, o grande desafio na área de sistemas de recomendação. Tanto a filtragem colaborativa quanto a baseada em conteúdo são afetadas pela partida fria (ADOMAVICIUS; TUZHILIN, 2005). Como forma de contornar este problema, outras técnicas têm sido propostas.

A filtragem demográfica (BOBADILLA et al., 2013) utiliza a ideia de que usuários que compartilham informações como idade, gênero, nacionalidade, escolaridade e área profissional podem compartilhar também as mesmas preferências. A filtragem baseada em conhecimento (BOBADILLA et al., 2013) utiliza a inferência sobre as necessidades do usuário a partir de consultas ou raciocínio baseado em casos. Outras informações que

podem ser usadas são aquelas referentes às características sociais (confiabilidade de usuários, seguidores, amigos, *posts* e *tags*) e os que envolvem elementos contextuais (horário, geolocalização, dispositivo utilizado) (BOBADILLA et al., 2013).

Visto que cada uma das abordagens tem suas vantagens e desvantagens, a maioria dos sistemas de recomendação atuais são chamados de híbridos, pois usam a combinação de duas ou mais técnicas na tentativa de contornar os problemas mais comuns e garantir qualidade nas recomendações (BOBADILLA et al., 2013). Ainda assim, a escolha da técnica para implementar um sistema de recomendação é fortemente dependente do contexto da aplicação, sendo que cada caso mostrará ser mais adequado a uma determinada técnica (LU et al., 2015).

2.2 Aprendizagem de Máquina e Mineração de Dados

Aprendizagem de máquina é a área de estudo que dá aos computadores a habilidade de aprender sem serem explicitamente programados (SAMUEL, 1959). Uma distinção frequentemente feita entre os processos de aprendizagem os classificam entre aprendizado por reforço, supervisionado ou não supervisionado (SATHYA; ABRAHAM, 2013).

O aprendizado por reforço é um processo de tentativas e erro. No qual um agente interage com seu meio e recebe penalizações ou recompensas de acordo com suas ações. Iterativamente, o agente busca maximizar e acumular recompensas, resultando em um processo de aprendizagem.

O aprendizado supervisionado utiliza uma amostra de dados, na qual a saída correspondente já é conhecida para cada elemento dessa amostra. Esse conjunto de treino é usado para aprender um modelo que mapeia cada entrada a uma aproximação do seu valor de saída. O modelo aprendido pode ser então usado para inferir as saídas de novas entradas e para expandir o conhecimento adquirido da relação entre os dados de entrada e de saída.

Na aprendizagem não supervisionada tem-se apenas um conjunto de dados de entrada sem algum valor de saída. Sendo assim, o objetivo é aprender uma estrutura, até então desconhecida, sobre os dados.

Apesar da aprendizagem de máquina e da mineração de dados muitas vezes tra-

tarem do mesmo tipo de problemas e até mesmo utilizarem as mesmas técnicas, elas são áreas de estudo com objetivos distintos. A aprendizagem de máquina foca em prever valores esperados com base em atributos de entrada de um conjunto de dados. A mineração de dados, por sua vez, se preocupa em descobrir e entender a relação entre os atributos (GOMIDE et al., 2015).

2.3 Agrupamento

Agrupamento é uma técnica de aprendizado não-supervisionado, usada com o objetivo de descobrir padrões ou correlações ocultas num conjunto de dados. Algoritmos de agrupamento particionam os dados em um certo número de grupos, de forma que objetos de um mesmo grupo devem ser similares e objetos de grupos diferentes devem ser dissimilares. A similaridade é calculada de acordo com uma métrica com base nas características dos objetos a serem agrupados (XU; WUNSCH, 2005).

Existem diversos algoritmos para agrupamento. Cada um utiliza diferentes estratégias para descobrir grupos bem definidos. De tal forma que cada algoritmo pode gerar um agrupamento diferente ou se adaptar melhor a um certo problema devido as características do conjunto de dados. Ainda assim, um mesmo algoritmo com diferentes configurações de parâmetros pode resultar em agrupamentos diferentes. Técnicas de avaliação e métricas de qualidade são importantes para indicar o grau de confiança de um agrupamento. Sendo também um meio de guiar na escolha do melhor algoritmo e do ajuste de parâmetros para um dado problema. Alguns exemplos de algoritmos são o *k-means*, *hierarchical clustering*, *mean shift*, DBSCAN e BIRCH. (XU; WUNSCH, 2005).

Em sistemas de recomendação, algoritmos de agrupamento são usados para melhorar a qualidade das recomendações. Podem ser feitos agrupamentos de usuários, agrupamentos de itens ou um método híbrido entre agrupamentos de usuários e itens (*bi-clustering*) (ZHU; GONG, 2009). Os valores de predição para cada usuário ou item são calculados sobre o grupo em que se encontram. Resultados mostram que o agrupamento de usuários (SHINDE; KULKARNI, 2012) e o agrupamento de itens (YAO; ZHANG, 2009) são boas alternativas para melhorar a escalabilidade e contornar problemas como a esparsidade dos dados e a partida fria.

2.4 Intervenções computadorizadas

O termo *e-health* é definido pela Organização Mundial de Saúde como a utilização de tecnologias da informação e comunicação na saúde. Englobando sistemas que apoiam tratamentos de pacientes, pesquisas, educação em saúde, rastreamento de doenças e monitoramento da saúde pública. (WHO, 2016).

No que se trata de mudanças comportamentais, existem obstáculos na busca de tratamento como a negação do problema, limitação de profissionais, disponibilidade de tempo, barreiras de distância e custo, além do estigma e vergonha com o tratamento em si. Intervenções assistidas por computador mostram-se como uma alternativa para quebrar esses obstáculos. Pacientes podem se sentir mais a vontade e preferir trabalhar com o problema em seu próprio ritmo. Por outro lado, somente é oferecido um conteúdo padronizado, que não considera as particularidades de cada caso. (LIEBERMAN et al., 2008).

As intervenções customizadas são um tipo específico de intervenções computadorizadas. Estas também objetivam disponibilizar conteúdo de forma auto-guiada e interativa. Porém, propõe a personalização do conteúdo de forma que se adapte às necessidades individuais de cada paciente. São comumente implementadas através de *websites* ou aplicativos de *smartphones*, dado a facilidade de coletar uma grande quantidade de dados e de modificar seu conteúdo em tempo real. Os principais componentes identificados em intervenções customizadas são o seu conteúdo; o uso de recursos multimídia; o uso de interatividade; e a oferta de acompanhamentos e *feedback* (GOMIDE et al., 2015).

Atualmente, percebe um grande crescimento no desenvolvimento de intervenções personalizadas. Ao mesmo tempo, cresce a preocupação com conteúdo de qualidade. Se desenvolvidas de forma correta, as intervenções customizadas podem impactar positivamente o tratamento e prevenção de problemas relacionados a saúde. Além de quebrar barreiras, tornam-se mais atrativas e por consequência podem encorajar o uso constante e aumentar o engajamento dos usuários com o tratamento. (LIEBERMAN et al., 2008).

2.5 Intervenções customizadas e sistemas de recomendação

Muitas das intervenções propõem a personalização através de uma abordagem baseada em regras. Estas regras são escritas por especialistas e implementadas através de instruções se-senão. Entre as limitações deste método estão a dificuldade de se trabalhar com muitas variáveis e a impossibilidade de se adaptar em tempo real (SADASIVAM et al., 2016).

Sistemas de recomendação geralmente usados em aplicações comerciais podem ser adaptados às intervenções. Ao invés da promoção de vendas, o objetivo é aumentar o interesse e participação do usuário (LIEBERMAN et al., 2008). Recomendações podem ser feitas na forma de conteúdos adequados a cada tipo de usuário.

Juntamente com técnicas de aprendizagem de máquina, sistemas de recomendação podem aprender e se adaptar de acordo com o *feedback* dos usuários. Quanto mais dados coletados, mais sofisticadas as recomendações podem se tornar (SADASIVAM et al., 2016).

Apesar de se adaptarem bem ao problema da personalização de intervenções, os sistemas de recomendação podem apresentar algumas limitações. Primeiramente, a coleta de dados e o uso contínuo dos usuários é um grande desafio. Quando comparados à *Netflix* e à *Amazon*, por exemplo, os sistemas de intervenções dificilmente terão tantos dados quanto essas empresas conseguem coletar. Além disso, os estudos em intervenções geralmente têm de cumprir as etapas de suas pesquisas em um intervalo de tempo mais curto. Outro fator limitante é a dificuldade em medir a efetividade do sistema. Enquanto a recomendação de um produto pode ser comparado com a decisão de compra, identificar se um determinado conteúdo da intervenção foi efetivo para o tratamento pode ser mais complexo (SADASIVAM et al., 2016).

Apesar da dificuldade nesse tipo de pesquisa, um modelo de usuário para intervenções ao consumo de álcool é proposto em (YASAVUR; AMINI; LISETTI, 2012). Este modelo inclui características comportamentais, contextuais e demográficas, de forma a caracterizar perfis de usuários para serem usados em sistemas de recomendação. Outro exemplo envolve a utilização de sistemas de recomendação em um aplicativo móvel para

a cessação do uso de tabaco (HORS-FRAILE et al., 2016). Este combina dois sistemas de recomendação, um seguindo uma abordagem híbrida (combinando a filtragem baseada em conteúdo, colaborativa e demográfica) para recomendar mensagens educativas e outro seguindo uma abordagem baseada em conteúdo para programar o envio das mensagens.

Embora existam limitações, sistemas de recomendação mostram-se como uma solução interessante para customização de intervenções para mudanças comportamentais. Porém, estes são muito dependentes do contexto da aplicação. Sendo assim, importante a observação dos dados disponíveis sobre usuários e itens a serem recomendados para a escolha da técnica que melhor se adapte ao problema.

3 Intervenções Para Tabaco e Álcool

As intervenções computadorizadas customizadas *Viva Sem Tabaco e Álcool e Saúde* foram desenvolvidas pelo Centro de Referência em Pesquisa, Intervenção e Avaliação em Álcool e outras Drogas (CREPEIA) em parceria com o Departamento de Ciência da Computação da Universidade Federal de Juiz de Fora. Para ambas foram desenvolvidos sistemas web utilizando a linguagem de programação *Java* e o *framework PrimeFaces*. Os sistemas estão licenciados sob a *GNU General Public License* e disponíveis no *GitHub*.

Apesar dos sistemas terem sido construídos sobre a mesma estrutural principal, cada um possui suas particularidades no que se refere a identidade visual, funcionalidades e conteúdo.

3.1 Viva Sem Tabaco

O *Viva Sem Tabaco (VST)* é uma intervenção para o tratamento de tabagismo implementada através de um *website*. Todo o seu conteúdo é baseado em pesquisas científicas, nos protocolos de tratamento do Ministério da Saúde e Instituto Nacional do Câncer e cumpre o Código de Ética da *Health on the Net Foundation*. Sua eficácia está sendo testada através de uma pesquisa clínica. O programa é gratuito e disponibilizado em diferentes línguas. A figura 3.1 mostra a tela principal do sistema.

A intervenção do *Viva Sem Tabaco* é dividida em três etapas principais: “Vale a pena parar”, “Quero parar” e “Já parei”. A primeira etapa consiste na disponibilização de conteúdo educativo e informativo. O objetivo desta etapa é a conscientização dos riscos do consumo de tabaco e das vantagens de parar de fumar. A segunda etapa tem como objetivo fornecer ajuda a quem deseja parar de fumar. Através de etapas menores e interativas, é possível escolher entre estratégias para parar e dicas para enfrentar a abstinência, de forma que no final se tenha um plano de parada completo e personalizado. Na terceira e última etapa é feito o acompanhamento daqueles que já tentaram parar de fumar através da etapa anterior. Com a mesma interatividade e personalização é possível



Figura 3.1: Tela principal do Viva Sem Tabaco.

montar um plano para recaídas e aprender a como evitá-las.

Os dados resultantes da interação do usuário com o sistema são armazenados em um banco de dados a fim de possibilitar pesquisas no que se refere à eficácia do programa, à sua usabilidade e à identificação de perfis dos usuários. Essa coleta de dados acontece principalmente durante as segunda e terceira etapas, onde o usuário responde aos formulários e escolhe entre diversas opções durante a criação dos planos. A figura 3.2 mostra um exemplo de um formulário durante a segunda etapa. Dados demográficos são extraídos do cadastro de usuários. Além disso, também há uma coleta implícita de dados, na qual todo o fluxo de navegação da sessão do usuário é rastreado e armazenado.

Existem técnicas que podem ajudar você a enfrentar a fissura.

Técnicas para enfrentar a fissura

As técnicas abaixo apesar de simples, apresentam resultados. Escolha as que você preferir para montarmos o seu plano:

- Beber um copo de água pausadamente.
- Comer alimentos com baixa quantidade de calorias como frutas cristalizadas (uva passas), balas dietéticas e chicletes dietéticos.
- Ler um cartão com suas razões para ter parado de fumar.
- Fazer exercício de relaxamento em áudio MP3 - link disponível no plano de parada.

Próximo passo

Saiba quais medicamentos podem aumentar suas chances de conseguir parar.

[VOLTAR](#) [AVANÇAR](#)

Vamos conversar? Online 

Figura 3.2: Exemplo de um formulário durante a etapa “Quero parar”.

3.2 Álcool e Saúde

O Álcool e Saúde(AES) tem como objetivo oferecer suporte e informações a pessoas que usam álcool. O *website* foi desenvolvido seguindo os princípios básicos da *Health on the Net Foundation* e todo seu conteúdo é baseado nas diretrizes para o tratamento de problemas relacionados ao uso de álcool da Organização Mundial de Saúde, Ministério da Saúde e do *National Institute on Alcohol Abuse and Alcoholism*. A figura 3.3 mostra a tela principal do Álcool e Saúde.

Assim como o *Viva sem Tabaco*, esta intervenção também é dividida em três etapas principais: “Avalie como você bebe”, “Querendo beber menos” e “Mantendo resultados”. A primeira etapa permite ao usuário avaliar seu hábito de beber. Após o teste, são apresentadas informações específicas a cada caso de uso do álcool: abstenção, uso não nocivo, uso de risco ou dependência. Se for o caso do usuário escolher parar de consumir ou diminuir seu consumo, a segunda etapa oferece mais conteúdos informativos e permite a criação de um plano personalizado para que o usuário beba menos. A terceira etapa refere-se ao acompanhamento para aqueles que cumpriram a etapa anterior. Nesta etapa, o usuário pode registrar diariamente suas dificuldades, seu consumo de álcool e verificar se suas metas preestabelecidas estão sendo atingidas.

Português

Cadastrar / Entrar

Página Inicial O Programa Sobre os autores Contato Para pesquisadores

Você sabe o quanto o álcool pode afetar a sua saúde ?

- Registre aqui o quanto você bebe e conheça melhor os efeitos do álcool no seu organismo.
- O conteúdo foi criado por profissionais de saúde e especialistas na área. [Saiba mais](#).
- A avaliação é gratuita e sempre será! Continue acessando ou faça seu cadastro agora mesmo [aqui](#).

Tela cheia indisponível. Saiba mais

Avalie como você bebe!

Querendo beber menos?

- Plano Personalizado
- Estratégias
- Dificuldades
- Benefícios

Mantendo resultados...

- Plano para revisar
- Revisão e ajuste
- Exatidão da avaliação
- Monitoramento

Dúvidas ou sugestões? Clique aqui

Olá, vamos conversar ? On-line

Figura 3.3: Tela principal do Álcool e Saúde.

Novamente como ocorre na intervenção previamente apresentada, são coletados dados da utilização do sistema para fins de pesquisa. Da mesma forma, os dados da interação do usuário nas três etapas e todo seu fluxo de navegação são armazenados em um banco de dados. Essa interação é exemplificada na figura 3.4.

O que é uma dose padrão?

Bebida	Volume	Alcool
Cerveja	340ml	5%
Vinho	140ml	12%
Uísque	40ml	40%
Cachaça	40ml	40%
Cerveja	600ml	40%

01. Com que frequência você consome bebidas alcoólicas?

- Nunca
- 1 vez ao mês ou menos
- De 2 a 4 vezes por mês
- De 2 a 3 vezes por semana
- 4 ou mais vezes por semana

02. Nas ocasiões em que bebe, quantas doses (*veja na ilustração acima quanto é uma dose) você consome tipicamente ao beber?

- 1 ou 2
- 3 ou 4
- 5 ou 6

Olá, vamos conversar ? On-line

Figura 3.4: Exemplo de um formulário durante a etapa “Avalie como você bebe”.

3.3 Avaliação das páginas informativas

Para a implementação dos sistemas de recomendação, foi adicionada em ambas as intervenções a funcionalidade de recomendação das páginas informativas. O componente de avaliação foi incluído nas páginas da etapa “Quero parar” do Viva Sem Tabaco e nas páginas da etapa “Querendo beber menos ?” do Álcool e Saúde.

Foi escolhida uma escala binária a fim de simplificar a avaliação para o usuário, facilitar a usabilidade e, por consequência, ter um maior número de avaliações coletadas. Desta forma, os usuários podem expressar sua opinião em relação aos conteúdos das páginas, definindo se as informações apresentadas foram úteis ou não foram úteis. As figuras 3.5 e 3.6 ilustram o sistema de avaliações.

Lembre-se!

1. Existem riscos para quaisquer números de cigarros fumados.
2. O uso de cigarros light não traz vantagens para sua saúde.

Fumo passivo

É o fumo que se dá através da respiração da fumaça produzida por um fumante e/ou pelo cigarro nas proximidades, sem que a pessoa fume de fato.

Danos à saúde

1. O fumo passivo também causa mal à saúde.
2. Em crianças o fumo passivo pode provocar a morte súbita infantil, problemas respiratórios e danos ao ouvido;
3. Em adultos, causa danos imediatos ao sistema cardiovascular, podendo causar doenças coronarianas e câncer de pulmão.

O que você pode fazer para ajudar

1. Fumar em lugares abertos e fora de locais públicos;
2. Fumar no exterior da casa para evitar o contato da fumaça, que fica no ambiente, entre familiares e animais domésticos.

Essa informação foi útil?  

Vamos conversar? Online 

Figura 3.5: Botões de avaliação em uma página do Viva Sem Tabaco.

Introdução

Ajuda Profissional

Buscando Apoio

Grupos Anônimos

Outros Recursos

Terapia

Ansiedade e Depressão

Não desista

Plano de Mudança

Parando de Beber - Apoio Profissional 76%

Medicamentos para tratar alcoolismo

Avanços no tratamento do alcoolismo fazem com que atualmente pacientes tenham mais alternativas de tratamento e que profissionais tenham mais ferramentas

- Novos medicamentos podem ajudá-lo a parar de beber atuando em vias neurais
- Estes novos medicamentos (ex. Naltrexona, Topiramato, etc) não o fazem se sentir com náuseas como os medicamentos mais antigos (ex. Disulfiram)
- Nenhuma destas medicações causam dependência, o que é a preocupação de alguns grupos de apoio e serviços de aconselhamento para álcool
- Estudos tem mostrado que pacientes com dependência de álcool podem receber tratamento efetivo combinando estas novas medicações e visitas breves à um profissional de saúde para acompanhamento

Essa página foi útil para você ?

VOLTAR PRÓXIMO

Atualizado em 27/07/2016

Olá, vamos conversar ? On-line 

Figura 3.6: Botões de avaliação em uma página do Álcool e Saúde.

4 Métodos

Neste capítulo é apresentado como foi conduzida a pesquisa deste trabalho. São introduzidos os algoritmos de agrupamento e a métrica de qualidade usada. Em seguida, são explicados o cálculo e avaliação das recomendações. Por fim, é feita uma breve explicação sobre o método de validação cruzada.

4.1 Agrupamento de Usuários

Em um sistema de recomendação com uma abordagem colaborativa, as recomendações de cada usuário são calculadas a partir de elementos coletados de usuários similares e que podem ter opiniões parecidas (EKSTRAND; RIEDL; KONSTAN, 2010). Uma alternativa para identificar esses elementos similares ao usuário que receberá a recomendação é definir uma espécie de vizinhança através de técnicas de agrupamento.

4.1.1 Caracterização dos usuários

O objetivo do agrupamento é gerar grupos de amostras de um conjunto de dados, no qual cada grupo contenha características parecidas. Neste caso, as amostras são os usuários e as características são os valores de atributos que definem o seu perfil. As Tabelas 4.1 e 4.2 mostram os atributos de usuários, assim como os tipos de dados, escolhidos para construir seus perfis para a intervenção VST e AES, respectivamente.

Tabela 4.1: Atributos para os usuários do VST.

atributo	tipo	valor
idade	inteiro	{16, ..., 61}
sexo	booleano	{0, 1}
ladder	inteiro	{1, ..., 10}
data_parar	booleano	{0, 1}
tentou_parar	booleano	{0, 1}
beber_agua	booleano	{0, 1}
comer	booleano	{0, 1}
ler_razoes	booleano	{0, 1}
relaxamento	booleano	{0, 1}

Tabela 4.2: Atributos para os usuários do AES.

atributo	tipo	valor
idade	inteiro	{22, ..., 69}
sexo	booleano	{0,1}
audit	inteiro	{0, ..., 50}
domingo	inteiro	{0, ..., 30}
segunda	inteiro	{0, ..., 30}
terça	inteiro	{0, ..., 30}
quarta	inteiro	{0, ..., 30}
quinta	inteiro	{0, ..., 30}
sexta	inteiro	{0, ..., 30}
sábado	inteiro	{0, ..., 30}

Na Tabela 4.1, o atributo *ladder* designa a motivação em parar de fumar indicada pelo usuário numa escala de 1 a 10, onde 1 é “nada motivado” e 10 é “totalmente motivado”. O atributo *data_parar* indica se o usuário escolheu uma data para parar de fumar diferente do dia em que fez o plano. O atributo *tentou_parar* indica se o usuário já tentou parar de fumar anteriormente. Os atributos *beber_agua*, *comer*, *ler_razoes* e *relaxamento* indicam, respectivamente, se o usuário escolheu adotar as seguintes técnicas para enfrentar a fissura: “beber um copo de água pausadamente”, “comer alimentos com baixa quantidade de calorias”, “ler um cartão com suas razões para ter parado de fumar” e “fazer exercício de relaxamento em áudio MP3”.

Na Tabela 4.2, o atributo *audit* é a pontuação no questionário *Alcohol Use Disorders Identification Test*, desenvolvido pela Organização Mundial de Saúde para avaliar o consumo, comportamentos e problemas relacionados ao álcool. Os demais campos são a quantidade de doses padrão consumidas em cada um dos dias da semana.

Para melhor se adequarem ao processo de aprendizagem, em ambos conjuntos de dados, os valores foram normalizados para uma escala de 0 a 1.

4.1.2 Métricas de similaridade

Um conceito importante no problema de agrupamento é a ideia de similaridade. Num bom agrupamento, amostras similares devem pertencer ao mesmo grupo e amostras distintas devem pertencer a grupos diferentes (XU; WUNSCH, 2005).

Existem diferentes métricas para o cálculo da similaridade. Para este trabalho,

foram consideradas a distância euclidiana, a distância de *Manhattan* e a similaridade do cosseno.

É importante notar que as métricas de distâncias denotam o inverso da similaridade como mostra a equação 4.1.

$$Dist(x, y) = \frac{1}{Sim(x, y)} \quad (4.1)$$

A equação 4.2 descreve a distância euclidiana entre as amostras x e y , onde x_f e y_f são os f -ésimos atributos de x e y . Quanto menor a distância euclidiana entre duas amostras mais similares elas são.

$$Dist(x, y) = \sqrt{\sum_{f=1}^d (x_f - y_f)^2} \quad (4.2)$$

A equação 4.3 descreve a distância de *Manhattan* entre as amostras x e y . Assim como na distância euclidiana, quanto menor a distância de *Manhattan* entre duas amostras mais similares elas são.

$$Dist(x, y) = \sum_{f=1}^d |x_f - y_f| \quad (4.3)$$

A equação 4.4 descreve a similaridade do cosseno entre as amostras x e y , definidas como vetores de atributos. A similaridade do cosseno está definida no intervalo $[-1, 1]$ e quanto mais próximo de 1 mais similares as amostras são.

$$Sim(x, y) = \cos\theta = \frac{x \cdot y}{\|x\| \times \|y\|} \quad (4.4)$$

Os atributos x_f e y_f usados nos cálculos são valores reais e, no caso deste trabalho, definidos nos intervalos mostrados nas tabelas 4.1 e 4.2 apresentadas anteriormente.

4.1.3 Algoritmos de agrupamento

Existem diversos métodos de agrupamento presentes na literatura. Para este estudo foram considerados os algoritmos *k-means* e o hierárquico aglomerativo com as diferentes

métricas de similaridade apresentadas.

O *k-means* baseia no cálculo iterativo de k centroides, à partir dos quais são definidos k grupos. Inicialmente, são definidos k centroides aleatórios. Cada ponto do conjunto de dados é iterativamente agrupado com o centroide mais próximo, de acordo com a métrica de similaridade adotada. A cada iteração, os centroides são recalculados como sendo a média das amostras de cada um dos k grupos. Usualmente, o processo é repetido até que os grupos estabilizem (MACQUEEN, 1967). O número k de grupos a serem gerados é dado como parâmetro. O algoritmo 1 apresenta cada uma das etapas do *k-means*.

Algoritmo 1: *k-means*

Entrada: conjunto de dados, número k de grupos.

Saída: conjunto de dados agrupado

1 início

2 | Escolha aleatoriamente k amostras do como os centroides iniciais;

3 | Agrupe cada amostra com o centróide mais próximo;

4 | Escolha novos centróides como sendo a média das amostras em cada um dos grupos;

5 | Repita 3 e 4 até que a diferença entre os novos e antigos centroides seja suficientemente pequena;

6 fim

Os algoritmos hierárquicos definem grupos unindo-os ou dividindo-os iterativamente. No caso do hierárquico aglomerativo, cada amostra inicialmente pertence a seu próprio grupo e a cada iteração os grupos mais similares são fundidos até que se tenha o número de grupos desejado. O número k de grupos é dado como parâmetro de entrada (KAUFMAN; ROUSSEEUW, 1990). A similaridade entre dois grupos pode ser calculada de diferentes formas, neste trabalho foi utilizado o critério *average link*. Este define a similaridade entre dois grupos como a média das similaridades de cada par de amostras formada por uma amostra de cada grupo. O algoritmo 2 mostra um pseudo-código do agrupamento hierárquico aglomerativo.

Algoritmo 2: *Agrupamento hierárquico aglomerativo***Entrada:** conjunto de dados, número k de grupos.**Saída:** conjunto de dados agrupado**1 início****2** | Defina um grupo para cada amostra;**3** | Agrupe cada grupo com o grupo mais próximo segundo o *average link*;**4** | Repita 3 até k grupos serem formados;**5 fim**

4.1.4 Métrica de qualidade

Para comparar os diferentes algoritmos e métricas de similaridade foi usado o método da Silhueta (ROUSSEEUW, 1987). O coeficiente de silhueta é calculado, para cada amostra, pela equação 4.5, onde para uma amostra x , a é a dissimilaridade média entre x e cada amostra do seu grupo, e b é a dissimilaridade média entre x e cada amostra do grupo mais próximo.

$$Silhueta(x) = \frac{a - b}{\max(a, b)} \quad (4.5)$$

O coeficiente de silhueta para um agrupamento é então definido como a média dos coeficientes de cada amostra. Os valores são limitados em $[-1, 1]$. Quanto mais próximo de 1, mais bem definido é o agrupamento.

4.2 Sistemas de Recomendação

4.2.1 Itens recomendados

Os itens dos sistemas de recomendação serão as páginas informativas apresentadas na seção anterior. Estas podem ser avaliadas pelos usuários positivamente ou negativamente. Avaliações positivas são representadas com o valor +1 e avaliações negativas com o valor -1. Os valores de avaliação são usados para o cálculo das recomendações como mostrado a seguir.

Desta forma, o sistema deve recomendar as páginas que correspondam com as preferências indicadas pelo usuário

4.2.2 Cálculo das predições e recomendações

Para a implementação do sistema de recomendação, foi adotada uma abordagem colaborativa baseada em usuários (EKSTRAND; RIEDL; KONSTAN, 2010). Neste modelo, para cada usuário, são calculadas as predições da sua avaliação para cada item ainda não avaliado. O cálculo da predição de um item i para um usuário u é dado pela equação 4.6, onde $r_{v,i}$ é a avaliação do item i pelo usuário v , $sim(u, v)$ é a similaridade entre os usuários u e v , e V é o conjunto de vizinhos usuário u .

$$P(u, i) = \frac{\sum_{v \in V} sim(u, v) \times r_{v,i}}{\sum_{v \in V} |sim(u, v)|} \quad (4.6)$$

A vizinhança é determinada pelos algoritmos de agrupamento apresentados na seção anterior. Desta forma, mesmo os usuários que ainda não avaliaram nenhum item podem ser recomendados de acordo com seus vizinhos. Esta é uma forma de evitar o problema da partida fria por novos usuários, a qual acontece quando a vizinhança é determinada pela similaridade a partir dos valores de avaliação.

Um valor de predição positivo designa um item que usuário pode gostar e, dessa forma, que pode ser recomendado. Já um valor de predição negativo indica um item que o usuário pode não gostar e este não deve ser recomendado.

4.2.3 Avaliação das recomendações

As recomendações geradas foram avaliadas através de avaliações internas ou avaliações *offline*, como definidas em (HERLOCKER et al., 2004). Os usuários são divididos em um grupo de treino e um grupo de testes. As avaliações dos usuários no grupo de treino são usadas para a construção do modelo. Já os usuários no grupo de teste têm as avaliações comparadas com as predições feitas pelo modelo.

Para a avaliação serão usadas métricas de *Precision/Recall*. Estas são comumente usadas em sistemas de recuperação da informação. Podendo considerar o objetivo de sistemas de recomendação como o de recuperar os itens mais apropriados ao usuário e cada recomendação a usuário como uma consulta, então essas métricas podem ser adaptadas a este contexto (EKSTRAND; RIEDL; KONSTAN, 2010). Elas são calculadas a partir da

Tabela 4.3: Matriz de confusão para o caso de recomendação

	Relevante	Irrelevante
Recomendado	VP	FP
Não recomendado	FN	VN

matriz de confusão apresentada na tabela 4.3, onde VP indica o número de verdadeiros positivos, FP de falsos positivos, FN de falsos negativos e VN de verdadeiros negativos. O *precision* é dado pela equação 4.7 e o *recall* pela equação 4.8.

$$P = \frac{VP}{VP + FP} \quad (4.7)$$

$$R = \frac{VP}{VP + FN} \quad (4.8)$$

O *precision* mostra a proporção de itens relevantes dentre os recomendados e pode avaliar a tarefa de “encontrar bons itens”. O *recall* indica a proporção de itens relevantes recomendados dentre todos os itens relevantes e pode ser usado para avaliar a tarefa de “encontrar todos os bons itens” (HERLOCKER et al., 2004).

É importante observar que para este trabalho, o *precision* é mais significativo que o *recall*. Já que o objetivo é que sejam recomendados um número fixo de bons itens, e não é necessário que todos os itens de interesse do usuário sejam recomendados imediatamente.

As duas métricas podem ser unidas em uma só através da média harmônica entre elas. Essa nova métrica chamada de F_1 ou *F-Measure* é dada pela equação 4.9 (HERLOCKER et al., 2004).

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (4.9)$$

4.3 Validação Cruzada

Para melhor avaliar a capacidade de generalização do modelo proposto foi utilizada validação cruzada. Em particular foi escolhido o método *K-fold* (KOHAVI, 1995).

Neste método, o conjunto de dados é dividido em K partes ou *folds*, dos quais k-1 *folds* são usados para treino e 1 *fold* é usado para teste. O conjunto de treinamento é usado para construir o modelo e o de teste para o cálculo de sua acurácia. O processo é então repetido K vezes, rotacionado a escolha dos *folds* para cada conjunto. Por fim, calcula-se

a acurácia média dentre todas as K iterações. Assim, tem-se um meio para avaliar a generalização dos modelos encontrados e para comparar as diferentes abordagens segundo essa generalização.

5 Experimentos

Aqui são apresentados os experimentos usados para avaliar os métodos propostos. Para cada uma das intervenções, foram propostos diferentes SRs combinando o modelo colaborativo com um algoritmo de agrupamento como:

- filtragem colaborativa + *k-means* usando distância euclidiana.
- filtragem colaborativa + hierárquico aglomerativo usando distância euclidiana.
- filtragem colaborativa + hierárquico aglomerativo usando distância de *Manhattan*.
- filtragem colaborativa + hierárquico aglomerativo usando similaridade do cosseno.

Cada algoritmo de agrupamento ainda foi testado com variando o parâmetro k (número de grupos) entre 2 e 10. Também foi proposto um recomendador aleatório para a comparação de resultados.

Os SRs propostos foram avaliados e comparados através dos três experimentos descritos nas próximas seções. O experimento 1 tem como objetivo avaliar o agrupamento e os experimentos 2 e 3 avaliam as recomendações geradas.

5.1 Conjunto de dados utilizado

Para o *Viva Sem Tabaco* foram usados registros de 40 usuários, 9 itens e 212 avaliações, enquanto que para o *Álcool e Saúde* foram usados registros de 81 usuários, 8 itens e 339 avaliações. Nesta amostragem, todos os usuários são avaliadores de pelo menos um item e todos os itens possuem ao menos uma avaliação.

As Figuras 5.1 e 5.2 apresentam comparações das distribuições quantitativas de avaliações positivas e negativas para cada uma das intervenções. Como pode ser visto nessas figuras, os dados são desbalanceados, de forma que existe uma quantidade de avaliações positivas muito superior às negativas.

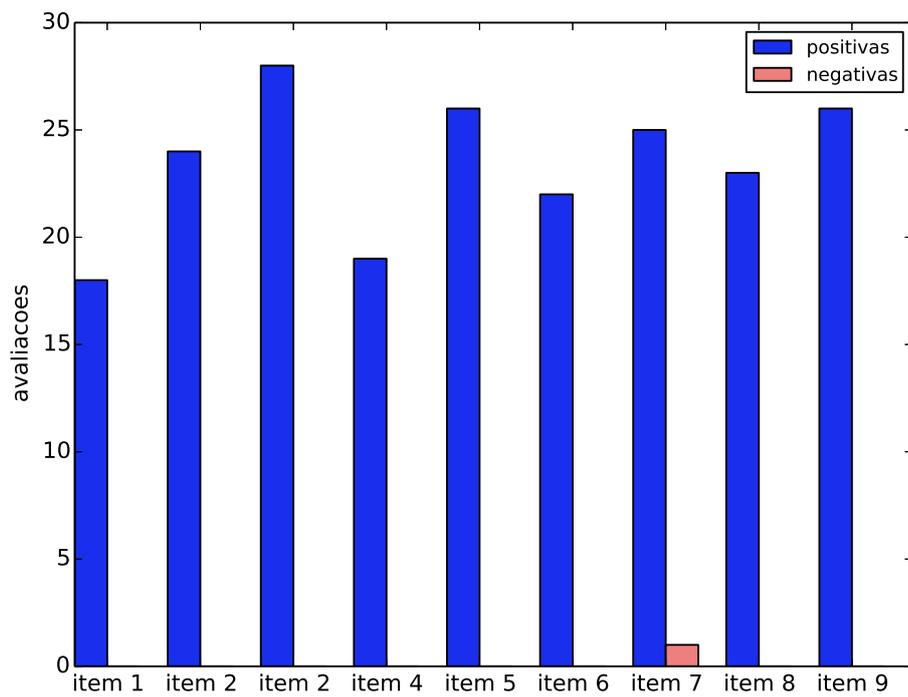


Figura 5.1: Distribuição de avaliações positivas e negativas para o VST.

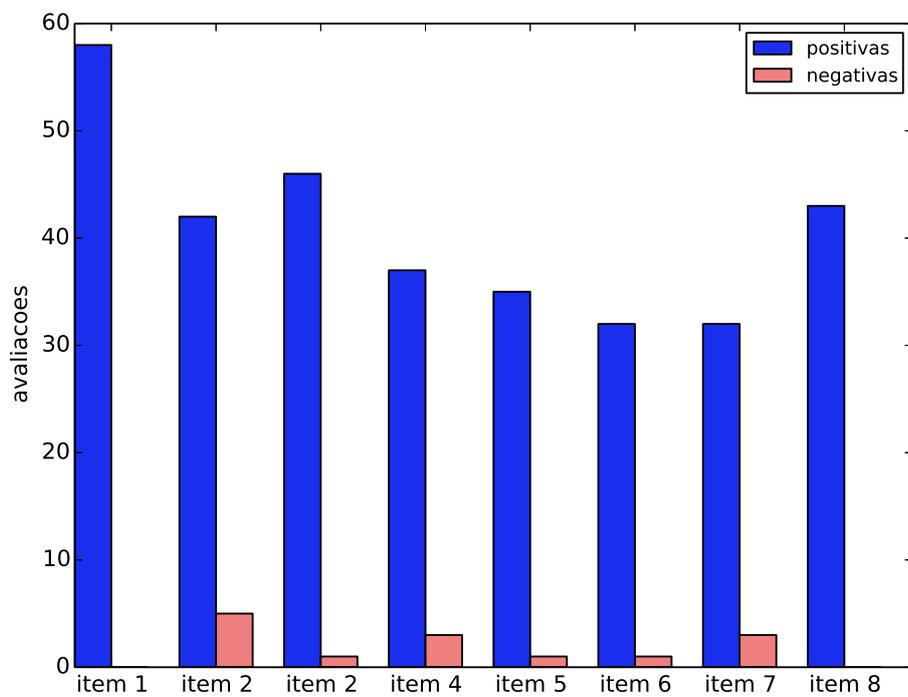


Figura 5.2: Distribuição de avaliações positivas e negativas para o AES.

5.1.1 Validação cruzada estratificada

Para compensar o desbalanceamento dos dados, foi utilizada a validação cruzada *stratified k-fold*. Nesta variação do *k-fold*, os folds são divididos de forma a manter a mesma proporção de amostras de cada classe que o conjunto de dados original. Neste, caso manter a mesma proporção de avaliações positivas e negativas.

5.2 Experimento 1

Este primeiro experimento teve como objetivo avaliar a qualidade dos agrupamentos gerados. Para tal, foi usado o coeficiente de silhueta (ROUSSEEUW, 1987). Apesar do algoritmo hierárquico aglomerativo ser determinístico, o k-means possui um aspecto aleatório. Portanto, para este foram realizadas 10 execuções independentes com sementes distintas. As Tabelas 5.1 e 5.2 mostram as médias do coeficiente de silhueta entre todas as execuções.

Tabela 5.1: Resultados do experimento 1 para o VST - coeficiente de silhueta.

grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.1852	0.2048	0.3011	0.4474
3	0.2050	0.1914	0.2804	0.4097
4	0.2107	0.2103	0.3019	0.4032
5	0.2231	0.2176	0.2942	0.3426
6	0.2391	0.2126	0.2953	0.4013
7	0.2624	0.2318	0.3220	0.4321
8	0.2854	0.2171	0.2979	0.4616
9	0.3042	0.2763	0.3674	0.4486
10	0.3186	0.2747	0.3670	0.4875

Tabela 5.2: Resultados do experimento 1 para o AES - coeficiente de silhueta.

grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.4093	0.4654	0.5177	0.6044
3	0.2976	0.4096	0.4200	0.5440
4	0.2009	0.3128	0.3022	0.5092
5	0.2018	0.3806	0.2642	0.3850
6	0.2010	0.3671	0.3137	0.3713
7	0.1967	0.3222	0.2729	0.3043
8	0.1989	0.3056	0.2770	0.2750
9	0.1918	0.2873	0.2568	0.3709
10	0.1928	0.2640	0.2454	0.3357

Analisando os resultados para o VST, pode-se perceber que os agrupamentos com

mais grupos mostraram maior coeficiente silhueta. O valor do coeficiente acompanha o crescimento do número de grupos com algumas oscilações.

Para o AES, o coeficiente de silhueta também mostrou oscilações, porém este decresce à medida que o número de grupos aumenta. Como resultado, observa-se que todos os algoritmos obtiveram melhores agrupamentos com apenas 2 grupos.

Em ambos os casos o hierárquico aglomerativo com a similaridade do cosseno mostrou um desempenho significativamente melhor. Pode-se concluir assim, que esta métrica de similaridade se adaptou melhor à característica dos dados utilizados.

5.3 Experimento 2

No primeiro experimento, os SRs propostos foram avaliados levando em consideração uma lista de tamanho fixo contendo as melhores predições para cada usuário. O *precision*, *recall* e *f-measure* foram calculados sobre a lista de recomendação em comparação aos valores reais de avaliação. A seguir, uma descrição do passo a passo do experimento:

1. Gerar grupos de usuários através do algoritmo de agrupamento.
2. Usando validação cruzada com 5 *folds*, separar os usuários em conjuntos de treino e teste.
3. Para cada usuário de teste, calcular as predições de cada item com base nos usuários de treino.
4. Para cada usuário de teste, construir uma lista de recomendação contendo os três itens com maior valor de predição.
5. Calcular o *precision*, *recall* e *f-measure* sobre cada lista de recomendação.
6. Calcular o valor médio das métricas entre todas as listas.
7. Calcular a média entre todos os *folds* da validação cruzada.

Além do aspecto estocástico do k-means, neste experimento, os usuários são embaralhados de forma aleatória para a formação dos *folds*. Sendo assim, o experimento foi executado 10 vezes com diferentes sementes. Como resultado final são retornadas as

médias de cada métrica em todas execuções. As próximas seções apresentam as tabelas com resultados e uma discussão para cada uma das intervenções.

5.3.1 Viva Sem Tabaco

Aqui são apresentados os resultados do experimento 2 para o VST. Para cada uma dos SRs propostos, a tabela 5.3 exibe os valores de *precision*, a tabela 5.4 exibe os valores de *recall* e a tabela 5.5 exibe os valores de *f-measure*. No cabeçalho de cada tabela também é mostrado o valor da respectiva métrica para o recomendado aleatório.

Podemos perceber que não houve concordância dos melhores valores de *precision* e *recall* com os melhores valores de coeficiente de silhueta. Ainda, os SRs com maior valor de *precision* não foram os mesmos que os SRs com maior valor de *recall*.

Todos os SRs mostraram-se melhor que o recomendador aleatório em ambas as métricas. Pode-se observar que o aleatório teve um valor de *precision* alto. Isso deve-se à pouca quantidade de avaliações negativas, sendo fácil, por chance, escolher três itens que tenham avaliações positivas. Avaliando ambas as métricas, pode-se observar pelo *f-measure* que o SR usando *k-means* com 4 grupos mostrou o melhor desempenho.

Tabela 5.3: Resultados do experimento 2 para o VST - *precision*.

precision		aleatório = 0.8809		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.9778	0.9772	0.9769	0.9772
3	0.9769	0.9769	0.9764	0.9764
4	0.9772	0.9775	0.9769	0.9769
5	0.9772	0.9775	0.9778	0.9764
6	0.9769	0.9769	0.9775	0.9772
7	0.9769	0.9772	0.9772	0.9775
8	0.9772	0.9526	0.9524	0.9775
9	0.9750	0.9513	0.9528	0.9518
10	0.9750	0.9525	0.9527	0.9531

Tabela 5.4: Resultados do experimento 2 para o VST - *recall*.

recall		aleatório = 0.4644		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.6061	0.6066	0.6070	0.6057
3	0.6051	0.6061	0.6069	0.6068
4	0.6071	0.6065	0.6052	0.5846
5	0.6057	0.5853	0.5860	0.5855
6	0.5996	0.5844	0.5844	0.5852
7	0.5932	0.5850	0.5863	0.5834
8	0.5871	0.5717	0.5725	0.5840
9	0.5815	0.5717	0.5722	0.5727
10	0.5832	0.5692	0.5701	0.5705

Tabela 5.5: Resultados do experimento 2 para o VST - *f-measure*.

f-measure		aleatório = 0.5714		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.7160	0.7161	0.7164	0.7155
3	0.7151	0.7157	0.7162	0.7161
4	0.7166	0.7161	0.7151	0.6989
5	0.7154	0.6995	0.7000	0.6994
6	0.7106	0.6987	0.6988	0.6993
7	0.7053	0.6993	0.7002	0.6981
8	0.7009	0.6820	0.6826	0.6984
9	0.6949	0.6818	0.6825	0.6826
10	0.6966	0.6787	0.6794	0.6813

5.3.2 Álcool e Saúde

Aqui são apresentados os resultados do experimento 2 para o AES. Para cada uma dos SRs propostos, a tabela 5.6 exibe os valores de *precision*, a tabela 5.7 exibe os valores de *recall* e a tabela 5.8 exibe os valores de *f-measure*. No cabeçalho de cada tabela também é mostrado o valor da respectiva métrica para o recomendado aleatório.

Diferente do caso do VST, pode-se perceber que o número de grupos que gerou os melhores valores de *precision* e *recall* também são aqueles que alcançaram os melhores valores de coeficiente de silhueta. O melhor desempenho foi obtido pelo SR com o algoritmo aglomerativo e 2 grupos, assim como o melhor coeficiente de silhueta.

Assim como no VST, todos os SRs mostraram-se melhor que o recomendador aleatório. Neste caso, dado que os dados são formados por um número maior de avaliações negativas, o aleatório teve um valor de *precision* ainda menor.

Tabela 5.6: Resultados do experimento 2 para o AES - *precision*.

precision		aleatório = 0.79098		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.9592	0.9467	0.9472	0.9601
3	0.9596	0.9350	0.9346	0.9598
4	0.9594	0.9220	0.9218	0.9349
5	0.9601	0.9226	0.9101	0.9096
6	0.9505	0.9102	0.9090	0.8969
7	0.9393	0.8970	0.8984	0.8974
8	0.9115	0.8989	0.8981	0.8854
9	0.9105	0.8984	0.8974	0.8849
10	0.8865	0.8485	0.8722	0.8725

Tabela 5.7: Resultados do experimento 2 para o AES - *recall*.

recall		aleatório = 0.4899		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.7047	0.6950	0.6932	0.7047
3	0.7039	0.6834	0.6840	0.7056
4	0.7056	0.6772	0.6767	0.6949
5	0.7049	0.6776	0.6734	0.6700
6	0.6973	0.6735	0.6727	0.6613
7	0.6876	0.6602	0.6619	0.6620
8	0.6613	0.6618	0.6623	0.6487
9	0.6633	0.6619	0.6549	0.6497
10	0.6442	0.6134	0.6385	0.6364

Tabela 5.8: Resultados do experimento 2 para o AES - *f-measure*.

f-measure		aleatório = 0.5727		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.7801	0.7692	0.7685	0.7805
3	0.7797	0.7576	0.7579	0.7810
4	0.7808	0.7493	0.7489	0.7662
5	0.7805	0.7498	0.7433	0.7412
6	0.7722	0.7434	0.7425	0.7309
7	0.7618	0.7301	0.7317	0.7315
8	0.7351	0.7318	0.7319	0.7186
9	0.7360	0.7318	0.7252	0.7192
10	0.7155	0.6828	0.7072	0.7061

5.4 Experimento 3

É possível perceber que as métricas de *precision* e *recall* são dependentes do tamanho da lista de recomendações, de tal forma que ao aumentar a lista de recomendações o *precision* tende a diminuir e o *recall* tende a aumentar (HERLOCKER et al., 2004). Sendo assim, foi proposto um segundo experimento, levando em conta todos os itens de predições positivas na lista de recomendação. O experimento é descrito a seguir:

1. Gerar grupos de usuários através do algoritmo de agrupamento.
2. Usando validação cruzada com 5 *folds*, separar os usuários em conjuntos de treino e teste.
3. Para cada usuário de teste, calcular as predições de cada item com base nos usuários de treino.
4. Para cada usuário de teste, construir uma lista de recomendação contendo todos os itens com valor de predição positivo.
5. Calcular o *precision*, *recall* e *f-measure* sobre cada lista de recomendação.
6. Calcular o valor médio das métricas entre todas as listas.
7. Calcular a média entre todos os *folds* da validação cruzada.

Assim como no experimento anterior, foram feitas 10 execuções com sementes diferentes. As próximas seções apresentam as tabelas com resultados e as análises para cada uma das intervenções.

5.4.1 Viva Sem Tabaco

Aqui são apresentados os resultados do experimento 3 para o VST. Para cada um dos SRs propostos, a tabela 5.6 exhibe os valores de *precision*, a tabela 5.7 exhibe os valores de *recall* e a tabela 5.8 exhibe os valores de *f-measure*. No cabeçalho de cada tabela também é mostrado o valor da respectiva métrica para o recomendador aleatório.

Assim como no experimento 2, não é possível notar alguma concordância entre os melhores agrupamentos e as melhores recomendações. Para todos os SRs, o *precision* não

apresentou variações consideráveis em relação ao experimento anterior. Porém, nota-se um aumento significativo do *recall*.

Todos os SRs mostraram-se melhores que o recomendador aleatório. Neste experimento, o desempenho do recomendador aleatório mostrou-se um pouco inferior, já que o número de itens a serem escolhidos é maior. Avaliando ambas as métricas, pode-se observar pelo *f-measure* que o SR usando *k-means* com 2 grupos mostrou o melhor desempenho.

Tabela 5.9: Resultados do experimento 3 para o VST - *precision*.

precision		aleatório = 0.8456		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.9778	0.9772	0.9769	0.9772
3	0.9769	0.9769	0.9764	0.9764
4	0.9772	0.9775	0.9769	0.9769
5	0.9772	0.9775	0.9778	0.9764
6	0.9769	0.9769	0.9775	0.9772
7	0.9769	0.9772	0.9772	0.9775
8	0.9772	0.9526	0.9524	0.9775
9	0.9750	0.9513	0.9528	0.9518
10	0.9750	0.9525	0.9527	0.9531

Tabela 5.10: Resultados do experimento 3 para o VST - *recall*.

recall		aleatório = 0.4693		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.9778	0.9741	0.9739	0.9744
3	0.9764	0.9738	0.9733	0.9701
4	0.9772	0.9715	0.9710	0.9469
5	0.9696	0.9408	0.9416	0.9467
6	0.9589	0.9350	0.9352	0.9446
7	0.9458	0.9260	0.9250	0.9348
8	0.9345	0.8953	0.8945	0.9292
9	0.9119	0.8944	0.8963	0.8989
10	0.9039	0.8706	0.8721	0.8956

Tabela 5.11: Resultados do experimento 3 para o VST - *f-measure*.

f-measure		aleatório = 0.5787		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.9778	0.9756	0.9753	0.9757
3	0.9767	0.9753	0.9747	0.9730
4	0.9772	0.9743	0.9738	0.9585
5	0.9728	0.9539	0.9546	0.9582
6	0.9661	0.9505	0.9508	0.9574
7	0.9569	0.9456	0.9450	0.9521
8	0.9509	0.9167	0.9161	0.9489
9	0.9337	0.9157	0.9174	0.9198
10	0.9300	0.8977	0.8987	0.9185

5.4.2 Álcool e Saúde

Aqui são apresentados os resultados do experimento 3 para o AES. Para cada um dos SRs propostos, a tabela 5.6 exibe os valores de *precision*, a tabela 5.7 exibe os valores de *recall* e a tabela 5.8 exibe os valores de *f-measure*. Como no experimento 2, foi possível perceber uma que o número de grupos que gera os melhores valores de recomendação são aqueles com os melhores valores de coeficiente de silhueta. Com uma lista de recomendação maior, o *precision* não sofreu grandes variações, o *recall* teve um aumento significativo e o aleatório teve uma queda de desempenho. Considerando ambas as métricas, o SR com o algoritmo hierárquico com a similaridade do cosseno e dois grupos mostrou novamente o melhor desempenho.

Tabela 5.12: Resultados do experimento 3 para o AES - *precision*.

precision		aleatório = 0.7574		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.9570	0.9442	0.9449	0.9577
3	0.9572	0.9324	0.9323	0.9573
4	0.9570	0.9194	0.9195	0.9324
5	0.9575	0.9200	0.9077	0.9319
6	0.9480	0.9077	0.9065	0.9198
7	0.9404	0.8945	0.8958	0.9196
8	0.9344	0.8965	0.8957	0.9080
9	0.9268	0.8969	0.8948	0.9080
10	0.9232	0.8961	0.8698	0.8947

Tabela 5.13: Resultados do experimento 3 para o AES - *recall*.

recall		aleatório = 0.4863		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.9758	0.9630	0.9636	0.9763
3	0.9757	0.9509	0.9507	0.9743
4	0.9755	0.9379	0.9382	0.9509
5	0.9762	0.9388	0.9260	0.9506
6	0.9650	0.9264	0.9253	0.9384
7	0.9527	0.9133	0.9144	0.9380
8	0.9452	0.9151	0.9142	0.9266
9	0.9363	0.9096	0.8980	0.9267
10	0.9288	0.9025	0.8885	0.9134

Tabela 5.14: Resultados do experimento 3 para o AES - *f-measure*.

f-measure		aleatório = 0.5638		
grupos	k-means	agl. dist. euclidiana	agl. dist. Manhattan	agl. sim. cos
2	0.9631	0.9503	0.9510	0.9638
3	0.9633	0.9384	0.9383	0.9626
4	0.9631	0.9255	0.9256	0.9385
5	0.9636	0.9262	0.9137	0.9380
6	0.9532	0.9139	0.9127	0.9259
7	0.9423	0.9006	0.9019	0.9256
8	0.9355	0.9026	0.9018	0.9141
9	0.9267	0.9001	0.8893	0.9141
10	0.9207	0.8951	0.8759	0.9008

5.5 Melhores Resultados

A tabela 5.15 apresenta os melhores resultados de cada uma das métricas para ambas as intervenções.

No experimento 1, observa-se que ambos tiveram valores positivos para o coeficiente de silhueta. Indicando que foram encontrados agrupamentos bem definidos.

Embora tenham sido avaliados o *precision* e o *recall* com a lista de recomendação limitada e ilimitada, é importante considerar quais métricas são mais relevantes para a avaliação dos sistemas.

O experimento 2 considera uma situação próxima à implementação real dos sistemas, na qual apenas três itens são recomendados. O *precision* é a métrica mais significativa, já que mostra que os sistemas têm capacidade de encontrar ao menos três itens que sejam relevantes. O *recall* tem valores mais baixos, porém é menos significativo para a avaliação, já que não é tarefa dos sistemas de recomendação encontrar todos os itens relevantes.

O experimento 3 não limita a lista de recomendações, avaliando os sistemas sob a tarefa de encontrar todos os itens relevantes. Apesar de não representar o que seria o objetivo real dos sistemas de recomendação, este serve para observar o comportamento das métricas. O *precision* não tem seu valor reduzido, o que corrobora a ideia da influência do desbalanceamento dos dados. Já o *recall* tem um aumento no seu valor como esperado, uma vez que não é limitado a somente três itens.

Tabela 5.15: Melhores resultados encontrados para ambas as intervenções.

	Viva Sem Tabaco	Álcool e Saúde
Experimento 1 - Coeficiente de Silhueta	0.4875	0.6044
Experimento 2 - Precision	0.9778	0.9601
Experimento 2 - Recall	0.6071	0.7056
Experimento 2 - F-Measure	0.7166	0.7805
Experimento 3 - Precision	0.9778	0.9577
Experimento 3 - Recall	0.9778	0.9763
Experimento 3 - F-Measure	0.9778	0.9638

6 Conclusões

6.1 Considerações Finais

Observando os resultados do *f-measure* nos experimentos apresentados, conclui-se que para o Viva Sem Tabaco, as melhores recomendações foram geradas com o uso do *k-means* com 2 e 4 grupos. Já para o Álcool e Saúde as melhores recomendações foram geradas com o uso do algoritmo hierárquico aglomerativo e similaridade do cosseno com 2 grupos. Desta forma, a escolha da melhor técnica pode ser dependente de cada aplicação.

No que se refere à customização de intervenções, este trabalho serviu para corroborar a ideia do uso de sistemas de recomendação como uma solução viável. A possibilidade de execução da proposta e os resultados, em geral, positivos mostram que os sistemas de recomendação podem contribuir positivamente para a escolha de conteúdo de acordo com o perfil do usuário.

A utilização de agrupamento para a tarefa de recomendação mostrou-se uma solução adequada. Apesar de não ser observado de maneira global, em certos casos bons agrupamentos resultaram em boas recomendações.

A qualidade das recomendações, segundo as métricas observadas, foi significativamente melhor do que a qualidade das recomendações aleatórias. Isso mostra que o sistema conseguiu aprender o que é positivo para cada usuário. Porém, devido a falta de avaliações negativas nos conjuntos de teste, não é possível afirmar que o sistema de recomendação tenha aprendido o que é negativo para cada usuário. Isso resulta em falsos positivos, ou seja, itens não relevantes mas que o sistema julga como o contrário. Pode-se então observar que existe um viés nos resultados originado do desbalanceamento dos dados.

Os sistemas de recomendação para as intervenções Viva Sem Tabaco e Álcool e Saúde são viáveis e têm potencial de resultados positivos. Porém o prosseguimento do estudo envolvendo outros algoritmos, conjuntos de dados mais completos e outros tipos de avaliações podem levar a resultados ainda mais promissores.

6.2 Trabalhos Futuros

Como trabalhos futuros, a principal proposta é realizar a avaliação dos sistemas de recomendação diretamente com o *feedback* dos usuários. Esta será feita enviando uma lista com as principais recomendações a cada usuário e coletando suas opiniões. Desta forma, será possível avaliar a qualidade das recomendações com mais confiabilidade.

Além disso, os experimentos realizados neste trabalho podem gerar resultados significativos à medida que mais dados são coletados pelos sistemas. Sendo então importante continuar com as avaliações feitas aqui e acompanhando seus resultados.

Outros algoritmos de agrupamento que exploram diferentes geometrias, como o DBSCAN que é baseado em densidades, podem apresentar resultados mais distintos. Combinando outros algoritmos e, ainda, outras métricas de distância é possível explorar mais os dados e talvez aprender outros padrões.

Também propõe-se estudar outros modelos de recomendação, como o colaborativo baseado em itens e o modelo baseado em conteúdo.

Outra proposta é utilizar os sistemas de recomendação para indicar mensagens contendo dicas de saúde e bem estar. Estas estão presentes no **Álcool e Saúde** em quantidade maior do que as páginas usadas neste trabalho.

Bibliografia

- ADOMAVICIUS, G.; TUZHILIN, a. Toward the Next Generation of Recommender Systems: a Survey of the State of the Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, v. 17, n. 6, p. 734–749, 2005. ISSN 10414347.
- BENNETT, J.; LANNING, S. The Netflix Prize. *KDD Cup and Workshop*, p. 3–6, 2007. ISSN 1554351X.
- BOBADILLA, J. et al. Recommender systems survey. *Knowledge-Based Systems*, v. 46, p. 109–132, 7 2013. ISSN 09507051.
- EKSTRAND, M. D.; RIEDL, J. T.; KONSTAN, J. A. Collaborative Filtering Recommender Systems. *Foundations and Trends® in Human–Computer Interaction*, v. 4, n. 2, p. 81–173, 2010. ISSN 1551-3955, 1551-3963.
- FOLTZ, P. W.; DUMAIS, S. T. Personalized information delivery: an analysis of information filtering methods. *Communications of the ACM*, vol.35,, n. 12, p. 51–60, 1992. ISSN 00010782.
- GOMIDE, H. et al. Intervenções computadorizadas customizadas para mudança de comportamentos. In: MURTA, S. G. et al. (Ed.). *Prevenção e Promoção em Saúde Mental: fundamentos, planejamento e estratégias de intervenção*. [S.l.]: Sinopsys, 2015. p. p. 381–404.
- HERLOCKER, J. L. et al. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, v. 22, n. 1, p. 5–53, 2004. ISSN 1046-8188.
- HORS-FRAILE, S. et al. Design of two combined health recommender systems for tailoring messages in a smoking cessation app. *Proceedings of the 10th ACM Conference on Recommender Systems - RecSys '16*, 2016.
- KAUFMAN, L.; ROUSSEEUW, F. P. *Groups in Data: An Introduction to Cluster Analysis: Wiley*. [S.l.: s.n.], 1990.
- KOHAVI, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *International Joint Conference on Artificial Intelligence*, 1995.
- LIEBERMAN, D. Z. et al. Predicting Content Preference: Applying Lessons Learned from the Commercial Web to Therapeutic Software. Cyberpsychology. *Journal of Psychosocial Research on Cyberspace*, v. 2, n. 2, p. article 1, 2008.
- LU, J. et al. Recommender system application developments: A survey. *Decision Support Systems*, v. 74, p. 12–32, 6 2015. ISSN 01679236.
- MACQUEEN, J. Some methods for classification and analysis of multivariate observations. *Proc. 5th Berkeley Symp.*, v. 1, p. 281–297, 1967.
- RESNICK, P. et al. GroupLens : An Open Architecture for Collaborative Filtering of Netnews. *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, p. 175–186, 1994. ISSN 00027863.

ROUSSEEUW, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis . *Journal of Computational and Applied Mathematics*, p. 53–65, 1987.

SADASIVAM, R. S. et al. Collective-intelligence recommender systems: Advancing computer tailoring for health behavior change into the 21st century. v. 18, n. 3, p. 42–3, 2016. ISSN 14388871.

SAMUEL, A. L. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, v. 3, n. 3, p. 210–229, 1959. ISSN 0018-8646.

SATHYA, R.; ABRAHAM, A. Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification. *International Journal of advanced Research in Artificial Intelligence*, v. 2, n. 2, p. 34–38, 2013. ISSN 21654069.

SHINDE, S. K.; KULKARNI, U. Hybrid personalized recommender system using centering-bunching based clustering algorithm. *Expert Systems with Applications*, Elsevier Ltd, v. 39, n. 1, p. 1381–1387, 2012. ISSN 09574174.

WHO, W. *eHealth*. 2016. Acessado: 2016-10-29. Disponível em: <<http://www.who.int/topics/ehealth/en/>>.

XU, R.; WUNSCH, D. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, v. 16, n. 3, p. 645–678, 5 2005. ISSN 10459227.

YAO, Z. Y. Z.; ZHANG, Q. Z. Q. Item-Based Clustering Collaborative Filtering Algorithm under High-Dimensional Sparse Data. *2009 International Joint Conference on Computational Sciences and Optimization*, v. 1, p. 787–790, 2009.

YASAVUR, U.; AMINI, R.; LISETTI, C. User Modeling for Pervasive Alcohol Intervention Systems. *First International Workshop on Recommendation Technologies for Lifestyle Change (LIFESTYLE 2012) and First International Workshop on Interfaces for Recommender Systems (InterfaceRS 2012)*, p. 29 – 33, 2012.

ZHU, R. R.; GONG, S. S. Analyzing of collaborative filtering using clustering technology. *2009 ISECS International Colloquium on Computing, Communication, Control, and Management*, v. 4, p. 57–59, 2009.