

REVISÃO BIBLIOGRÁFICA SOBRE SISTEMAS DE RECOMENDAÇÃO

CAPÍTULO 2 DO TRABALHO DE CONCLUSÃO DE CURSO DO MBA CIÊNCIA DE DADOS DA USP 2021

2 Sistemas de Recomendação

Os sistemas de recomendação representam uma solução inovadora para tentar absorver informações de comportamento dos consumidores, com a finalidade de se identificar preferências e, assim, sugerir produtos com maior potencial de venda (HWANGBO, KIM e CHA, 2018).

Os sistemas de recomendação são organizados para auxiliar pessoas a selecionarem e realizarem escolhas baseados nas informações disponíveis (TAGHAVI, BAKHTIYARI e SCAVINO, 2013). Por promover indicações a usuários, auxilia a sociedade como um todo a evitar desperdício de tempo, energia e custos, em um processo de otimização de recursos (ASOSHEHA, BAGHERPOUR e YAHYAPOUR, 2008).

De acordo com Zibriczky (2016), os sistemas de recomendação são úteis em diversas aplicações em que existe uma quantidade significativa de escolhas e, ao mesmo tempo, os usuários estão interessados em apenas uma pequena parcela de produtos.

A recomendação de produtos tem sido cada vez mais utilizada pelas empresas ao redor do mundo em inúmeras áreas. Estima-se que 60% dos vídeos visualizados no Youtube e 80% dos filmes assistidos no Netflix sejam fruto de recomendações (SHOKEEN e RANA, 2019)

Recorrentemente, os algoritmos de recomendação enfrentam um conjunto de desafios, tais como: lidar com uma grande quantidade de informações (de usuários e produtos); a necessidade de se obter respostas em tempo real; limitação das informações a respeito dos clientes; excessivo conjunto de informações a respeito de clientes recorrentes; e, volatilidade das informações a cada interação do usuário (LINDEN, SMITH e YORK, 2003). Sendo assim, diversas alternativas dentro do escopo de indicação ao usuário têm sido historicamente testadas para minimizar tais dificuldades.

O presente capítulo busca sintetizar as abordagens que envolvem os sistemas de recomendação, as vantagens e desvantagens de cada um deles e distintas aplicações reais voltadas para o setor financeiro.

2.1 Regras de Associação

As regras de associação são aplicadas para se obter padrões frequentes, associações e correlações entre conjuntos de dados (PASSI, JAIN e SINGH, 2019). O algoritmo Apriori, de Agrawal, Imielinski e Swami (1993), é considerado o primeiro algoritmo aplicado para mineração de conjuntos de produtos e regras de associação (FACELI et al., 2011).

Considerando um conjunto de itens, por exemplo, X, Y e Z, o objetivo técnico das regras de associação é estabelecer relações do tipo: quem compra o item X, compra o item Y ou quem compra Z, compra o item X.

Como a metodologia a ser aplicada no algoritmo Apriori envolve, normalmente, um número substancial de varreduras nos dados para se chegar nas conclusões, o algoritmo FP-Growth, de Han, Pei e Mao (2004), também ganhou destaque na literatura por reduzir o custo computacional (FACELI et al., 2011)

Alguns autores consideram as regras de associação como um tipo de sistema de recomendação, como Bogaert et al. (2019). Entretanto, o mais comum é encontrar estas regras fora da tipologia de sistemas de recomendação. Um motivo fundamental para isto seria que, na esmagadora maioria dos casos, as regras de associação são utilizadas como ferramenta exploratória, e não preditiva; além do fato de não existir qualquer teste de validação para obtenção dos padrões entre os itens (RUDIN et al., 2011).

Apesar da vantagem na facilidade de aplicação e interpretação presente nas regras de associação, conforme Passi, Jain e Singh (2019), há diversos aspectos problemáticos neste tipo de abordagem. Em primeiro lugar, os autores apontam elevado custo computacional para bancos de dados extensos, tanto no tradicional algoritmo Apriori quanto no algoritmo FP-Growth; em segundo lugar, a recomendação seria considerada fraca, no sentido de que poucas regras são encontradas quando o nível de confiança é alto, de modo que, para alguns produtos, nenhuma regra é gerada; e, em terceiro lugar, as recomendações em regras de associação não são personalizadas. Este último ponto implica que são consideradas as transações de todo o conjunto de informações, não considera as avaliações dos usuários e, portanto, a recomendação direcionada a um usuário está em função do comportamento de outros, os quais podem ser significativamente distintos.

2.2 Tipos de sistemas de recomendação

Os sistemas de recomendação são estruturados em três possibilidades básicas: (1) baseados em conteúdo, (2) baseados em filtros colaborativos ou (3) híbridos. Nesta seção serão apresentadas as características e comparativos entre os tipos, de modo a se identificar o funcionamento, as vantagens e desvantagens de cada um deles.

2.2.1 Sistema de recomendação baseado em conteúdo

Os sistemas de recomendação baseados em conteúdo (*content-based*) tem por característica fundamental utilizar a descrição dos atributos dos produtos para promover recomendações. As avaliações e comportamentos de consumo dos usuários são combinados às informações disponíveis do conteúdo dos itens para se gerar indicações (AGGARWAL, 2016). Conforme Izicki e Santos (2020), a partir das

variáveis presentes nos produtos que o usuário tem interesse, objetiva-se encontrar produtos com características semelhantes.

Os métodos envolvidos na recomendação baseada em conteúdo representam aqueles aplicados em aprendizado supervisionado. A descrição dos itens juntamente às avaliações aplicadas por um usuário específico – conhecido como usuário ativo – são empregadas como dados de treinamento e aplicados para criar uma classificação ou modelo de regressão personalizados. A partir dos dados de treinamento, busca-se prever se o usuário ativo em questão terá ou não ter interesse pelo item a ser sugerido (AGGARWAL, 2016). Os sistemas baseados em conteúdo são largamente estruturados a partir de métodos de regressão (IZBICKI e SANTOS, 2020).

Assim como qualquer sistema de recomendação, as indicações baseadas em conteúdo possuem vantagens e desvantagens.

Segundo Aggarwal (2016), a abordagem baseada em conteúdo possui a vantagem de lidar bem com o ingresso de novos produtos no catálogo. Isto se deve ao fato de que outros produtos com características similares já foram avaliados pelo usuário ativo e, assim, o modelo supervisionado aplicado a este seria capaz de valer-se do histórico de avaliações promovidas por ele. Além desta questão, Lops, Gemmis e Semeraro (2011) levantam outras duas vantagens em relação a sistemas de filtro colaborativo: (1) há independência nas recomendações ao usuário ativo, de modo a não depender dos interesses de outros usuários, e (2) a modelagem aplicada a cada usuário ativo é inteligível a ponto de não ser considerado um sistema caixa preta.

Por outro lado, Aggarwal (2016) alerta para duas desvantagens dos sistemas baseados em conteúdo. Em primeiro lugar, as recomendações são consideradas demasiadamente óbvias e, em segundo lugar, não são capazes de gerar recomendações efetivas para novos usuários, dada a indisponibilidade de avaliações por parte destes.

2.2.2 Sistema de recomendação por filtragem colaborativa

2.2.2.1 A filtragem colaborativa

A introdução do termo “filtragem colaborativa” se deu a partir da criação do primeiro sistema de recomendação para uso comercial, o *Tapestry*. Este sistema foi construído para gerar recomendações de documentos extraídos de grupos de notícias (MELVILLE e SINDHWANI, 2010).

De acordo com Schafer et al. (2007, p. 294) a aplicação de técnicas puramente baseadas em conteúdo seria frequentemente inadequada. O sistema *Tapestry*, criado pela *Xerox PARC* incluiu as ações e opiniões de usuários, a partir do conteúdo das mensagens dos usuários, o que posteriormente seria reconhecido como

o primeiro passo na direção da construção de indicações por filtros colaborativos. O *Tapestry* ficou conhecido como um sistema em que o usuário que possuía interesse na recomendação resgatava-a ativamente no banco de dados. A motivação do sistema foi a necessidade de se filtrar listas de e-mails de acordo com os interesses do usuário, dado o contexto da larga expansão na comunicação via mensagens eletrônicas (GOLDBERGM, NICHOLS e TERRY, 1992).

Em um esforço adicional, Maltz e Ehrlich (1995) promoveram um sistema de recomendação por filtragem colaborativa capaz de facilitar a transmissão de indicações de um usuário para outro (*pointers*).

Ao refletir sobre o tema, Schafer et al. (2007) colocam os filtros colaborativos com a mesma essência da forma de comunicação utilizada historicamente pelos seres humanos, que, a partir do compartilhamento de opiniões entre si sobre a experiência com os produtos comprados (no “boca-a-boca”), acabam por gerar sugestões uns aos outros.

Conforme Zibriczky (2016), apesar do surgimento em meados dos anos 1990, os sistemas de recomendação obtiveram destaque no *Netflix Prize*. Este prêmio foi vinculado a um desafio para a comunidade de pesquisadores de aprendizado de máquina promovido pela Netflix, em outubro de 2006, com o objetivo de se desenvolver um sistema de recomendação o mais efetivo possível (BENNETT et al., 2007).

De forma ampla, devido ao substancial aumento da capacidade de processamento dos computadores nas últimas décadas foi permitida a coleta das opiniões em tempo real, não somente observando as preferências de cada comunidade, mas criando indicações personalizadas de determinado produto para grupos específicos de usuários. (SCHAFFER et al., 2007).

A filtragem colaborativa envolve utilizar a avaliação de produtos a partir da opinião de usuários/clientes e pode criar recomendações para um ou mais itens. De acordo com Schafer et al. (2007) com as avaliações em um sistema de filtragem colaborativa: (1) podem ser escalares com aspecto numérico (com notas de 1 a 5, por exemplo) ou ordinais (concordo, concordo plenamente, neutro, discordo, discordo plenamente, por exemplo); (2) avaliações binárias; e, (3) avaliações unárias, ao verificar se um usuário observou, comprou ou avaliou de forma positiva determinado produto.

Em relação à obtenção das avaliações, Schafer et al. (2007) apontam que estas podem ser realizadas de modo explícito, implícito ou de ambas as formas. As avaliações explícitas podem ser obtidas a partir das opiniões de usuários acerca de cada produto, ao passo que as avaliações implícitas poderiam ser provenientes das ações dos potenciais clientes (por exemplo, itens consultados no site ou compras recorrentes)

Em um sistema de filtragem colaborativa, de acordo com os interesses, preferências, atividades e comportamentos de uma quantidade significativa de usuários e os gostos de um particular usuário, a predição normalmente é realizada com base na similaridade com outros usuários. A premissa fundamental desse tipo de sistema de recomendação é que as preferências dos indivíduos são temporalmente fixas, de

modo que aqueles que preferiram um produto no passado persistirão com a preferência no futuro e, adicionalmente, vão preferir itens similares (KANAUJIA, PANDEY e SWARUP RAUTARAY, 2017).

As tarefas de dos sistemas de filtragem colaborativa podem ser sintetizados em: (I) recomendação de produtos, (II) predição para determinado produto; e (III) recomendação restrita, em que se insere um conjunto de regras para filtrar dentro de um grupo específico de produtos (SCHAFER et al., 2007). Para cumprir tais tarefas, a próxima subseção é dedicada a compreender os principais tipos de filtros colaborativos.

2.2.2.2 Tipos de filtros colaborativos

O desenvolvimento de sistemas de recomendação por filtragem colaborativa envolve duas ramificações: *user-based collaborative filtering systems* e *item-based collaborative filtering systems*.

Na abordagem *user-based* (ou *user-to-user*) indica-se produtos escolhidos pela maior parte dos consumidores considerados similares. A similaridade é calculada a partir das avaliações realizadas no passado dos usuários. As medidas de correlação e semelhança de cosseno são largamente utilizadas neste contexto (HWANGBO, KIM e CHA, 2018).

Um tradicional algoritmo de filtragem colaborativa é aquele em que representa determinado usuário com um vetor N-dimensional de produtos. Neste cenário, os componentes são positivos quando comprados – ou bem avaliados – e negativos, para produtos avaliados de forma negativa. Por característica, o algoritmo multiplica o vetor de componentes pelo vetor inverso com a informação das frequências dos itens comprados ou avaliados. O algoritmo constrói recomendações a partir de cálculo de similaridade a partir do cosseno do ângulo entre dois vetores (LINDEN, SMITH e YORK, 2003, p. 77).

Entretanto, segundo Linden, Smith e York (2003) a aplicação de tal modelo representa dificuldades significativas, tais como, esparsidade, escalabilidade e *performance*. O quadro 1 apresenta uma síntese das desvantagens associadas ao uso de algoritmos baseados em filtros colaborativos.

Cold Start	As técnicas de filtragem colaborativa exigem expressivo número de usuários para se aplicar as recomendações com acurácia e eficiência. Os filtros colaborativos possuem dificuldade em lidar com novos produtos e usuários.
Escalabilidade	Se o número de usuários for muito expressivo, maior capacidade computacional é exigida para que se possa realizar as recomendações.
Esparsidade	O número de produtos e itens de sites de e-commerce normalmente são extensos. Em muitas situações, os usuários mais ativos podem ter avaliado um conjunto muito pequeno de produtos comprados.

Fonte: Kanaujia, Pandey e Swarup Rautaray (2017, p. 20) e Koren e Volinsk (2009)

Quadro 1 – Dificuldades enfrentadas em algoritmos baseados em filtros colaborativos

Para além dos problemas apresentados por Kanaujia, Pandey e Swarup Rautaray (2017) e Koren e Volinsk (2009), outras questões devem ser consideradas na visão de Passi, Jain e Singh (2019). Estes colocam dois pontos adicionais: (1) a presença de falsos vizinhos próximos (em métodos de vizinhança) e (2) a presença de “ovelhas-negras”. Estes últimos se referem a usuários com gostos não-convencionais comparados aos demais, de modo que as recomendações para estes indivíduos tendem a ser pouco adequadas, como também podem atrapalhar a recomendação aos demais.

Ao longo das próximas seções será possível identificar algumas possíveis soluções para o enfrentamento dos problemas relacionados aos algoritmos associados a filtros colaborativos. Uma delas seria a aplicação da abordagem *item-based* (ou *item-to-item*). Este tipo de sistema verifica similaridades entre os produtos e recomenda os itens que são mais similares aos itens selecionados pelo usuário (Hwangbo, Kim e Cha, 2018, p. 95).

Linden, Smith e York (2003) formularam um tipo de algoritmo de filtragem colaborativa *item-to-item*, com a intensão de lidar com um conjunto massivo de dados – com muitos produtos e usuários – e produzir recomendações em tempo real e, ao mesmo tempo, com alta qualidade, para o *Amazon.com*. Os autores explicam que ao invés de combinarem usuários similares, o algoritmo criado por eles, segue as seguintes etapas:

1º) equipara-se cada compra ou avaliação dos consumidores (normalmente pela similaridade por cosseno);

2º) combina-se os itens similares em uma lista de recomendações, a qual pode ser obtida buscando itens que os consumidores tendem a comprar em conjunto;

3º) recomenda-se itens mais similar ao(s) comprado(s) ou avaliado(s) pelo usuário, a partir da tabela de itens similares. O algoritmo busca produtos similares a cada compra ou avaliação do usuário.

Na figura 1 está o resumo do funcionamento do algoritmo.

Figura 1 – Modelo básico do sistema *item-to-item*

```
For each item in product catalog,  $I_1$ 
  For each customer  $C$  who purchased  $I_1$ 
    For each item  $I_2$  purchased by
      customer  $C$ 
      Record that a customer purchased  $I_1$ 
        and  $I_2$ 
  For each item  $I_2$ 
    Compute the similarity between  $I_1$  and  $I_2$ 
```

Fonte: Linden, Smith e York (2003, p.79)

De acordo Linden, Smith e York (2003), o ponto crucial para geração de escalabilidade e performance por meio do algoritmo de filtragem colaborativa *item-to-item* é criar uma tabela de itens similares *offline*, sendo que o único componente *online* seria a quantidade de produtos que o consumidor já tiver comprado ou avaliado. A busca *online* estaria restrita a buscar bens similares (etapa 3), de modo a não ter problemas de escalabilidade em função do número de produtos catalogados ou quantidade de usuários; inclusive, aos consumidores que compram/avaliam de forma muito recorrente, poderia ser realizada uma amostragem das ações destes para haver maior eficiência computacional.

2.2.2.3 Abordagem da filtragem colaborativa neural (NFC)

O sistema de filtragem colaborativa tem sido explorado, também, a partir de *deep learning*, conhecida como filtragem colaborativa neural (NCF).

A recomendação por redes neurais é aplicada em ambientes de elevado volume de informações, como textos e músicas. Entretanto, He, Liao e Zang (2017) realizaram um esforço em aplicar a filtragem colaborativa com uma arquitetura neural, a qual aprende uma função arbitrária a partir dos dados. As evidências empíricas alcançadas pelos autores mostraram uma melhor performance de recomendação, quando comparado a métodos tradicionais, como a decomposição matricial e aplicação do produto interno – os quais são utilizados para extrair a interação entre usuários e características dos produtos.

De acordo com Shokeen e Rana (2019), a aplicação de *deep learning* no contexto de sistemas de recomendação têm se mostrado úteis para amenizar problemas de cold-start e esparsidade de dados.

2.2.3 Sistema de recomendação híbrido

A filtragem híbrida é uma abordagem que contempla um *mix* entre a filtragem baseada em conteúdo e a colaborativa, com o objetivo de aumentar a acurácia e eficácia das recomendações (KANAUJIA, PANDEY e SWARUP RAUTARAY, 2017).

Em geral, as técnicas de filtragem híbrida utilizam algumas características da filtragem colaborativa dentro do contexto da abordagem baseada em conteúdo; ou, de modo contrário, aplicam recursos da filtragem baseada em conteúdo dentro da abordagem da filtragem colaborativa (SHOKEEN e RANA, 2019).

De acordo com Kanaujia, Pandey e Swarup Rautaray (2017, p. 21), a *Netflix* recomenda filmes a partir de uma abordagem híbrida, pois realiza indicações a partir de um comparativo entre os hábitos de cada usuário e o interesse de usuários similares (filtragem colaborativa), bem como recomenda filmes com características próximas às preferências de cada usuário apoiado naquilo que já assistiu e/ou avaliou.

Entretanto, alguns problemas podem surgir com a aplicação de sistemas híbridos. Shokeen e Rana (2019) apontam que os algoritmos criados em torno da abordagem híbrida lidam, geralmente, com problemas como *cold-start*, esparsidade dos dados, privacidade, escalabilidade e confiança.

2.3 Sistema de recomendação com filtragem baseada em conteúdo versus colaborativa

2.3.1 Comparativo entre sistema de recomendação por filtragem colaborativa e sistema de filtragem baseada em conteúdo

Schafer et al. (2007) apontam premissas e características fundamentais nas quais baseiam-se a filtragem colaborativa parte e a filtragem baseada em conteúdo, cuja síntese está apresentada no Quadro 2.

	Filtragem colaborativa	Filtragem baseada em conteúdo
Avaliação e Similaridade	Usuários com gostos similares irão avaliar itens de forma similar	Os produtos similares serão avaliados similarmente pelo usuário que os consome
Predição	Necessita de avaliações para realizar predições	Pode realizar predições sem avaliações de produtos
Conteúdo	Não requer conteúdo informações sobre o produto)	Requer conteúdo (informações sobre o produto)
Variabilidade das recomendações	Maior capacidade de gerar recomendações inesperadas	Tendência a recomendar produtos extremamente similares aos consumidos pelo usuário (<i>over-specialize</i>).

Fonte: Adaptado de Schafer et al. (2007, p. 300)

Quadro 2 – Comparativo entre Filtragem colaborativa e Filtragem baseada em conteúdo

A partir das características específicas dos sistemas de recomendação por filtragem colaborativa e filtragem baseada em conteúdo, é possível combiná-las de como a se ter uma abordagem híbrida. É importante apontar que não existe consenso na literatura sobre o melhor caminho a ser seguido para combinar os sistemas. Entretanto, uma possibilidade é analisar o conteúdo para identificar produtos que combinam diretamente à necessidade do usuário e utilizam a filtragem colaborativa para obter características mais difíceis de serem extraídas.

2.3.2 Principais métodos aplicados em sistemas de recomendação por filtragem colaborativa e baseados em conteúdo

Os primeiros sistemas de recomendações estavam associados a correlações estatísticas e modelagem preditiva, mas não vinculadas a um conjunto amplo de práticas estatísticas e literatura de aprendizado de máquina (MELVILLE e SINDHWANI, 2010). Com o passar do tempo, a literatura sobre sistemas de recomendação foi estendida e a síntese dos algoritmos aplicados está sintetizada no quadro 3.

Filtragem colaborativa	Filtragem baseada em conteúdo
K-vizinhos próximos	Modelo de representação de espaço vetorial (TF-IDF)
Correlação de Pearson	Feedback de relevância
Desvio médio quadrático	Algoritmo Rocchio
Similaridade por cosseno	Classificadores lineares
Fatoração (decomposição) matricial	Métodos probabilísticos
Classificadores bayesianos	Naive Bayes
Métodos baseados em regressão	Função de similaridade por cosseno
Técnicas de redução de dimensionalidade	Árvores de decisão

Fonte: Kanaujia, Pandey e Swarup Rautaray (2017, p. 21)

Quadro 3 - Exemplos de algoritmos aplicados em filtragem colaborativa e filtragem baseada em conteúdo

Em termos de metodologias, atualmente, as principais áreas que envolvem filtragem colaborativa são os métodos de vizinhança e os modelos de fator latente. Os métodos de vizinhança estão focados em cálculos ligados à relação entre produtos ou usuários. Na abordagem orientada pelos produtos (*item-to-item*), sistema recomenda itens vizinhos aos de preferência do mesmo indivíduo. Por outro lado, a abordagem baseada no usuário (*user-to-user*) considera as avaliações de um usuário em relação às de outros. A partir do cálculo de similaridade entre as avaliações dos usuários, o sistema busca recomendar produtos bem avaliados por outros usuários “vizinhos” (KOREN, BELL e VOLINSK, 2009).

A partir da metade da década dos anos 2000, um conjunto de esforços para aprimorar os métodos para criação de sistemas de recomendação deu origem a aplicação de modelos de fatores latentes, entre os quais, os mais bem sucedidos, de acordo com Koren, Bell e Volinsk (2009), são os estruturados a partir de fatoração matricial.

Segundo Melville e Sindhwani (2010), ao invés de se basearem na vizinhança, tal abordagem parte do princípio que a similaridade entre usuários e produtos são simultaneamente causadas por alguma dimensão oculta dentro do conjunto de dados.

Na abordagem de modelos de fator latente tenta-se explicar classificações dadas pelos usuários a partir dos produtos e usuários em um conjunto de fatores (de 20 a 100) obtidos dos padrões de avaliação (KOREN, BELL e VOLINSK, 2009). A ideia é criar dimensões com características pessoais e/ou do

conteúdo do produto e segmentar os usuários de acordo com as intersecções das características reveladas por eles, agrupando-os.

2.4 Sistemas de recomendação no setor financeiro

A dinâmica do mercado financeiro tem sido alterada de forma expressiva, a partir do rápido desenvolvimento da tecnologia da informação. Os serviços financeiros oferecidos no mercado online criaram oportunidades para os clientes poderem acompanhar as modificações em tempo real e, assim, poderem operar de forma mais eficiente (TAGHAVI, BAKHTIARI e SCAVINO, 2013).

Esta seção é dedicada a discutir os desafios específicos de sistemas de recomendação que envolvam o setor financeiro. Para tal, optou-se por inserir os aspectos singulares relacionados à modelagem aplicada em recomendações de produtos financeiros tendo como pano de fundo, para efeito comparativo, o formato de indicações para produtos convencionais. Posteriormente é feita a apresentação de desafios e algumas aplicações recentes envolvendo recomendações de produtos financeiros.

2.4.1 Desafios dos sistemas de recomendação de produtos financeiros

De acordo com Musto et al (2015), a criação de um sistema de recomendação para investimentos financeiros representa uma tarefa complexa, pois normalmente envolve o encontro entre as necessidades e restrições dos investidores. As instituições financeiras – como, bancos e corretoras, por exemplo –, ao buscarem responder aos interesses de seus clientes, trabalham na direção de oferecer a estes os serviços mais indicados. Entre os problemas envolvidos, Musto et al. (2015) apontam a ausência de captura do histórico real do usuário em investimentos (seja com feedback positivo ou não), além do fato dos clientes manterem, normalmente, os ativos por um longo período. Este horizonte de investimento dificulta a obtenção de uma amostra suficientemente significativa com a finalidade de se construir um sistema de recomendação convencional, o qual está vinculado às experiências anteriores dos clientes. Zibriczky (2016) argumenta que a recomendação de produtos financeiros exige, em geral, um comprometimento de longo prazo, em função de uma série de fatores externos, como riscos de mercado, risco de liquidez, riscos legais, entre outros. Para além destes pontos, o autor alerta, também, para a necessidade de conhecimento para otimização das escolhas por parte dos usuários; deve-se considerar que, para reduzir os riscos, os usuários tendem a atuar de forma mais conservadora, comparado ao consumo de produtos provenientes de e-commerce.

Outra diferença substancial entre recomendação de produtos financeiros e os produtos convencionais, de acordo com Zibriczky (2016), é a elevada restrição de acesso a dados de clientes, o que, em geral, tende a tornar mais complexa a tarefa de recomendação. A privacidade é uma questão que, por gerar ausência de dados, em muitos casos, gera *cold-start problem*.

Para complementar os desafios envolvidos, Oyeboade e Orji (2020) atribuem parte da complexidade à ausência de explícita disponibilidade de avaliações por parte dos usuários. Este aspecto, na visão dos autores se deve a dois fatores fundamentais: (1) as instituições não possuem uma listagem de avaliações de produtos bancários de seus clientes e (2) dificuldade em gerar recomendações para novos clientes e clientes potenciais (*cold-start problem*).

2.4.2 Aplicações de sistemas de recomendação no setor financeiro

Com a finalidade de obter um sistema de recomendação adequado para produtos bancários, Oyeboade e Orji (2020) buscaram contornar as dificuldades provenientes de um sistema de filtragem colaborativa ou baseado em conteúdo puros, a partir de um sistema híbrido, com a consideração de duas abordagens: filtragem colaborativa baseada em produtos (*item-based*) e demográfica. Os autores utilizaram uma base com 393.816 clientes de um banco e alcançaram acurácia preditiva superior a partir do sistema híbrido, em comparação às abordagens de forma isolada.

Os problemas existentes na aplicação isolada de sistemas de filtragem colaborativa ou de sistemas baseados em conteúdo também levaram Musto et. al (2015) a procurar outra forma de atuar sobre um conjunto de dados 1.172 usuários. Os autores combinaram um sistema de recomendação *case-based* com um algoritmo de diversificação de investimentos. Para tal guiaram-se a partir de diversos atributos, os quais envolveram perfil de risco, objetivos de rentabilidade dos ativos, tempo de investimento, experiência com investimentos, ativos financeiros disponíveis, sexo e idade.

Gigli, Lillo e Regoli (2017) analisaram um conjunto de dados de um banco europeu de grande porte, o qual continha um numeroso conjunto de ativos de investimentos de 200 mil clientes. Os autores aplicaram três algoritmos para recomendação: (1) ranking bayesiano personalizado, a partir do método de fatoração matricial com a maximização da probabilidade a posteriori da estrutura de preferência do usuário, (2) mínimos quadrados usando método de fatores latentes e (3) algoritmo Word2Vec com 300 dimensões espaciais.

Gonzalez-Carrasco et al. (2012) modelaram um sistema de recomendação de portfólios de investimentos a clientes de instituições financeiras, denominado PB-ADVISOR. O sistema é baseado em lógica *fuzzy* e tecnologia semântica. O PB-ADVISOR levou em consideração dois aspectos fundamentais de cada cliente: o viés comportamental e a tolerância a risco.

REFERÊNCIAS

- AGGARWAL, C. C. **Recommender Systems: the textbook**. Cham: Springer. 2016. E-book. 518 p. DOI 10.1007/978-3-319-29659-3.
- AGRAWAL, R., IMIELINSKI, T.; SWAMI, A. Mining association rules between sets of items in large databases. In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 1993, Washington D.C., p. 207–216.
- ASOSHEHA, A.; BAGHERPOUR, S.; YAHYAPOUR, N. Extended acceptance models for recommender system adaption, case of retail and banking service in Iran. **WSEAS transactions on business and economics**, v. 5, n. 5, p. 189-200, 2008.
- BENNETT, J. et al. The netflix prize. In: **Proceedings of KDD cup and workshop**. 2007. p. 35.
- BOGAERT, M. et al. Evaluating multi-label classifiers and recommender systems in the financial service sector. **European Journal of Operational Research**, v. 279, n. 2, p. 620-634, 2019.
- FACELI, K. et al. **Inteligência artificial: uma abordagem de aprendizado de máquina**. 2. ed. Rio de Janeiro: LTC, 2021.
- HAN, J., PEI, J., Yin, Y. e MAO, R. *Mining frequent patterns without candidate generation*. **Data Mining and Knowledge Discovery**, p. 53-87, 2004.
- IZBICKI, R.; SANTOS, T. M. dos. **Aprendizado de máquina: uma abordagem estatística**. São Carlos: Rafael Izbicki, 2020. E-book. Disponível em: < <http://www.rizbicki.ufscar.br/AME.pdf>>. Acesso em: 08 abr. 2021.
- GIGLI, A.; LILLO, F.; REGOLI, D. Recommender Systems for Banking and Financial Services. In: **RecSys Posters**. 2017.
- GONZALEZ-CARRASCO, I. et al. PB-ADVISOR: A private banking multi-investment portfolio advisor. **Information Sciences**, v. 206, p. 63-82, 2012.
- HAN, J., PEI, J., Yin, Y.; MAO, R. *Mining frequent patterns without candidate generation*. **Data Mining and Knowledge Discovery**, p. 53-87, 2004.
- HE, X. et al. Neural collaborative filtering. In: **Proceedings of the 26th international conference on world wide web**. 2017. p. 173-182.
- HWANGBO, H; KIM, Y. S.; CHA, K. J. Recommendation system development for fashion retail e-commerce. **Electronic Commerce Research and Applications**, v. 28, p. 94-101, 2018.
- KANAUJIA, P. K. M.; PANDEY, M.; SWARUP RAUTARAY, S. A Framework for Development of Recommender System for Financial Data Analysis. **International Journal of Information Engineering & Electronic Business**, v. 9, n. 5, 2017.
- KOREN, Y.; BELL, R.; VOLINSKY, C. Matrix factorization techniques for recommender systems. **Computer**, v. 42, n. 8, p. 30-37, 2009.
- LINDEN, G.; SMITH, B.; YORK, J. Amazon. com recommendations: Item-to-item collaborative filtering. **IEEE Internet computing**, v. 7, n. 1, p. 76-80, 2003.

LOPS, P.; DE GEMMIS, M.; SEMERARO, G. Content-based recommender systems: State of the art and trends. **Recommender systems handbook**, p. 73-105, 2011.

MALTZ, D.; EHRLICH, K. Pointing the way: Active collaborative filtering. In: **Proceedings of the SIGCHI conference on Human factors in computing systems**. 1995. p. 202-209.

MELVILLE, P.; SINDHWANI, V. Recommender systems. **Encyclopedia of machine learning**, v. 1, p. 829-838, 2010.

MUSTO, C. et al. Personalized finance advisory through case-based recommender systems and diversification strategies. **Decision Support Systems**, v. 77, p. 100-111, 2015.

OYEBODE, Oladapo; ORJI, Rita. A hybrid recommender system for product sales in a banking environment. **Journal of Banking and Financial Technology**, p. 1-11, 2020.

RUDIN, C. et al. Sequential event prediction with association rules. In: **Proceedings of the 24th annual conference on learning theory**. JMLR Workshop and Conference Proceedings, 2011. p. 615-634.

SCHAFER, J. B. et al. Collaborative filtering recommender systems. In: **The adaptive web**. Springer, Berlin, Heidelberg, 2007. p. 291-324.

SHOKEEN, J.; RANA, C. An application-oriented review of deep learning in recommender systems. **International Journal of Intelligent Systems and Applications**, v. 10, n. 5, p. 46, 2019.

TAGHAVI, M.; BAKHTIYARI, K.; SCAVINO, E. Agent-based computational investing recommender system. In: **Proceedings of the 7th ACM conference on recommender systems**. 2013. p. 455-458.

ZIBRICZKY, D. Recommender systems meet finance: a literature review. In: **Proceedings of the 2nd Workshop Personalization Recommender System**. 2016. p. 1-10.