

---

# Capítulo 1

## Uma introdução ao Recomendador Sistemas

---

“Muitos recebem conselhos, somente os sábios lucram com eles.” – Harper Lee

### 1.1 Introdução

---

A crescente importância da Web como meio para transações eletrônicas e comerciais tem servido como força motriz para o desenvolvimento da tecnologia de sistemas de recomendação. Um importante catalisador nesse sentido é a facilidade com que a Web permite que os usuários forneçam feedback sobre o que gostam ou não gostam. Por exemplo, considere o cenário de um provedor de conteúdo como a Netflix. Nesses casos, os usuários podem facilmente fornecer feedback com um simples clique do mouse. Uma metodologia típica para fornecer feedback é na forma de classificações, nas quais os usuários selecionam valores numéricos de um sistema de avaliação específico (por exemplo, sistema de classificação de cinco estrelas) que especificam o que gostam e não gostam de vários itens.

Outras formas de feedback não são tão explícitas, mas são ainda mais fáceis de coletar no paradigma centrado na web. Por exemplo, o simples ato de um usuário comprar ou navegar por um item pode ser visto como uma recomendação para aquele item. Essas formas de feedback são comumente usadas por comerciantes online como a Amazon.com, e a coleta desse tipo de dados é completamente descomplicada em termos do trabalho exigido do cliente. A ideia básica dos sistemas de recomendação é utilizar essas diversas fontes de dados para inferir os interesses do cliente.

A entidade à qual a recomendação é fornecida é chamada de usuário, e o produto recomendado também é chamado de item. Portanto, a análise de recomendações geralmente se baseia na interação anterior entre usuários e itens, pois interesses e propensões passados costumam ser bons indicadores de escolhas futuras. Uma exceção notável

é o caso dos sistemas de recomendação baseados em conhecimento, nos quais as recomendações são sugeridas com base nos requisitos especificados pelo usuário, em vez do histórico do usuário.

Então, qual é o princípio básico que fundamenta o funcionamento dos algoritmos de recomendação?

O princípio básico das recomendações é que existem dependências significativas entre a atividade centrada no usuário e a atividade centrada no item. Por exemplo, um usuário interessado em um documentário histórico tem maior probabilidade de se interessar por outro documentário histórico ou por um programa educacional do que por um filme de ação. Em muitos casos, várias categorias de itens podem apresentar correlações significativas, que podem ser aproveitadas para fazer recomendações mais precisas.

Alternativamente, as dependências podem estar presentes na granularidade mais fina de itens individuais, em vez de categorias. Essas dependências podem ser aprendidas de forma orientada por dados a partir da matriz de classificações, e o modelo resultante é usado para fazer previsões para usuários-alvo.

Quanto maior o número de itens avaliados disponíveis para um usuário, mais fácil será fazer previsões robustas sobre seu comportamento futuro. Diversos modelos de aprendizagem podem ser utilizados para realizar essa tarefa. Por exemplo, o comportamento coletivo de compra ou avaliação de vários usuários pode ser aproveitado para criar grupos de usuários semelhantes interessados em produtos semelhantes. Os interesses e as ações desses grupos podem ser aproveitados para fazer recomendações a membros individuais desses grupos.

A descrição acima mencionada baseia-se em uma família muito simples de algoritmos de recomendação, conhecidos como modelos de vizinhança. Essa família pertence a uma classe mais ampla de modelos, conhecida como filtragem colaborativa. O termo “filtragem colaborativa” refere-se ao uso de avaliações de múltiplos usuários de forma colaborativa para prever avaliações ausentes. Na prática, os sistemas de recomendação podem ser mais complexos e ricos em dados, com uma ampla variedade de tipos de dados auxiliares. Por exemplo, em sistemas de recomendação baseados em conteúdo, o conteúdo desempenha um papel fundamental no processo de recomendação, no qual as avaliações dos usuários e as descrições dos atributos dos itens são utilizadas para fazer previsões. A ideia básica é que os interesses dos usuários podem ser modelados com base nas propriedades (ou atributos) dos itens que eles avaliaram ou acessaram no passado. Uma estrutura diferente é a dos sistemas baseados em conhecimento, nos quais os usuários especificam seus interesses interativamente, e a especificação do usuário é combinada com o conhecimento do domínio para fornecer recomendações. Em modelos avançados, dados contextuais, como informações temporais, conhecimento externo, informações de localização, informações sociais ou informações de rede, podem ser usados.

Este livro estudará todos os tipos de sistemas básicos, incluindo sistemas colaborativos, baseados em conteúdo e baseados em conhecimento. Também discutiremos os modelos básicos e aprimorados de sistemas de recomendação em diferentes domínios. Estudaremos vários aspectos da robustez dos sistemas de recomendação, como modelos de ataque e a construção de modelos confiáveis. Além disso, uma variedade de modelos de avaliação e hibridização para sistemas de recomendação serão estudados exaustivamente. Neste capítulo, o objetivo é fornecer uma visão geral da ampla diversidade de trabalhos na área de sistemas de recomendação e também relacionar os diversos tópicos aos capítulos individuais deste livro.

Este capítulo está organizado da seguinte forma. A Seção 1.2 discute os principais objetivos dos sistemas de recomendação. A Seção 1.3 apresenta os modelos básicos e métodos de avaliação utilizados em sistemas de recomendação. O uso de sistemas de recomendação em diversos domínios de dados é discutido na Seção 1.4. Modelos avançados para sistemas de recomendação são discutidos na Seção 1.5. A Seção 1.6 discute as conclusões e o resumo.

## 1.2 Objetivos dos Sistemas de Recomendação

---

Antes de discutir os objetivos dos sistemas de recomendação, apresentamos as diversas maneiras pelas quais o problema de recomendação pode ser formulado. Os dois modelos principais são os seguintes:

1. Versão preditiva do problema: A primeira abordagem consiste em prever o valor da classificação para uma combinação usuário-item. Supõe-se que os dados de treinamento estejam disponíveis, indicando as preferências do usuário pelos itens. Para  $m$  usuários e  $n$  itens, isso corresponde a uma matriz  $m \times n$  incompleta, onde os valores especificados (ou observados) são usados para o treinamento. Os valores ausentes (ou não observados) são previstos usando este modelo de treinamento. Este problema também é conhecido como problema de complementação de matriz, pois temos uma matriz de valores incompletamente especificada, e os valores restantes são previstos pelo algoritmo de aprendizado.
2. Versão de classificação do problema: Na prática, não é necessário prever as avaliações dos usuários para itens específicos para fazer recomendações aos usuários. Em vez disso, um comerciante pode desejar recomendar os  $k$  principais itens para um usuário específico ou determinar os  $k$  principais usuários a serem segmentados para um item específico. A determinação dos  $k$  principais itens é mais comum do que a determinação dos  $k$  principais usuários, embora os métodos nos dois casos sejam exatamente análogos. Ao longo deste livro, discutiremos apenas a determinação dos  $k$  principais itens, porque é a configuração mais comum. Este problema também é conhecido como o problema da recomendação dos  $k$  principais e é a formulação de classificação do problema da recomendação.

No segundo caso, os valores absolutos das classificações previstas não são importantes. A primeira formulação é mais geral, pois as soluções para o segundo caso podem ser derivadas resolvendo a primeira formulação para várias combinações usuário-item e, em seguida, classificando as previsões. No entanto, em muitos casos, é mais fácil e natural projetar métodos para resolver a versão de classificação do problema diretamente. Tais métodos serão discutidos no Capítulo 13.

Aumentar as vendas de produtos é o principal objetivo de um sistema de recomendação. Afinal, os sistemas de recomendação são utilizados por comerciantes para aumentar seus lucros. Ao recomendar itens cuidadosamente selecionados aos usuários, os sistemas de recomendação chamam a atenção dos usuários para itens relevantes. Isso aumenta o volume de vendas e os lucros do comerciante. Embora o objetivo principal de um sistema de recomendação seja aumentar a receita do comerciante, isso geralmente é alcançado de maneiras menos óbvias do que pode parecer à primeira vista. Para atingir o objetivo mais amplo, centrado no negócio, de aumentar a receita, os objetivos operacionais e técnicos comuns dos sistemas de recomendação são os seguintes:

1. Relevância: O objetivo operacional mais óbvio de um sistema de recomendação é recomendar itens relevantes para o usuário em questão. Os usuários têm maior probabilidade de consumir itens que considerem interessantes. Embora a relevância seja o principal objetivo operacional de um sistema de recomendação, ela não é suficiente por si só. Portanto, discutimos abaixo vários objetivos secundários, que não são tão importantes quanto a relevância, mas são importantes o suficiente para ter um impacto significativo.
2. Novidade: Sistemas de recomendação são realmente úteis quando o item recomendado é algo que o usuário nunca viu antes. Por exemplo, filmes populares de um gênero preferido raramente seriam novidade para o usuário. A recomendação repetida de itens populares também pode levar à redução da diversidade de vendas [203].
3. Serendipidade: Uma noção relacionada é a de serendipidade [229], em que os itens recomendados são algo inesperados e, portanto, há um elemento modesto de sorte

descoberta, em oposição a recomendações óbvias. A serendipidade difere da novidade, pois as recomendações são verdadeiramente surpreendentes para o usuário, em vez de simplesmente algo que ele desconhecia. Muitas vezes, um usuário específico pode estar consumindo apenas itens de um tipo específico, embora possa existir um interesse latente por itens de outros tipos, o que pode ser surpreendente para o próprio usuário. Ao contrário da novidade, os métodos serendipitosos concentram-se na descoberta de tais recomendações.

Por exemplo, se um novo restaurante indiano abre em um bairro, a recomendação desse restaurante a um usuário que normalmente come comida indiana é nova, mas não necessariamente fortuita. Por outro lado, quando o mesmo usuário recebe recomendação de comida etíope, e não sabia que tal comida poderia lhe interessar, a recomendação é fortuita. A serendipidade tem o efeito colateral benéfico de aumentar a diversidade de vendas ou iniciar uma nova tendência de interesse no usuário. O aumento da serendipidade geralmente traz benefícios estratégicos e de longo prazo para o comerciante, devido à possibilidade de descobrir áreas de interesse inteiramente novas. Por outro lado, algoritmos que fornecem recomendações fortuitas geralmente tendem a recomendar itens irrelevantes. Em muitos casos, os benefícios estratégicos e de longo prazo dos métodos fortuitos superam essas desvantagens de curto prazo.

4. Aumento da diversidade de recomendações: Os sistemas de recomendação normalmente sugerem uma lista dos  $k$  itens mais recomendados. Quando todos esses itens recomendados são muito semelhantes, aumenta o risco de o usuário não gostar de nenhum deles. Por outro lado, quando a lista de recomendações contém itens de tipos diferentes, há uma chance maior de o usuário gostar de pelo menos um desses itens. A diversidade tem a vantagem de garantir que o usuário não se canse de recomendações repetidas de itens semelhantes.

Além desses objetivos concretos, uma série de objetivos indiretos também são alcançados pelo processo de recomendação, tanto da perspectiva do usuário quanto do comerciante. Da perspectiva do usuário, as recomendações podem ajudar a melhorar a satisfação geral do usuário com o site. Por exemplo, um usuário que recebe repetidamente recomendações relevantes da Amazon.com ficará mais satisfeito com a experiência e terá maior probabilidade de usar o site novamente. Isso pode melhorar a fidelidade do usuário e aumentar ainda mais as vendas no site. Do lado do comerciante, o processo de recomendação pode fornecer insights sobre as necessidades do usuário e ajudar a personalizar ainda mais a experiência do usuário. Por fim, fornecer ao usuário uma explicação sobre por que um determinado item é recomendado costuma ser útil. Por exemplo, no caso da Netflix, as recomendações são fornecidas juntamente com filmes assistidos anteriormente. Como veremos mais adiante, alguns algoritmos de recomendação são mais adequados para fornecer explicações do que outros.

Há uma grande diversidade nos tipos de produtos recomendados por esses sistemas. Alguns sistemas de recomendação, como o Facebook, não recomendam produtos diretamente. Em vez disso, podem recomendar conexões sociais, o que traz um benefício indireto para o site, aumentando sua usabilidade e os lucros com publicidade. Para entender a natureza desses objetivos, discutiremos alguns exemplos populares de sistemas de recomendação históricos e atuais.

Esses exemplos também mostrarão a ampla diversidade de sistemas de recomendação que foram construídos como protótipos de pesquisa ou estão disponíveis hoje como sistemas comerciais em vários cenários de problemas.

## Sistema de recomendação do GroupLens

O GroupLens foi um sistema de recomendação pioneiro, desenvolvido como um protótipo de pesquisa para recomendação de notícias da Usenet. O sistema coletava avaliações de leitores da Usenet e as utilizava para prever se outros leitores gostariam ou não de um artigo antes de lê-lo.

Alguns dos primeiros algoritmos automatizados de filtragem colaborativa foram desenvolvidos no ambiente do GroupLens<sup>1</sup>. As ideias gerais desenvolvidas por esse grupo também foram estendidas a outros ambientes de produtos, como livros e filmes. Os sistemas de recomendação correspondentes foram chamados de BookLens e MovieLens, respectivamente. Além de suas contribuições pioneiras para a pesquisa de filtragem colaborativa, a equipe de pesquisa do GroupLens se destacou por divulgar diversos conjuntos de dados durante os primeiros anos desse campo, quando os conjuntos de dados não estavam facilmente disponíveis para benchmarking. Exemplos proeminentes incluem três conjuntos de dados [688] do sistema de recomendação MovieLens. Esses conjuntos de dados têm tamanho sucessivamente crescente e contêm 105, 106 e 107 avaliações, respectivamente.

## Sistema de recomendação da Amazon.com

A Amazon.com [698] também foi uma das pioneiras em sistemas de recomendação, especialmente no ambiente comercial. Durante os primeiros anos, foi um dos poucos varejistas que teve a visão de perceber a utilidade dessa tecnologia. Originalmente fundada como uma livraria eletrônica, o negócio se expandiu para praticamente todos os tipos de produtos. Consequentemente, a Amazon.com agora vende praticamente todas as categorias de produtos, como livros, CDs, software, eletrônicos e assim por diante. As recomendações na Amazon.com são fornecidas com base em classificações explicitamente fornecidas, comportamento de compra e comportamento de navegação. As classificações na Amazon.com são especificadas em uma escala de 5 pontos, com a classificação mais baixa sendo 1 estrela e a classificação mais alta sendo 5 estrelas. Os dados de compra e navegação específicos do cliente podem ser facilmente coletados quando os usuários estão conectados com um mecanismo de autenticação de conta suportado pela Amazon. As recomendações também são fornecidas aos usuários na página principal do site, sempre que eles acessam suas contas. Em muitos casos, explicações para as recomendações são fornecidas. Por exemplo, a relação de um item recomendado com itens comprados anteriormente pode ser incluída na interface do sistema de recomendação.

O comportamento de compra ou navegação de um usuário pode ser visto como um tipo de classificação implícita, em oposição a uma classificação explícita, que é especificada pelo usuário. Muitos sistemas comerciais permitem a flexibilidade de fornecer recomendações com base em feedback explícito e implícito. De fato, vários modelos foram projetados (cf. seção 3.6.4.6 do Capítulo 3) para considerar conjuntamente o feedback explícito e implícito no processo de recomendação. Alguns dos algoritmos utilizados pelas primeiras versões do sistema de recomendação da Amazon.com são discutidos em [360].

## Sistema de recomendação de filmes da Netflix

A Netflix foi fundada como uma empresa de aluguel de discos de vídeo digitais (DVDs) por correspondência [690] de filmes e programas de televisão, que eventualmente foi expandida para entrega em streaming. Atualmente, o principal negócio da Netflix é fornecer entrega em streaming de filmes e programas de televisão por assinatura. A Netflix oferece aos usuários a capacidade de classificar os filmes e programas de televisão em uma escala de 5 pontos. Além disso, as ações do usuário em termos de assistir a vários itens também são armazenadas pela Netflix. Essas classificações e ações são então usadas pela Netflix para fazer recomendações. A Netflix faz um excelente trabalho ao fornecer explicações para os itens recomendados. Ela fornece explicitamente exemplos de recomendações com base em itens específicos que foram assistidos pelo usuário. Essas informações fornecem ao usuário informações adicionais

---

<sup>1</sup> O termo "GroupLens" refere-se atualmente ao grupo acadêmico da Universidade de Minnesota [687] que desenvolveu esses algoritmos. Este grupo continua a trabalhar na área de sistemas de recomendação e fez muitas contribuições pioneiras ao longo dos anos.

informações para decidir se assiste ou não a um filme específico. Apresentar explicações significativas é importante para que o usuário entenda por que ele pode achar um determinado filme interessante. Essa abordagem também aumenta a probabilidade de o usuário agir de acordo com a recomendação e melhora verdadeiramente a experiência do usuário. Esse tipo de abordagem interessante também pode ajudar a aumentar a fidelidade e a retenção de clientes.

A Netflix contribuiu significativamente para a comunidade de pesquisa por meio do concurso Prêmio Netflix. Este concurso foi criado para fornecer um fórum de competição entre diversos algoritmos de filtragem colaborativa contribuídos pelos participantes. Um conjunto de dados de avaliações de filmes da Netflix foi divulgado, e a tarefa era prever as avaliações de combinações específicas de itens e usuários.

Para esse fim, a Netflix forneceu um conjunto de dados de treinamento e um conjunto de dados de qualificação. O conjunto de dados de treinamento continha 100.480.507 avaliações de 480.189 usuários para 17.770 filmes.

O conjunto de treinamento incluiu um conjunto de sondas menor contendo 1.408.395 avaliações. O conjunto de sondas foi baseado em avaliações mais recentes do que os dados de treinamento restantes e foi estatisticamente semelhante à parte do conjunto de dados com avaliações ocultas. Essa parte do conjunto de dados foi chamada de conjunto de dados de qualificação e continha mais de 2.817.131 tripletos do formato

Usuário, Filme, Data da Avaliação. Observe que o tripleto não continha a classificação real, que era conhecida apenas pelos juízes. Os usuários precisavam prever as avaliações no conjunto de dados de qualificação com base em modelos dos dados de treinamento. Essa previsão foi pontuada pelos juízes (ou um sistema automatizado equivalente), e os usuários foram (continuamente) informados dos resultados da previsão em apenas metade do conjunto de dados de qualificação no quadro de líderes. Essa metade do conjunto de dados de qualificação foi chamada de conjunto de questionários. A metade restante foi usada como conjunto de teste para calcular a pontuação final e determinar os vencedores dos prêmios. As pontuações da metade restante nunca foram reveladas aos usuários até o final. Além disso, não foi revelado aos competidores quais dos trios do conjunto classificatório pertenciam ao conjunto de quiz e quais pertenciam ao conjunto de teste. A razão para esse arranjo incomum no conjunto de teste foi garantir que os usuários não aproveitassem as pontuações na tabela de classificação para superajustar seus algoritmos ao conjunto de teste. Questões relacionadas ao superajuste serão descritas no Capítulo 7, sobre algoritmos de avaliação. De fato, a estrutura da Netflix para lidar com as inscrições dos competidores é um excelente exemplo de design de avaliação adequado de algoritmos de recomendação.

O conjunto de sondas, o conjunto de questionários e o conjunto de testes foram projetados para apresentar características estatísticas semelhantes. Os prêmios foram concedidos com base na melhoria do algoritmo de recomendação da Netflix, conhecido como Cinematch, ou na melhoria da melhor pontuação anterior em um determinado limite. Muitos algoritmos de recomendação conhecidos, como modelos de fatores latentes, foram popularizados pelo concurso da Netflix. O concurso Netflix Prize é notável por suas inúmeras contribuições à pesquisa de recomendação [71, 373].

## Sistema de Personalização do Google Notícias

O sistema de personalização do Google Notícias [697] é capaz de recomendar notícias aos usuários com base em seu histórico de cliques. Os cliques são associados a usuários específicos com base em mecanismos de identificação habilitados pelas contas do Gmail. Nesse caso, os artigos de notícias são tratados como itens. O ato de um usuário clicar em um artigo de notícias pode ser visto como uma avaliação positiva para esse artigo. Essas classificações podem ser vistas como classificações unárias, nas quais existe um mecanismo para um usuário expressar sua afinidade por um item, mas não existe nenhum mecanismo para ele mostrar sua antipatia. Além disso, as classificações são implícitas, pois são inferidas a partir de ações do usuário, em vez de serem explicitamente especificadas por ele. No entanto, variações da abordagem também podem ser aplicadas a casos em que as classificações são explicitamente especificadas. Algoritmos de recomendação colaborativa são aplicados às classificações coletadas, para que inferências possam ser feitas sobre o

Tabela 1.1: Exemplos de produtos recomendados por vários sistemas de recomendação do mundo real

Sistema	Objetivo do produto
Amazon.com [698]	Livros e outros produtos
Netflix [690]	DVDs, streaming de vídeo
Bobo da corte [689]	Piadas
GroupLens [687]	Notícias
MovieLens [688]	Filmes
last.fm [692]	Música
Google Notícias [697]	Notícias
Pesquisa Google [696]	Anúncios
Facebook [691]	Amigos, Anúncios
Pandora [693]	Música
YouTube [694]	Vídeos online
TripAdvisor [695]	Produtos de viagem
IMDb [699]	Filmes

artigos personalizados para usuários específicos. Uma descrição de um sistema de filtragem colaborativa para O Google Notícias é fornecido em [175]. Mais detalhes sobre o mecanismo de personalização do Google Notícias são discutidos na seção 13.8.1.2 do Capítulo 13.

Recomendações de amigos do Facebook

Os sites de redes sociais geralmente recomendam amigos em potencial aos usuários para aumentar o número de conexões sociais no site. O Facebook [691] é um exemplo de site de rede social. Este tipo de recomendação tem objetivos ligeiramente diferentes dos uma recomendação de produto. Enquanto uma recomendação de produto aumenta diretamente o lucro do comerciante, ao facilitar a venda de produtos, o aumento do número de conexões sociais melhora a experiência do usuário em uma rede social. Isso, por sua vez, incentiva a crescimento da rede social. As redes sociais dependem fortemente do crescimento da rede para aumentar suas receitas de publicidade. Portanto, a recomendação de amigos em potencial (ou links) permite um melhor crescimento e conectividade da rede. Este problema também é conhecido como previsão de links no campo da análise de redes sociais. Essas formas de As recomendações são baseadas em relações estruturais e não em dados de classificação. Portanto, a natureza dos algoritmos subjacentes é completamente diferente. A recomendação do link O problema é explorado em detalhes no Capítulo 10. A relação da publicidade computacional a tecnologia do sistema de recomendação é discutida no Capítulo 13.

1.2.1 O espectro de aplicações de recomendação

A seguir, forneceremos uma breve visão geral dos objetivos específicos da aplicação alcançados por diversas implementações de sistemas de recomendação. Uma breve visão geral dos produtos sugeridos e dos objetivos alcançados por diversos sistemas de recomendação são ilustrados. na Tabela 1.1. Muitos desses sistemas de recomendação são focados no comércio eletrônico tradicional aplicações para vários produtos, incluindo livros, filmes, vídeos, viagens e outros bens e serviços. A aplicabilidade mais ampla dos sistemas de recomendação a aplicações de comércio eletrônico é discutido em [530]. No entanto, os sistemas de recomendação expandiram-se para além do tradicional domínio das recomendações de produtos. Vale ressaltar que alguns dos sistemas da Tabela 1.1

pode não recomendar produtos específicos. Um exemplo é o aplicativo de busca Google, que pode anunciar produtos junto com seus resultados de busca. Esta é a área da publicidade computacional, que é uma área distinta por si só, mas, ainda assim, está intimamente relacionada aos sistemas de recomendação. Esta área é discutida em detalhes na seção 13.8.2 do Capítulo 13.

Da mesma forma, o Facebook recomenda amigos, e sites de recrutamento online recomendam empregadores e candidatos a emprego uns aos outros. Este último sistema também é conhecido como recomendação recíproca. Os modelos para alguns desses algoritmos de recomendação são bem diferentes daqueles dos sistemas de recomendação tradicionais. Este livro estudará muitas dessas variações em detalhes.

## 1.3 Modelos básicos de sistemas de recomendação

---

Os modelos básicos para sistemas de recomendação trabalham com dois tipos de dados, que são (i) as interações usuário-item, como avaliações ou comportamento de compra, e (ii) as informações de atributos sobre os usuários e itens, como perfis textuais ou palavras-chave relevantes. Os métodos que usam os primeiros são chamados de métodos de filtragem colaborativa, enquanto os métodos que usam os últimos são chamados de métodos de recomendação baseados em conteúdo. Observe que os sistemas baseados em conteúdo também usam as matrizes de classificação na maioria dos casos, embora o modelo geralmente se concentre nas avaliações de um único usuário, em vez das de todos os usuários. Em sistemas de recomendação baseados em conhecimento, as recomendações são baseadas em requisitos do usuário explicitamente especificados. Em vez de usar dados históricos de avaliação ou compra, bases de conhecimento externas e restrições são usadas para criar a recomendação. Alguns sistemas de recomendação combinam esses diferentes aspectos para criar sistemas híbridos. Sistemas híbridos podem combinar os pontos fortes de vários tipos de sistemas de recomendação para criar técnicas que podem ter um desempenho mais robusto em uma ampla variedade de cenários.

A seguir, discutiremos brevemente esses modelos básicos e também forneceremos indicações para os capítulos relevantes do livro onde eles são discutidos.

### 1.3.1 Modelos de filtragem colaborativa

Modelos de filtragem colaborativa utilizam o poder colaborativo das avaliações fornecidas por vários usuários para fazer recomendações. O principal desafio no desenvolvimento de métodos de filtragem colaborativa é que as matrizes de classificação subjacentes são esparsas. Considere o exemplo de um aplicativo de filme no qual os usuários especificam avaliações que indicam se gostam ou não de filmes específicos.

A maioria dos usuários teria assistido apenas a uma pequena fração do grande universo de filmes disponíveis. Como resultado, a maioria das classificações não é especificada. As classificações especificadas também são chamadas de classificações observadas. Ao longo deste livro, os termos "especificado" e "observado" serão usados de forma intercambiável. As classificações não especificadas serão chamadas de "não observadas" ou "ausentes".

A ideia básica dos métodos de filtragem colaborativa é que essas classificações não especificadas podem ser imputadas porque as classificações observadas são frequentemente altamente correlacionadas entre vários usuários e itens. Por exemplo, considere dois usuários chamados Alice e Bob, que têm gostos muito semelhantes. Se as classificações, que ambos especificaram, forem muito semelhantes, então sua similaridade pode ser identificada pelo algoritmo subjacente. Nesses casos, é muito provável que as classificações nas quais apenas um deles especificou um valor também sejam semelhantes. Essa similaridade pode ser usada para fazer inferências sobre valores especificados de forma incompleta. A maioria dos modelos para filtragem colaborativa concentra-se em alavancar correlações entre itens ou entre usuários para o processo de predição. Alguns modelos usam ambos os tipos de correlações. Além disso, alguns modelos usam técnicas de otimização cuidadosamente projetadas para criar um modelo de treinamento com praticamente o mesmo



A maneira como um classificador cria um modelo de treinamento a partir dos dados rotulados. Esse modelo é então usado para imputar os valores ausentes na matriz, da mesma forma que um classificador imputa os rótulos de teste ausentes. Existem dois tipos de métodos comumente usados na filtragem colaborativa, chamados de métodos baseados em memória e métodos baseados em modelo:

1. Métodos baseados em memória: Métodos baseados em memória também são chamados de algoritmos de filtragem colaborativa baseados em vizinhança. Estes estavam entre os primeiros algoritmos de filtragem colaborativa, nos quais as classificações de combinações usuário-item são previstas com base em suas vizinhanças. Essas vizinhanças podem ser definidas de duas maneiras:
  - Filtragem colaborativa baseada no usuário: neste caso, as classificações fornecidas por usuários com ideias semelhantes de um usuário-alvo A são usadas para fazer recomendações para A. Assim, a ideia básica é determinar usuários semelhantes ao usuário-alvo A e recomendar classificações para as avaliações não observadas de A, calculando médias ponderadas das avaliações desse grupo de pares. Portanto, se Alice e Bob avaliaram filmes de forma semelhante no passado, é possível usar as avaliações observadas de Alice no filme Exterminador do Futuro para prever as avaliações não observadas de Bob nesse filme. Em geral, os k usuários mais semelhantes a Bob podem ser usados para fazer previsões de classificação para Bob. Funções de similaridade são calculadas entre as linhas da matriz de avaliações para descobrir usuários semelhantes.
  - Filtragem colaborativa baseada em itens: Para fazer as previsões de classificação para o item alvo B pelo usuário A, o primeiro passo é determinar um conjunto S de itens que são mais semelhantes ao item alvo B. As classificações no conjunto de itens S, especificadas por A, são usadas para prever se o usuário A gostará do item B. Portanto, as classificações de Bob em filmes de ficção científica semelhantes, como Alien e Predador, podem ser usadas para prever sua classificação em O Exterminador do Futuro. Funções de similaridade são calculadas entre as colunas da matriz de classificações para descobrir itens semelhantes.

As vantagens das técnicas baseadas em memória são a simplicidade de implementação e a facilidade de explicação das recomendações resultantes. Por outro lado, algoritmos baseados em memória não funcionam muito bem com matrizes de classificação esparsas. Por exemplo, pode ser difícil encontrar usuários suficientemente semelhantes a Bob, que tenham avaliado Gladiador. Nesses casos, é difícil prever com precisão a classificação de Bob para Gladiador. Em outras palavras, esses métodos podem não cobrir completamente as previsões de classificação. No entanto, a falta de cobertura geralmente não é um problema quando apenas os k itens mais importantes são necessários.

Os métodos baseados em memória são discutidos em detalhes no Capítulo 2.

2. Métodos baseados em modelos: Em métodos baseados em modelos, métodos de aprendizado de máquina e mineração de dados são usados no contexto de modelos preditivos. Nos casos em que o modelo é parametrizado, os parâmetros deste modelo são aprendidos dentro do contexto de uma estrutura de otimização. Alguns exemplos de tais métodos baseados em modelos incluem árvores de decisão, modelos baseados em regras, métodos bayesianos e modelos de fatores latentes. Muitos desses métodos, como os modelos de fatores latentes, apresentam um alto nível de cobertura mesmo para matrizes de classificação esparsas. Algoritmos de filtragem colaborativa baseados em modelos são discutidos no Capítulo 3.

Embora os algoritmos de filtragem colaborativa baseados em memória sejam valorizados por sua simplicidade, eles tendem a ser heurísticos por natureza e não funcionam bem em todos os cenários. No entanto, a distinção entre métodos baseados em memória e métodos baseados em modelos é um tanto artificial, pois



Figura 1.1: Exemplo de classificações de intervalo de 5 pontos

Classificações gerais

- 1. A qualidade do conteúdo do curso
- 2. O ensino geral do instrutor



Figura 1.2: Exemplo de classificações ordinais usadas em avaliações de cursos da Universidade de Stanford

métodos baseados em memória também podem ser considerados modelos baseados em similaridade, embora sejam heurísticos. Na seção 2.6 do Capítulo 2, também será demonstrado que algumas variações de métodos baseados em vizinhança podem ser formalmente expressas como modelos baseados em regressão. Modelos de fatores latentes foram popularizados anos mais tarde como resultado do concurso Netflix Prize, embora algoritmos semelhantes tenham sido propostos muito antes no contexto de conjuntos de dados incompletos (genéricos) [24]. Recentemente, foi demonstrado que algumas combinações de métodos baseados em memória e em modelos [309] fornecem resultados muito precisos.

1.3.1.1 Tipos de classificações

O design dos algoritmos de recomendação é influenciado pelo sistema utilizado para rastrear as classificações. As classificações são frequentemente especificadas em uma escala que indica o nível específico de gosto ou desgosto do item em questão. É possível que as classificações sejam valores contínuos, como no caso do mecanismo de recomendação de piadas Jester [228, 689], no qual as classificações podem assumir qualquer valor entre -10 e 10. Isso, no entanto, é relativamente raro. Normalmente, as classificações são baseadas em intervalos, onde um conjunto discreto de números ordenados é usado para quantificar o gosto ou desgosto.

Essas classificações são chamadas de classificações baseadas em intervalos. Por exemplo, uma escala de classificação de 5 pontos pode ser extraída do conjunto  $\{\bar{y}_2, \bar{y}_1, 0, 1, 2\}$ , em que uma classificação de  $\bar{y}_2$  indica uma aversão extrema e uma classificação de 2 indica uma forte afinidade com o item. Outros sistemas podem extrair as classificações do conjunto  $\{1, 2, 3, 4, 5\}$ .

O número de classificações possíveis pode variar de acordo com o sistema em questão. O uso de classificações de 5, 7 e 10 pontos é particularmente comum. O sistema de classificação de 5 estrelas, ilustrado na Figura 1.1, é um exemplo de classificação por intervalo. Ao longo de cada uma das classificações possíveis, indicamos uma interpretação semântica do nível de interesse do usuário. Essa interpretação pode variar ligeiramente entre diferentes comerciantes, como Amazon ou Netflix. Por exemplo, a Netflix usa um sistema de classificação de 5 estrelas em que o ponto de 4 estrelas corresponde a "gostei muito" e o ponto central de 3 estrelas corresponde a "gostei". Portanto, há três classificações favoráveis e duas classificações desfavoráveis na Netflix, o que leva a uma escala de classificação desequilibrada. Em alguns

Em alguns casos, pode haver um número par de classificações possíveis, e a classificação neutra pode estar ausente. Essa abordagem é chamada de sistema de classificação de escolha forçada.

Também é possível usar valores categóricos ordenados, como {Discordo Totalmente, Discordo, Neutro, Concordo, Concordo Totalmente}, para atingir os mesmos objetivos. Em geral, essas classificações são chamadas de classificações ordinais, e o termo deriva do conceito de atributos ordinais. Um exemplo de classificações ordinais, usadas nos formulários de avaliação de cursos da Universidade Stanford, é ilustrado na Figura 1.2. Em classificações binárias, o usuário pode representar apenas uma preferência ou desgosto pelo item e nada mais. Por exemplo, as classificações podem ser 0, 1 ou valores não especificados. Os valores não especificados precisam ser previstos para valores de 0 a 1. Um caso especial de classificações é o das classificações unárias, nas quais há um mecanismo para um usuário especificar uma preferência por um item, mas nenhum mecanismo para especificar uma desgosto. Classificações unárias são particularmente comuns, especialmente no caso de conjuntos de dados de feedback implícito [259, 260, 457]. Nesses casos, as preferências dos clientes são derivadas de suas atividades e não de suas classificações explicitamente especificadas. Por exemplo, o comportamento de compra de um cliente pode ser convertido em avaliações unárias. Quando um cliente compra um item, isso pode ser visto como uma preferência pelo item. No entanto, o ato de não comprar um item dentro um amplo universo de possibilidades nem sempre indica uma aversão. Da mesma forma, muitas redes sociais, como o Facebook, utilizam botões de "curtir", que permitem expressar apreço por um item. No entanto, não existe um mecanismo para especificar a não preferência por um item. A configuração de feedback implícito pode ser vista como o análogo de complementação matricial do problema de aprendizagem positiva não rotulada (PU) na classificação de dados [259].

#### Exemplos de classificações explícitas e implícitas

Um exemplo quantitativo de avaliações explícitas é ilustrado na Figura 1.3(a). Neste caso, há 6 usuários, rotulados como U1 ... U6, e 6 filmes com títulos especificados. Avaliações mais altas indicam feedback mais positivo na Figura 1.3(a). As entradas ausentes correspondem a preferências não especificadas. O exemplo desta figura representa um pequeno exemplo de brinquedo. Em geral, as classificações podem ser representadas como uma matriz  $m \times n$ , onde  $m$  e  $n$  são tipicamente muito grandes e podem variar na ordem de centenas de milhares. Embora este exemplo específico utilize uma matriz  $6 \times 6$ , os valores de  $m$  e  $n$  normalmente não são os mesmos em cenários do mundo real. Uma matriz de classificações é às vezes chamada de matriz de utilidade, embora as duas nem sempre sejam iguais.

A rigor, quando a utilidade se refere ao valor do lucro, a utilidade de uma combinação usuário-item se refere ao valor do lucro obtido ao recomendar aquele item a um usuário específico. Embora as matrizes de utilidade sejam frequentemente definidas como as mesmas que as matrizes de classificação, é possível que o aplicativo transforme explicitamente as classificações em valores de utilidade com base em critérios específicos do domínio. Todos os algoritmos de filtragem colaborativa são então aplicados à matriz de utilidade em vez da matriz de classificação. No entanto, essa abordagem raramente é usada na prática, e a maioria dos algoritmos de filtragem colaborativa trabalha diretamente com a matriz de classificação.

Um exemplo de matriz de classificações unárias é ilustrado na Figura 1.3(b). Para os casos em que as classificações são unárias, a matriz é chamada de matriz de utilidade de preferência positiva, pois permite apenas a especificação de preferências positivas. As duas matrizes na Figura 1.3 têm o mesmo conjunto de entradas observadas, mas fornecem insights muito diferentes. Por exemplo, os usuários U1 e U3 são muito diferentes na Figura 1.3(a) porque têm classificações muito diferentes para suas entradas mutuamente especificadas. Por outro lado, esses usuários seriam considerados muito semelhantes na Figura 1.3(b) porque expressaram uma preferência positiva pelos mesmos itens. A utilidade baseada em classificações fornece uma maneira para os usuários expressarem preferências negativas por itens. Por exemplo, o usuário U1 não gosta do filme Gladiador na Figura 1.3(a). Não há mecanismo para especificar isso na matriz de utilidade de preferência positiva da Figura 1.3(b) além de

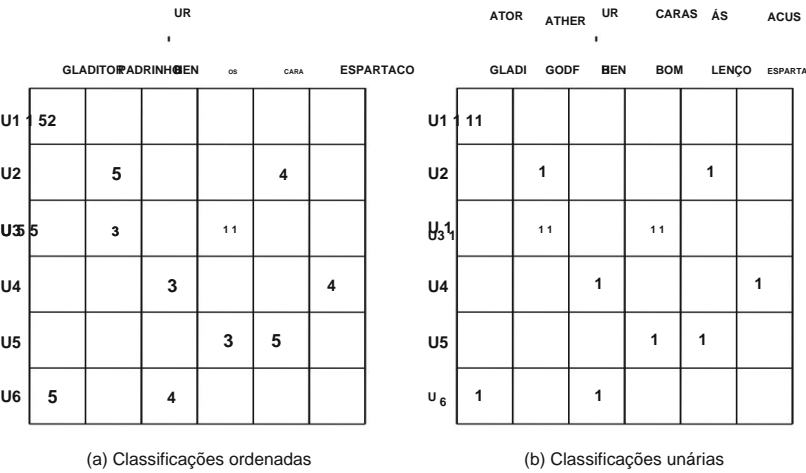


Figura 1.3: Exemplos de matrizes de utilidade

uma entrada ausente relativamente ambígua. Em outras palavras, a matriz na Figura 1.3(b) é menos expressiva. Embora a Figura 1.3(b) forneça um exemplo de matriz binária, é possível que as entradas diferentes de zero sejam valores positivos arbitrários. Por exemplo, elas podem corresponder às quantidades de itens comprados pelos diferentes usuários. Em geral, matrizes unárias são criadas por ações do usuário, como a compra de um item, e, portanto, também são chamadas de matrizes de feedback implícito.

Classificações unárias têm um efeito significativo no algoritmo de recomendação em questão, porque não há informações disponíveis sobre se um usuário não gosta de um item. No caso de matrizes unárias, muitas vezes é recomendado [260] realizar a análise de forma simples, tratando as entradas ausentes como 0s na fase inicial. No entanto, o valor final previsto pelo algoritmo de aprendizagem pode ser muito maior que 0, especialmente se o item corresponder aos interesses do usuário. Os itens recomendados são, portanto, baseados nas entradas com o maior erro de previsão positivo sobre a suposição inicial de "zero". De fato, se as entradas ausentes não forem substituídas por 0s, é possível um sobreajuste significativo. Esse tipo de sobreajuste é um artefato do fato de que muitas vezes não há um nível suficiente de discriminação entre os vários valores observados das classificações. Em matrizes de feedback explícitas, as classificações correspondem a preferências (altamente discriminadas), enquanto em matrizes de feedback implícitas, as classificações correspondem a confidências (menos discriminadas). Em um capítulo posterior, forneceremos um exemplo específico de overfitting com matrizes de feedback implícitas quando entradas ausentes não são tratadas como zeros (cf. seção 3.6.6.2 do Capítulo 3).

A pré-substituição de classificações ausentes não é recomendada em matrizes de classificações explícitas. Em matrizes de classificações explícitas com gostos e desgostos, a substituição de entradas ausentes por qualquer valor (como 0 ou a média da linha/coluna/dados) sempre leva a um viés significativo na análise. No caso unário, a substituição de entradas ausentes por 0s também leva a algum viés [457, 467, 468], embora frequentemente pequeno, pois a suposição padrão em dados de feedback implícito, como dados de compra, é que o usuário não comprará a maioria dos itens.

Frequentemente, aceita-se conviver com esse viés no caso unário, pois uma quantidade significativa de sobreajuste é reduzida pela substituição. Há também alguns efeitos computacionais interessantes dessas escolhas. Essas compensações são discutidas nos Capítulos 2 e 3.

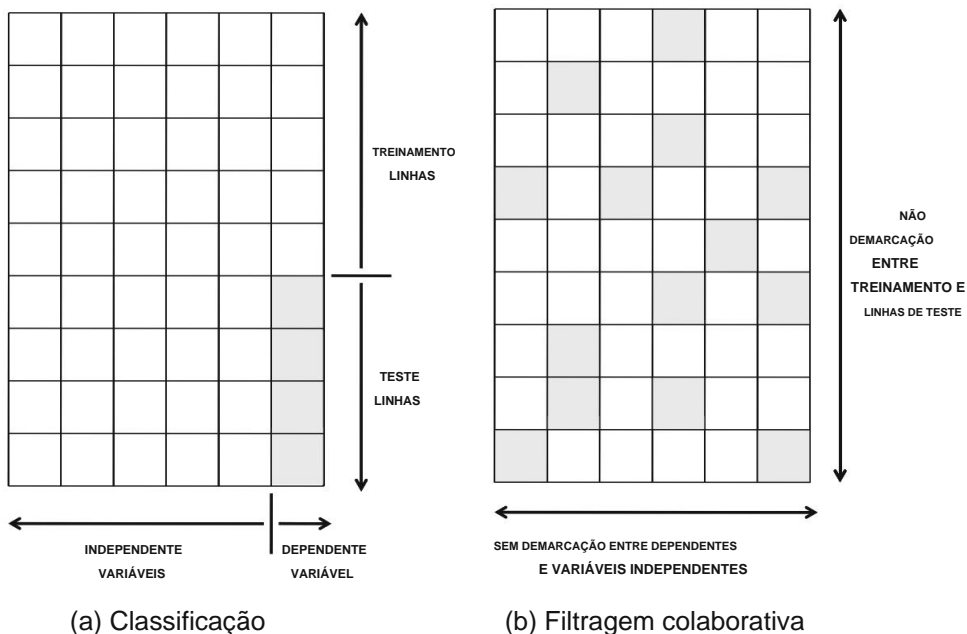


Figura 1.4: Comparando o problema de classificação tradicional com a filtragem colaborativa. Entradas sombreadas estão faltando e precisam ser previstas.

#### 1.3.1.2 Relação com Análise de Valor Ausente

Modelos de filtragem colaborativa estão intimamente relacionados à análise de valores ausentes. A literatura tradicional sobre análise de valores ausentes estuda o problema da imputação de entradas em uma matriz de dados incompletamente especificada. A filtragem colaborativa pode ser vista como um caso especial (difícil) desse problema, no qual a matriz de dados subjacente é muito grande e esparsa. Uma discussão detalhada dos métodos para análise de valores ausentes na literatura estatística pode ser encontrada em [362]. Muitos desses métodos também podem ser usados para sistemas de recomendação, embora alguns deles possam exigir adaptações especializadas para matrizes muito grandes e esparsas. De fato, algumas das classes recentes de modelos para sistemas de recomendação, como modelos de fatores latentes, foram estudadas anteriormente no contexto da análise de valores ausentes [24]. Métodos semelhantes foram propostos independentemente no contexto de sistemas de recomendação [252, 309, 313, 500, 517, 5]. Em geral, muitos métodos clássicos de estimativa de valores ausentes [362] também podem ser usados para filtragem colaborativa.

#### 1.3.1.3 Filtragem colaborativa como uma generalização da classificação e modelagem de regressão

Os métodos de filtragem colaborativa podem ser vistos como generalizações da modelagem de classificação e regressão. Nos problemas de classificação e regressão, a classe/variável dependente pode ser vista como um atributo com valores ausentes. Outras colunas são tratadas como características/variáveis independentes. O problema de filtragem colaborativa pode ser visto como uma generalização desta estrutura, pois qualquer coluna pode ter valores ausentes, em vez de (apenas) a variável de classe. No problema de recomendação, uma distinção clara não é necessária.

não existe entre variáveis de classe e variáveis de característica porque cada característica desempenha o papel duplo de variável dependente e independente. Essa distinção existe no problema de classificação apenas porque as entradas ausentes são restritas a uma coluna especial. Além disso, não há distinção entre linhas de treinamento e teste na filtragem colaborativa, pois qualquer linha pode conter entradas ausentes. Portanto, é mais significativo falar de entradas de treinamento e teste na filtragem colaborativa do que de linhas de treinamento e teste. A filtragem colaborativa é uma generalização da modelagem de classificação/regressão na qual a predição é realizada de forma entrada a entrada, em vez de linha a linha. Essa relação entre modelagem de classificação/regressão e filtragem colaborativa é importante de se ter em mente, pois muitos princípios dos métodos de classificação e modelagem de regressão podem ser generalizados para sistemas de recomendação. A relação entre os dois problemas é ilustrada na Figura 1.4. Esta figura é particularmente útil para relacionar a filtragem colaborativa com a classificação e será revisitada várias vezes neste livro, sempre que as semelhanças entre esses dois problemas forem aproveitadas de alguma forma para o desenvolvimento algorítmico ou teórico.

O problema de complementação de matrizes também compartilha uma série de características com o cenário transdutivo em classificação e regressão. No cenário transdutivo, as instâncias de teste também são incluídas no processo de treinamento (tipicamente com o uso de um algoritmo semissupervisionado), e muitas vezes é difícil fazer previsões para instâncias de teste que não estão disponíveis no momento do treinamento. Por outro lado, modelos nos quais previsões podem ser facilmente feitas para novas instâncias são chamados de indutivos. Por exemplo, um modelo Bayesiano ingênuo em classificação é inerentemente indutivo porque pode ser facilmente usado para prever o rótulo de uma instância de teste para a qual as características não eram conhecidas no momento da construção do modelo Bayesiano.

A configuração para o preenchimento de matrizes é inerentemente transdutiva, pois os dados de treinamento e teste são fortemente integrados entre si na matriz de classificações  $m \times n \times R$ , e muitos modelos não conseguem prever facilmente as classificações para usuários e/ou itens fora da amostra. Por exemplo, se John for adicionado à matriz de classificações (com muitas classificações especificadas) após o modelo de filtragem colaborativa já ter sido construído, muitos métodos prontos para uso não serão capazes de fazer previsões para John. Isso é especialmente verdadeiro para métodos de filtragem colaborativa baseados em modelos. No entanto, alguns modelos recentes de preenchimento de matrizes também foram projetados para serem indutivos, nos quais as classificações podem ser previstas para usuários e/ou itens fora da amostra.

### 1.3.2 Sistemas de recomendação baseados em conteúdo

Em sistemas de recomendação baseados em conteúdo, os atributos descritivos dos itens são usados para fazer recomendações. O termo "conteúdo" refere-se a essas descrições. Em métodos baseados em conteúdo, as avaliações e o comportamento de compra dos usuários são combinados com as informações de conteúdo disponíveis nos itens. Por exemplo, considere uma situação em que John avaliou o filme O Exterminador do Futuro com uma nota alta, mas não temos acesso às avaliações de outros usuários. Portanto, métodos de filtragem colaborativa são descartados. No entanto, a descrição do item O Exterminador do Futuro contém palavras-chave de gênero semelhantes às de outros filmes de ficção científica, como Alien e Predador. Nesses casos, esses filmes podem ser recomendados a John.

Em métodos baseados em conteúdo, as descrições dos itens, que são rotuladas com classificações, são usadas como dados de treinamento para criar um problema de classificação ou modelagem de regressão específico do usuário. Para cada usuário, os documentos de treinamento correspondem às descrições dos itens que ele comprou ou avaliou. A variável de classe (ou dependente) corresponde às avaliações ou comportamento de compra especificados. Esses documentos de treinamento são usados para criar um modelo de classificação ou regressão específico para o usuário em questão (ou usuário ativo). Esse modelo específico do usuário é usado para prever se o indivíduo correspondente gostará de um item para o qual sua avaliação ou comportamento de compra é desconhecido.

Métodos baseados em conteúdo apresentam algumas vantagens na hora de fazer recomendações para novos itens quando não há dados de classificação suficientes disponíveis para o item em questão. Isso ocorre porque outros itens com atributos semelhantes podem ter sido avaliados pelo usuário ativo. Portanto, o modelo supervisionado poderá utilizar essas classificações em conjunto com os atributos do item para fazer recomendações mesmo quando não houver histórico de classificações para o item em questão.

Os métodos baseados em conteúdo também apresentam diversas desvantagens:

1. Em muitos casos, métodos baseados em conteúdo fornecem recomendações óbvias devido ao uso de palavras-chave ou conteúdo. Por exemplo, se um usuário nunca consumiu um item com um conjunto específico de palavras-chave, tal item não tem chance de ser recomendado. Isso ocorre porque o modelo construído é específico para o usuário em questão e o conhecimento da comunidade de usuários semelhantes não é aproveitado. Esse fenômeno tende a reduzir a diversidade dos itens recomendados, o que é indesejável.
2. Embora os métodos baseados em conteúdo sejam eficazes para fornecer recomendações de novos itens, eles não são eficazes para fornecer recomendações para novos usuários. Isso ocorre porque o modelo de treinamento para o usuário-alvo precisa usar o histórico de suas avaliações. Na verdade, geralmente é importante ter um grande número de classificações disponíveis para o usuário-alvo, a fim de fazer previsões robustas sem overfitting.

Portanto, os métodos baseados em conteúdo têm compensações diferentes dos sistemas de filtragem colaborativa.

Embora a descrição acima mencionada forneça a visão convencional baseada em aprendizagem de métodos baseados em conteúdo, uma visão mais ampla desses métodos às vezes é usada. Por exemplo, os usuários podem especificar palavras-chave relevantes em seus próprios perfis. Esses perfis podem ser combinados com descrições de itens para fazer recomendações. Tal abordagem não usa classificações no processo de recomendação e, portanto, é útil em cenários de inicialização a frio. No entanto, tais métodos são frequentemente vistos como uma classe distinta de sistemas de recomendação, conhecidos como sistemas baseados em conhecimento, porque as métricas de similaridade são frequentemente baseadas no conhecimento do domínio principal. Os sistemas de recomendação baseados em conhecimento são frequentemente considerados intimamente relacionados aos sistemas de recomendação baseados em conteúdo, e às vezes é questionado se existe uma demarcação clara entre as duas classes de métodos [558]. Métodos para sistemas de recomendação baseados em conteúdo são discutidos no Capítulo 4.

### 1.3.3 Sistemas de recomendação baseados em conhecimento

Sistemas de recomendação baseados em conhecimento são particularmente úteis no contexto de itens que não são comprados com muita frequência. Exemplos incluem itens como imóveis, automóveis, solicitações de turismo, serviços financeiros ou bens de luxo caros. Nesses casos, classificações suficientes podem não estar disponíveis para o processo de recomendação. Como os itens são comprados raramente e com diferentes tipos de opções detalhadas, é difícil obter um número suficiente de classificações para uma instanciação específica (ou seja, combinação de opções) do item em questão. Esse problema também é encontrado no contexto do problema de inicialização a frio, quando classificações suficientes não estão disponíveis para o processo de recomendação. Além disso, a natureza das preferências do consumidor pode evoluir ao longo do tempo ao lidar com esses itens. Por exemplo, o modelo de um carro pode evoluir significativamente ao longo de alguns anos, como resultado das preferências pode mostrar uma evolução correspondente. Em outros casos, pode ser difícil capturar totalmente o interesse do usuário com dados históricos, como classificações. Um item específico pode ter atributos associados a ele que correspondem às suas várias propriedades, e um usuário pode estar interessado apenas em itens com propriedades específicas. Por exemplo, carros podem ter várias marcas, modelos,

Tabela 1.2: Os objetivos conceituais de vários sistemas de recomendação

Abordagem	Objetivo Conceitual	Entrada
Colaborativo	Dê-me recomendações com base em uma abordagem colaborativa que alavanca as avaliações e ações dos meus colegas/minhas.	Avaliações do usuário + classificações da comunidade
Conteúdo-baseado	Dê-me recomendações com base no conteúdo (atributos) Eu tenho favorecido em minhas avaliações e ações passadas.	Avaliações do usuário + atributos do item
Baseado no conhecimento	Dê-me recomendações com base em minhas especificações explícitas Especificação do tipo de conteúdo (atributos) que desejo. atributos do item +	usuário + conhecimento de domínio

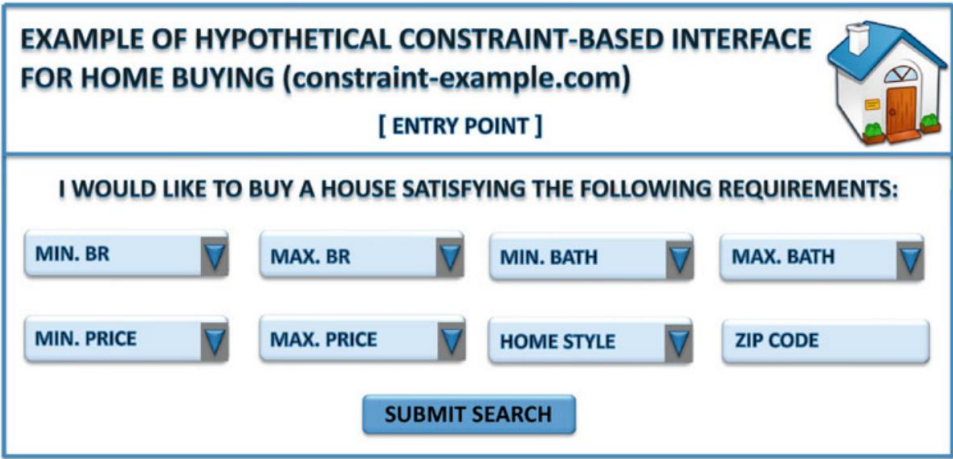


Figura 1.5: Um exemplo hipotético de uma interface de usuário inicial para um recomendador baseado em restrições)

cores, opções de motor e opções de interior, e os interesses do usuário podem ser regulados por um combinação específica dessas opções. Assim, nesses casos, o domínio do item tende a ser complexo em termos de suas propriedades variadas, e é difícil associar classificações suficientes com o grande número de combinações disponíveis.

Tais casos podem ser abordados com sistemas de recomendação baseados em conhecimento, nos quais as classificações não são utilizadas para fins de recomendação. Em vez disso, o processo de recomendação é realizado com base nas semelhanças entre os requisitos do cliente e as descrições dos itens, ou o uso de restrições que especificam os requisitos do usuário. O processo é facilitado com o uso de bases de conhecimento, que contêm dados sobre regras e funções de similaridade para uso durante o processo de recuperação. Na verdade, as bases de conhecimento são tão importantes para o funcionamento eficaz desses métodos que a abordagem toma seu nome desse fato. A especificação explícita de requisitos resulta em maior controle dos usuários sobre o processo de recomendação. Em ambos sistemas colaborativos e baseados em conteúdo, as recomendações são decididas inteiramente por as ações/avaliações anteriores do usuário, as ações/avaliações de seus colegas ou uma combinação das duas. Os sistemas baseados no conhecimento são únicos porque permitem que os usuários especifiquem explicitamente o que eles querem. Essa diferença é ilustrada na Tabela 1.2.

Os sistemas de recomendação baseados em conhecimento podem ser classificados com base no tipo de interface (e conhecimento correspondente) usada para atingir os objetivos acima mencionados:

1. Sistemas de recomendação baseados em restrições: Em sistemas baseados em restrições [196, 197], os usuários normalmente especificam requisitos ou restrições (por exemplo, limites inferiores ou superiores) no item



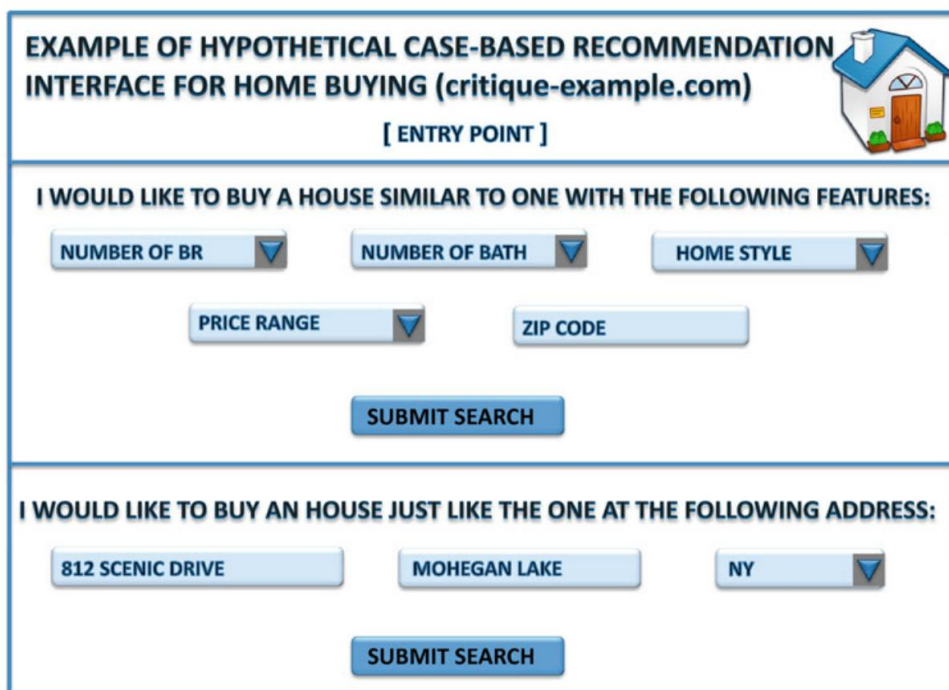


Figura 1.6: Um exemplo hipotético de uma interface de usuário inicial para um recomendador baseado em casos)

atributos. Um exemplo de tal interface é ilustrado na Figura 1.5. Regras específicas de domínio são usadas para corresponder os requisitos do usuário aos atributos do item. Essas regras representam o conhecimento específico de domínio usado pelo sistema. Tais regras podem assumir a forma de restrições específicas de domínio nos atributos do item (por exemplo, "Carros anteriores ao ano de 1970 não têm controle de cruzeiro"). Além disso, sistemas baseados em restrições frequentemente criam regras que relacionam atributos do usuário a atributos do item (por exemplo, "Investidores mais velhos não investem em produtos de risco ultra-alto"). Nesses casos, os atributos do usuário também podem ser especificados no processo de busca. Dependendo do número e do tipo de resultados retornados, o usuário pode ter a oportunidade de modificar seus requisitos originais. Por exemplo, ele pode relaxar algumas de suas restrições quando poucos resultados são retornados ou pode adicionar mais restrições. Esse processo de busca é repetido iterativamente até que o usuário chegue aos resultados desejados.

2. Sistemas de recomendação baseados em casos: Em sistemas de recomendação baseados em casos [102, 116, 377, 558], casos específicos são especificados pelo usuário como alvos ou pontos de ancoragem. Métricas de similaridade são definidas nos atributos do item para recuperar itens semelhantes a esses casos. Um exemplo de tal interface é ilustrado na Figura 1.6. As métricas de similaridade são frequentemente definidas cuidadosamente de forma específica ao domínio. Portanto, as métricas de similaridade formam o conhecimento do domínio que é usado em tais sistemas. Os resultados retornados são frequentemente usados como novos casos-alvo com algumas modificações interativas pelo usuário. Por exemplo, quando um usuário vê um resultado retornado, que é quase semelhante ao que ele deseja, ele pode reemitir uma consulta com esse alvo, mas com alguns dos atributos alterados de acordo com a preferência do usuário. Esse processo iterativo é usado para guiar o usuário em direção aos itens de interesse.

Observe que, em ambos os casos, o sistema oferece ao usuário a oportunidade de alterar seus requisitos especificados. No entanto, a maneira como isso é feito é diferente nos dois casos.

Em sistemas baseados em casos, exemplos (ou casos) são usados como pontos de ancoragem para orientar a busca, em combinação com métricas de similaridade. Interfaces de crítica são particularmente populares para expressar feedback nesses sistemas, onde os usuários modificam iterativamente um ou mais atributos de um item preferido a cada iteração. Em sistemas baseados em restrições, regras (ou restrições) são usadas para orientar a busca. A forma da orientação pode frequentemente assumir a forma de sistemas baseados em busca, onde os usuários especificam suas restrições com uma interface baseada em busca.

Como a interatividade é alcançada em sistemas de recomendação baseados em conhecimento? Essa orientação ocorre por meio de um ou mais dos seguintes métodos:

1. **Sistemas conversacionais:** Neste caso, as preferências do usuário são determinadas iterativamente no contexto de um ciclo de feedback. A principal razão para isso é que o domínio do item é complexo e as preferências do usuário só podem ser determinadas no contexto de um sistema conversacional iterativo.
2. **Sistemas baseados em pesquisa:** em sistemas baseados em pesquisa, as preferências do usuário são obtidas usando uma sequência predefinida de perguntas, como a seguinte: "Você prefere uma casa em uma área suburbana ou dentro da cidade?" Em alguns casos, interfaces de pesquisa específicas podem ser configuradas para fornecer a capacidade de especificar restrições do usuário.
3. **Recomendação baseada em navegação:** na recomendação baseada em navegação, o usuário especifica uma série de solicitações de alteração para o item que está sendo recomendado no momento. Por meio de um conjunto iterativo de solicitações de mudança, é possível chegar a um item desejável. Um exemplo de solicitação de mudança especificada pelo usuário, quando uma casa específica está sendo recomendada, é o seguinte: "Gostaria de uma casa semelhante a cerca de 8 km a oeste da casa atualmente recomendada". Esses sistemas de recomendação também são chamados de sistemas de recomendação críticos [417].

Vale ressaltar que tanto os sistemas baseados em conhecimento quanto os baseados em conteúdo dependem significativamente dos atributos dos itens. Devido ao uso de atributos de conteúdo, os sistemas baseados em conhecimento herdam algumas das mesmas desvantagens dos sistemas baseados em conteúdo. Por exemplo, assim como os sistemas baseados em conteúdo, as recomendações em sistemas baseados em conhecimento podem, às vezes, ser óbvias, pois o uso de avaliações da comunidade (ou seja, de pares) não é aproveitado. De fato, os sistemas baseados em conhecimento são, às vezes, considerados "primos" dos sistemas baseados em conteúdo [558]. A principal diferença é que os sistemas baseados em conteúdo aprendem com o comportamento anterior do usuário, enquanto os sistemas de recomendação baseados em conhecimento recomendam com base na especificação ativa de suas necessidades e interesses pelo usuário. Portanto, na maior parte da literatura sobre recomendações, os recomendadores baseados em conhecimento são considerados uma categoria distinta dos recomendadores baseados em conteúdo. Essas distinções se baseiam tanto nos objetivos de tais sistemas quanto no tipo de dados de entrada utilizados (ver Tabela 1.2). As diferentes formas de sistemas de recomendação baseados em conhecimento são discutidas no Capítulo 5.

#### 1.3.3.1 Sistemas de recomendação baseados em utilitários

Em sistemas de recomendação baseados em utilidade, uma função de utilidade é definida com base nas características do produto para calcular a probabilidade de um usuário gostar do item [239]. O desafio central nos métodos baseados em utilidade é definir uma função de utilidade apropriada para o usuário em questão.

Vale ressaltar que todos os esquemas de recomendação, sejam eles colaborativos, baseados em conteúdo ou em conhecimento, classificam implicitamente os itens recomendados com base em seu valor percebido (ou utilidade) para o usuário-alvo. Em sistemas baseados em utilidade, esse valor de utilidade é

com base em uma função conhecida a priori. Nesse sentido, tais funções podem ser vistas como um tipo de conhecimento externo. Portanto, sistemas baseados em utilidade podem ser vistos como um caso específico de sistemas de recomendação baseados em conhecimento. De fato, será mostrado no Capítulo 5 que funções de utilidade são frequentemente usadas de diversas maneiras para classificar itens em sistemas de recomendação baseados em conhecimento.

### 1.3.4 Sistemas de recomendação demográfica

Em sistemas de recomendação demográfica, as informações demográficas sobre o usuário são aproveitadas para aprender classificadores que podem mapear dados demográficos específicos para classificações ou propensões de compra. Um sistema de recomendação antigo, conhecido como Grundy [508], recomendava livros com base em uma biblioteca de estereótipos montados manualmente. As características do usuário foram coletadas com o uso de um diálogo interativo. O trabalho em [320] observou que os grupos demográficos da pesquisa de marketing podem ser usados para recomendar itens. Outro trabalho [475] faz recomendações de páginas da Web com base nas características demográficas dos usuários que classificaram uma página específica como positiva. Em muitos casos, as informações demográficas podem ser combinadas com contexto adicional para orientar o processo de recomendação. Essa abordagem está relacionada à metodologia de sistemas de recomendação sensíveis ao contexto. Alguns desses métodos são discutidos na seção 8.5.3 do Capítulo 8.

Técnicas mais recentes têm se concentrado no uso de classificadores para fazer recomendações. Um dos sistemas interessantes a esse respeito foi uma técnica que extraía características das páginas iniciais dos usuários para prever a probabilidade de eles gostarem de determinados restaurantes. Classificadores baseados em regras [31, 32] são frequentemente usados para relacionar o perfil demográfico ao comportamento de compra de forma interativa. Embora a abordagem em [31, 32] não tenha sido usada especificamente para recomendar itens específicos, ela pode ser facilmente combinada com um sistema de recomendação. Tais sistemas de recomendação não são muito diferentes do problema de classificação e modelagem de regressão vanilla, no qual as variáveis de características correspondem aos perfis demográficos e as variáveis dependentes correspondem às classificações ou ao comportamento de compra. Embora os sistemas de recomendação demográfica geralmente não forneçam os melhores resultados de forma independente, eles aumentam significativamente o poder de outros sistemas de recomendação como um componente de modelos híbridos ou de conjunto. Técnicas demográficas às vezes são combinadas com sistemas de recomendação baseados em conhecimento para aumentar sua robustez.

### 1.3.5 Sistemas de recomendação híbridos e baseados em conjuntos

Os três sistemas mencionados exploram diferentes fontes de entrada e podem funcionar bem em diferentes cenários. Por exemplo, sistemas de filtragem colaborativa dependem de avaliações da comunidade, métodos baseados em conteúdo dependem de descrições textuais e das próprias avaliações do usuário-alvo, e sistemas baseados em conhecimento dependem de interações com o usuário no contexto de bases de conhecimento. Da mesma forma, sistemas demográficos usam os perfis demográficos dos usuários para fazer recomendações. Vale ressaltar que esses diferentes sistemas usam diferentes tipos de entrada e têm diferentes pontos fortes e fracos. Alguns sistemas de recomendação, como sistemas baseados em conhecimento, são mais eficazes em configurações de inicialização a frio, onde uma quantidade significativa de dados não está disponível. Outros sistemas de recomendação, como métodos colaborativos, são mais eficazes quando muitos dados estão disponíveis.

Em muitos casos em que uma maior variedade de entradas está disponível, existe a flexibilidade de usar diferentes tipos de sistemas de recomendação para a mesma tarefa. Nesses casos, existem muitas oportunidades para hibridização, onde os vários aspectos de diferentes tipos de sistemas são combinados para alcançar o melhor de todos os mundos. Os sistemas de recomendação híbridos estão intimamente relacionados

para o campo da análise de conjuntos, em que o poder de vários tipos de aprendizagem de máquina algoritmos são combinados para criar um modelo mais robusto. Os sistemas de recomendação baseados em ensemble são capazes de combinar não apenas o poder de múltiplas fontes de dados, mas também são capazes para melhorar a eficácia de uma classe específica de sistemas de recomendação (por exemplo, sistemas colaborativos sistemas) combinando vários modelos do mesmo tipo. Este cenário não é muito diferente daquela da análise de conjunto no campo da classificação de dados. O Capítulo 6 estuda vários estratégias de hibridização para sistemas de recomendação.

### 1.3.6 Avaliação de Sistemas de Recomendação

Dado um conjunto de algoritmos de recomendação, qual é o seu desempenho? Como podemos avaliar sua eficácia relativa? Os sistemas de recomendação compartilham diversas semelhanças conceituais. com o problema de classificação e modelagem de regressão. Em classificação e regressão modelagem, a variável de classe ausente precisa ser prevista a partir das variáveis de recurso. sistemas de recomendação, qualquer uma das entradas da matriz pode estar faltando e precisa ser prevista de forma orientada por dados a partir das entradas observadas na matriz restante. Nesse sentido, a O problema de recomendação pode ser visto como uma generalização do problema de classificação. Portanto, muitos dos modelos utilizados para avaliação de classificadores podem ser utilizados para avaliação sistemas de recomendação, embora com algumas modificações. Existem variações significativas em as técnicas de avaliação utilizadas para diferentes aspectos dos sistemas de recomendação, como classificação previsão ou classificação. A primeira está intimamente relacionada à classificação e à modelagem de regressão, enquanto a segunda está intimamente relacionada à avaliação da eficácia da recuperação em mecanismos de busca. e aplicações de recuperação de informação. Os métodos de avaliação para sistemas de recomendação são discutido em detalhes no Capítulo 7.

## 1.4 Desafios específicos de domínio no Recommender Sistemas

---

Em diferentes domínios, como dados temporais, dados baseados em localização e dados sociais, o contexto da recomendação desempenha um papel crítico. Portanto, a noção de sistemas de recomendação contextual foi desenvolvida para abordar as informações adicionais que surgem em esses domínios. Essa noção é usada com diferentes modificações para vários tipos de dados, como dados temporais, dados de localização ou dados sociais.

### 1.4.1 Sistemas de recomendação baseados em contexto

Sistemas de recomendação baseados em contexto ou com reconhecimento de contexto levam em consideração vários tipos de informações contextuais ao fazer recomendações. Essas informações contextuais podem incluem tempo, localização ou dados sociais. Por exemplo, os tipos de roupas recomendados por um varejista pode depender tanto da estação quanto da localização do cliente. Outro exemplo é o caso em que um determinado tipo de festival ou feriado afeta o cliente subjacente atividade.

Observou-se, de modo geral, que o uso de tais informações contextuais pode melhorar significativamente a eficácia do processo de recomendação. Sistemas de recomendação baseados em contexto são incrivelmente poderosas porque as ideias subjacentes são relevantes para uma ampla variedade de contextos específicos de domínio. De fato, um tema recorrente nos capítulos posteriores do livro será a utilização de um modelo multidimensional [7] para recomendações específicas de contexto em diferentes

Configurações específicas de domínio. Sistemas de recomendação sensíveis ao contexto serão discutidos no Capítulo 8 de forma geral. No entanto, aspectos individuais do contexto, como tempo, localização e informações sociais, são estudados em detalhes em outros capítulos. Uma revisão geral desses diferentes aspectos é fornecida abaixo.

### 1.4.2 Sistemas de recomendação sensíveis ao tempo

Em muitos cenários, as recomendações para um item podem evoluir com o tempo. Por exemplo, as recomendações para um filme podem ser muito diferentes no momento do lançamento em relação às recomendações recebidas vários anos depois. Nesses casos, é extremamente importante incorporar o conhecimento temporal no processo de recomendação. O aspecto temporal nesses sistemas de recomendação pode ser refletido de diversas maneiras:

1. A classificação de um item pode evoluir com o tempo, à medida que as atitudes da comunidade evoluem e os interesses dos usuários mudam. Os interesses, gostos, desgostos e modas dos usuários inevitavelmente evoluem com o tempo.
2. A classificação de um item pode depender da hora do dia, dia da semana, mês ou estação específica. Por exemplo, não faz muito sentido recomendar roupas de inverno durante o verão ou capas de chuva durante a estação seca.

O primeiro tipo de sistema de recomendação é criado pela incorporação do tempo como um parâmetro explícito em sistemas de filtragem colaborativa. O segundo tipo pode ser visto como um caso especial de sistemas de recomendação baseados em contexto. Sistemas de recomendação temporais são desafiadores devido ao fato de que a matriz de classificações é esparsa, e o uso de um contexto temporal específico agrava o problema de escassez. Portanto, é particularmente importante ter acesso a grandes conjuntos de dados nesses cenários.

Outra configuração comum é a de conjuntos de dados de feedback implícito, como fluxos de cliques na Web. A atividade do usuário na web e em outras plataformas da internet gera muitos dados úteis que podem ser minerados para fazer recomendações sobre atividades futuras. Nesses casos, a mineração de padrões sequenciais discretos e os modelos de Markov são úteis. O problema da recomendação sensível ao tempo é discutido em detalhes no Capítulo 9.

### 1.4.3 Sistemas de recomendação baseados em localização

Com a crescente popularidade dos celulares com GPS, os consumidores frequentemente se interessam por recomendações baseadas em localização. Por exemplo, um usuário viajante pode querer determinar o restaurante mais próximo com base em seu histórico de avaliações de outros restaurantes. Em geral, a recomendação de lugares sempre inclui um aspecto de localização.

Um exemplo desse sistema é o Foursquare<sup>2</sup>, que recomenda vários tipos de lugares, como restaurantes ou casas noturnas. Há dois tipos de localidade espacial comuns a esses sistemas:

1. Localidade específica do usuário: A localização geográfica de um usuário desempenha um papel importante em suas preferências. Por exemplo, um usuário de Wisconsin pode não ter as mesmas preferências de filmes que um usuário de Nova York. Esse tipo de localidade é chamado de localidade de preferência.

---

<sup>2</sup><http://foursquare.com>

2. Localidade específica do item: A localização geográfica de um item (por exemplo, restaurante) pode ter impacto na relevância do item, dependendo da localização atual do usuário. Os usuários geralmente não estão dispostos a viajar para muito longe de sua localização atual.  
Esse tipo de localidade é chamado de localidade de viagem.

Os algoritmos para localidade de preferência e localidade de viagem são bastante diferentes. Os primeiros estão mais próximos de sistemas sensíveis ao contexto, enquanto os últimos são geralmente projetados como heurísticas ad hoc. Sistemas de recomendação baseados em localização têm despertado interesse crescente nos últimos anos devido à crescente prevalência de celulares e outros dispositivos com GPS. Sistemas de recomendação baseados em localização são discutidos em detalhes no Capítulo 9.

### 1.4.4 Sistemas de Recomendação Social

Os sistemas de recomendação social são baseados em estruturas de rede, sinais e etiquetas sociais, ou uma combinação desses vários aspectos da rede. Em geral, os sistemas de recomendação que são baseados em sinais e etiquetas sociais são ligeiramente diferentes daqueles que são baseados puramente em aspectos estruturais. Os sistemas de recomendação, que são baseados puramente em aspectos estruturais, são usados para sugerir nós e links dentro da própria rede. Por outro lado, os sistemas de recomendação social também podem ser usados para recomendar vários produtos com o uso de sinais sociais. Ambas as formas de sistemas de recomendação serão estudadas neste livro. No entanto, essas formas de recomendação são suficientemente diferentes para serem estudadas em diferentes capítulos deste livro. É importante notar que a utilidade dos sistemas de recomendação estrutural se estende além das redes sociais, porque tais métodos são aplicados a vários tipos de redes habilitadas para a Web.

#### 1.4.4.1 Recomendação Estrutural de Nós e Links

Vários tipos de redes, incluindo redes sociais, são compostos por nós e links. Em muitos casos, é desejável recomendar nós e links. Por exemplo, uma busca personalizada na web pode exigir a recomendação de material relacionado a um tópico específico.

Como a Web pode ser vista como um grafo, tais métodos podem ser vistos como um problema de recomendação de nós. O problema da recomendação de nós está intimamente relacionado ao problema da busca na Web. De fato, ambos os problemas exigem o uso de várias formas de algoritmos de classificação.

Um componente-chave desses métodos é o uso do algoritmo PageRank, embora a personalização desses algoritmos esteja mais relacionada aos algoritmos de recomendação.

Portanto, tais algoritmos também são chamados de algoritmos personalizados de PageRank. Nos casos em que exemplos de nós de interesse estão disponíveis, tais nós podem ser usados como dados de treinamento para determinar outros nós de interesse. Esse problema é chamado de classificação coletiva. Um problema intimamente relacionado é o da recomendação de links ou predição de links, em que é desejável sugerir amigos (ou links em potencial) para um usuário em uma rede social. O problema da predição de links também tem inúmeras aplicações além das redes sociais. Curiosamente, os problemas de classificação, classificação coletiva e recomendação de links estão intimamente relacionados. De fato, soluções para um problema são frequentemente usadas como sub-rotinas para outros problemas. Por exemplo, métodos de classificação e predição de links são frequentemente usados para recomendações tradicionais de produtos em gráficos de itens de usuário. De fato, esses métodos podem ser usados para realizar recomendações em muitas configurações de problemas, que podem ser transformadas em gráficos. Métodos para recomendações de nós e links são discutidos no Capítulo 10.

#### 1.4.4.2 Recomendações de produtos e conteúdo com influência social

Muitas formas de recomendação de produtos e conteúdos são realizadas com a ajuda de conexões de rede e outras dicas sociais. Esse problema também é conhecido como marketing viral. No marketing viral, os produtos são recomendados com o uso de sistemas de boca a boca. Para atingir esse objetivo, é importante ser capaz de determinar entidades influentes e relevantes em termos de tópicos na rede. Esse problema é conhecido como análise de influência em redes sociais [297]. Muitas variações desse problema foram propostas, nas quais os influenciadores são encontrados de forma sensível a tópicos, no cenário de fluxo social. Por exemplo, determinar os usuários influentes em um fluxo do Twitter para tópicos específicos pode ser muito útil para o marketing viral. Em outros casos, dicas sociais são coletadas de redes sociais para fazer recomendações. Esses métodos são discutidos no Capítulo 10.

#### 1.4.4.3 Sistemas de recomendação confiáveis

Muitos sites de mídia social, como Epinions [705] ou Slashdot [706], permitem que os usuários expressem sua confiança e desconfiança uns nos outros, seja de forma direta ou por meio de vários mecanismos de feedback. Por exemplo, os usuários podem expressar sua confiança ou desconfiança em avaliações de outros usuários ou podem especificar diretamente suas relações de confiança ou desconfiança com outros usuários. Essas informações de confiança são muito úteis para fazer recomendações mais robustas. Por exemplo, é evidente que um método de vizinhança baseado no usuário deve ser computado com o uso de pares confiáveis para obter recomendações robustas. Pesquisas recentes mostraram [221, 588, 616] que a incorporação de informações de confiança pode levar a recomendações mais robustas.

Sistemas de recomendação confiáveis são apresentados no Capítulo 11.

#### 1.4.4.4 Aproveitando o feedback de marcação social para recomendações

Os usuários têm vários métodos para incorporar seu feedback em sistemas de recomendação. A forma mais comum de feedback é a marcação social. Essas formas de feedback são particularmente comuns em sites de compartilhamento de conteúdo na Web, como Flickr (compartilhamento de fotos) [692], last.fm [692] (compartilhamento de músicas) e Bibsonomy [708] (compartilhamento de literatura científica). As tags são metadados que os usuários utilizam para adicionar palavras-chave informativas curtas ao conteúdo. Por exemplo, um usuário em um site de música pode marcar o álbum Thriller de Michael Jackson como "rock". Essas tags fornecem informações úteis sobre os interesses do usuário e o conteúdo do item, porque a tag está associada a ambos. As tags servem como contexto útil para a execução das recomendações. Métodos para recomendações sensíveis ao contexto podem ser usados diretamente para incorporar esse feedback no processo de recomendação. Outros métodos especializados também foram desenvolvidos para usar o feedback de marcação social no processo de recomendação.

Esses métodos são discutidos em detalhes no Capítulo 11.

## 1.5 Tópicos e aplicações avançadas

---

Este livro também apresentará uma série de tópicos e aplicações avançadas. A maioria desses tópicos é discutida nos Capítulos 12 e 13, embora alguns deles estejam espalhados pelo livro, quando apropriado. Nesta seção, fornecemos uma breve introdução a esses tópicos.

### 1.5.1 O problema de partida a frio em sistemas de recomendação

Um dos principais problemas em sistemas de recomendação é que o número de classificações inicialmente disponíveis é relativamente pequeno. Nesses casos, torna-se mais difícil aplicar modelos tradicionais de filtragem colaborativa. Embora métodos baseados em conteúdo e conhecimento sejam mais robustos do que modelos colaborativos na presença de partidas a frio, tal conteúdo ou conhecimento pode nem sempre estar disponível. Portanto, diversos métodos específicos foram desenvolvidos para amenizar o problema da partida a frio no contexto de sistemas de recomendação. A suscetibilidade de vários modelos ao problema da partida a frio também é destacada ao longo deste livro, juntamente com possíveis soluções.

### 1.5.2 Sistemas de recomendação resistentes a ataques

O uso de sistemas de recomendação tem um impacto significativo na venda de diversos produtos e serviços. Como resultado, os vendedores de produtos e serviços têm incentivos econômicos significativos para manipular a produção dos sistemas de recomendação. Um exemplo de tal manipulação seria submeter avaliações inflacionadas de seus próprios produtos aos sistemas de recomendação. Um rival malicioso pode submeter avaliações tendenciosas e negativas sobre os produtos de um concorrente. Ao longo dos anos, inúmeras estratégias sofisticadas foram desenvolvidas para atacar sistemas de recomendação. Tais ataques são altamente indesejáveis, pois reduzem a eficácia geral do sistema de recomendação e a qualidade da experiência para usuários legítimos. Portanto, são necessários métodos que permitam recomendações robustas na presença de tais ataques. Os métodos de ataque, incluindo a suscetibilidade de vários tipos de algoritmos a ataques, são discutidos em detalhes no Capítulo 12. Além disso, o Capítulo 12 fornecerá uma série de estratégias para a construção de sistemas de recomendação robustos na presença de tais ataques.

### 1.5.3 Sistemas de recomendação de grupo

Uma extensão interessante dos sistemas de recomendação tradicionais é a noção de sistemas de recomendação de grupo [168]. Nesses casos, o sistema de recomendação é adaptado para recomendar uma atividade específica a um grupo de usuários, em vez de a um único usuário. Exemplos podem incluir assistir a um filme ou televisão por um grupo [408, 653], a seleção de música em uma academia ou as recomendações de viagem a um grupo de turistas. Os primeiros sistemas, como o PolyLens [168], projetaram modelos que agregavam as preferências de usuários individuais para criar recomendações de grupo. No entanto, o consenso ao longo dos anos evoluiu para o desenvolvimento de sistemas de recomendação, que são melhores do que a soma de suas partes e podem levar em consideração as interações entre os vários usuários para o desenvolvimento de recomendações [272, 413]. Estratégias simples de média não funcionam bem quando os grupos são heterogêneos e contêm usuários com gostos diversos [653]. Isso ocorre porque os usuários geralmente têm um impacto nos gostos uns dos outros com base em fenômenos da psicologia social, como contágio emocional e conformidade. Pesquisas detalhadas sobre o assunto podem ser encontradas em [45, 271, 407]. Os sistemas de recomendação de grupo são discutidos na seção 13.4 do Capítulo 13.

### 1.5.4 Sistemas de recomendação multicritério

Em sistemas multicritério, as classificações podem ser especificadas com base em diferentes critérios por um único usuário. Por exemplo, um usuário pode classificar filmes com base no enredo, na música, nos efeitos especiais e assim por diante. Essas técnicas frequentemente fornecem recomendações modelando a utilidade do usuário para um item como um vetor de classificações correspondentes a vários critérios. Em sistemas multicritério



Em sistemas de recomendação, pode-se frequentemente obter resultados enganosos usando apenas a classificação geral em conjunto com um sistema de recomendação tradicional. Por exemplo, se dois usuários têm a mesma classificação geral para um filme, mas suas classificações de componentes para o enredo e a música são muito diferentes, então os dois usuários não devem ser considerados semelhantes da perspectiva de um algoritmo de filtragem colaborativa baseado em similaridade. Em alguns dos sistemas multicritério, os usuários podem não especificar uma classificação geral. Nesses casos, o problema é ainda mais desafiador porque é necessário apresentar listas classificadas de itens para vários usuários com base em múltiplos critérios. Excelentes visões gerais de sistemas de recomendação multicritério podem ser encontradas em [11, 398, 604] de várias perspectivas.

Foi demonstrado [271, 410] que alguns dos métodos para sistemas de recomendação de grupo também podem ser adaptados a sistemas de recomendação multicritério. No entanto, os dois tópicos são geralmente considerados distintos, pois enfatizam aspectos distintos do processo de recomendação. Métodos para sistemas de recomendação multicritério são discutidos na seção 13.5 do Capítulo 13.

### 1.5.5 Aprendizagem Ativa em Sistemas de Recomendação

Um grande desafio em sistemas de recomendação é a aquisição de classificações suficientes para fazer previsões robustas. A escassez da matriz de classificações continua sendo um impedimento significativo para o funcionamento eficaz dos sistemas de recomendação. A aquisição de classificações suficientes pode reduzir o problema da escassez. Diversos sistemas de recomendação do mundo real possuem mecanismos para incentivar os usuários a inserir classificações para popular o sistema. Por exemplo, os usuários podem receber incentivos para avaliar determinados itens. Em geral, é difícil obter muitas classificações de um único usuário devido ao alto custo do processo de aquisição. Portanto, é necessário selecionar criteriosamente os itens a serem avaliados por usuários específicos. Por exemplo, se um usuário já avaliou muitos filmes de ação, pedir que ele avalie outro filme de ação não ajuda muito a prever as classificações de outros filmes de ação e ajuda ainda menos a prever as classificações de filmes pertencentes a gêneros não relacionados. Por outro lado, pedir que o usuário avalie filmes pertencentes a gêneros menos populares ajudará significativamente a prever as classificações de filmes pertencentes a esse gênero. É claro que, se um usuário for solicitado a avaliar um filme não relacionado, não é necessário que ele forneça feedback, pois pode nem ter assistido ao filme. Portanto, existem muitas compensações interessantes no problema de aprendizado ativo de sistemas de recomendação, que não são encontradas em outros domínios de problemas, como classificação. Uma revisão de métodos de aprendizado ativo para sistemas de recomendação pode ser encontrada em [513]. Métodos de aprendizado ativo são discutidos na seção 13.6 do Capítulo 13.

### 1.5.6 Privacidade em Sistemas de Recomendação

Os sistemas de recomendação baseiam-se fortemente no feedback dos utilizadores, que pode ser implícito ou explícito. Este feedback contém informações significativas sobre os interesses do utilizador e pode revelar informações sobre as suas opiniões políticas, orientações sexuais e preferências pessoais. Em muitos casos, essas informações podem ser altamente sensíveis, o que leva a preocupações com a privacidade. Tais preocupações com a privacidade são significativas, uma vez que impedem a divulgação de dados necessários para o avanço dos algoritmos de recomendação. A disponibilidade de dados reais é crucial para os avanços algorítmicos. Por exemplo, a contribuição do conjunto de dados do Prémio Netflix para a comunidade de sistemas de recomendação é inestimável, na medida em que pode ser creditada por motivar o desenvolvimento de muitos algoritmos de última geração [373]. Nos últimos anos, o tema da privacidade tem sido explorado no contexto de uma ampla variedade de

Problemas de mineração de dados [20]. O domínio de recomendação não é exceção, e vários algoritmos de preservação de privacidade foram desenvolvidos [133, 484, 485]. O tópico de privacidade em sistemas de recomendação é discutido em detalhes na seção 13.7 do Capítulo 13.

### 1.5.7 Domínios de Aplicação

Sistemas de recomendação são utilizados em diversos domínios de aplicação, como varejo, música, conteúdo, busca na Web, consultas e anúncios computacionais. Alguns desses domínios exigem métodos especializados para adaptação de sistemas de recomendação. Em particular, o Capítulo 13 estudará três domínios específicos correspondentes a recomendações de notícias, publicidade computacional e sistemas de recomendação recíproca. Todos esses domínios de aplicação são centrados na Web por natureza. Um aspecto importante dos sistemas de recomendação é que eles pressupõem a existência de mecanismos fortes de identificação do usuário para rastrear e identificar interesses de longo prazo dos usuários. Em muitos domínios da Web, mecanismos para identificação forte do usuário podem não estar disponíveis. Nesses casos, o uso direto da tecnologia de recomendação pode não ser viável. Além disso, como novos itens (anúncios) entram e saem continuamente do sistema, certos tipos de métodos, como bandidos multiarmados, são particularmente adequados. Portanto, o Capítulo 13 discutirá os cenários em que a tecnologia de recomendação pode ser utilizada nesses domínios de aplicação. As mudanças específicas que precisam ser feitas nos sistemas de recomendação prontos para uso serão discutidas neste capítulo, juntamente com técnicas avançadas, como bandidos multiarmados.

## 1.6 Resumo

---

Este livro fornecerá uma visão geral das classes mais importantes de algoritmos para sistemas de recomendação, suas vantagens e desvantagens e os cenários específicos em que são mais eficazes. O problema da recomendação será estudado no contexto de diferentes cenários específicos de domínio e com diferentes tipos de informações de entrada e bases de conhecimento. Como este livro mostrará, o problema da recomendação é rico e tem muitas manifestações diferentes, dependendo da natureza dos dados de entrada e do cenário em questão. Além disso, a eficácia relativa de diferentes algoritmos pode variar de acordo com a configuração específica do problema. Essas compensações também serão exploradas por este livro. Em muitos casos, sistemas híbridos podem ser desenvolvidos, explorando essas compensações de forma eficaz.

Vários tópicos avançados, como modelos de ataque, sistemas de recomendação de grupo, sistemas multicritério e sistemas de aprendizagem ativa, serão estudados em capítulos posteriores deste livro. Também exploraremos uma série de aplicações específicas, como recomendações de notícias e publicidade computacional. Esperamos que este livro forneça uma visão geral abrangente e uma compreensão dos diferentes cenários que surgem no campo dos sistemas de recomendação.

## 1.7 Notas Bibliográficas

---

Os sistemas de recomendação tornaram-se cada vez mais populares em meados da década de 1990, com o desenvolvimento de sistemas de recomendação como o GroupLens [501]. Desde então, este tópico tem sido amplamente explorado no contexto de uma ampla variedade de modelos, como sistemas colaborativos, sistemas baseados em conteúdo e sistemas baseados em conhecimento. Pesquisas e livros detalhados sobre o tema podem ser encontrados em [5, 46, 88, 275, 291, 307, 364, 378, 505, 529, 570]. Entre eles, o trabalho em [5] é uma pesquisa muito bem escrita, que fornece uma excelente visão geral das ideias básicas. Pesquisas mais recentes podem ser encontradas em [88, 378, 570]. Uma pesquisa sobre o uso de fontes não tradicionais de

Informações para recomendações, como dados sociais, temporais, secundários ou contextuais, são fornecidas em [544]. Uma classificação recente de várias facetas da pesquisa de sistemas de recomendação pode ser encontrada em [462]. Um excelente livro introdutório pode ser encontrado em [275], enquanto um manual detalhado [505] discute vários aspectos dos sistemas de recomendação em detalhes.

O problema da filtragem colaborativa com matrizes de classificação incompletas está intimamente relacionado à literatura tradicional sobre análise de dados ausentes [362], embora os dois campos tenham sido frequentemente estudados de forma independente. Os primeiros modelos de filtragem colaborativa baseados no usuário foram estudados em [33, 98, 501, 540]. Os métodos baseados no usuário utilizam as classificações de usuários semelhantes no mesmo item para fazer previsões. Embora tais métodos tenham sido inicialmente bastante populares, eles não são facilmente escaláveis e, às vezes, imprecisos. Posteriormente, métodos baseados em itens [181, 360, 524] foram propostos, os quais calculam as classificações previstas como uma função das classificações do mesmo usuário em itens semelhantes. Outra abordagem popular para fazer recomendações é o uso de modelos de fatores latentes. Os primeiros trabalhos em modelos de fatores latentes aparecem independentemente nos contextos de recomendação [525] e análise de valor ausente [24]. Eventualmente, esses métodos foram redescobertos como a classe mais eficaz de métodos para realizar recomendações [252, 309, 313, 500, 517]. Além de seu uso em modelos baseados em fatores, métodos de redução de dimensionalidade também são usados para reduzir a dimensionalidade da matriz de classificações para melhorar a eficiência da similaridade usuário-para-usuário ou item-para-item no espaço reduzido [228, 525]. No entanto, o trabalho sobre análise de dados ausentes é igualmente relevante para a literatura de recomendações. Outros modelos relevantes para filtragem colaborativa incluem o uso de modelos de mineração de dados, como agrupamento [167, 360, 608], classificação ou mineração de padrões de associação [524]. A esparsidade é um grande problema em tais sistemas, e vários sistemas baseados em gráficos foram projetados para aliviar o problema da esparsidade [33, 204, 647].

Os métodos baseados em conteúdo estão intimamente relacionados à literatura de recuperação de informações [144, 364, 400], na qual métodos de recuperação de similaridade são usados no processo de recomendação. Os métodos de classificação de texto também são particularmente úteis no processo de recomendação. Uma discussão detalhada sobre vários métodos de classificação de texto pode ser encontrada em [22]. Alguns dos primeiros trabalhos sobre recomendações baseadas em conteúdo são encontrados em [60, 69]. O levantamento geral em [5] também discute recomendações baseadas em conteúdo de forma bastante extensa.

Existem muitos casos em que métodos colaborativos e baseados em conteúdo não são úteis para obter recomendações significativas devido ao alto grau de complexidade e às restrições no espaço de itens. Nesses casos, sistemas de recomendação baseados em conhecimento [116] são particularmente úteis. Sistemas de recomendação demográficos são discutidos em [320, 475, 508], enquanto sistemas de recomendação baseados em utilidade são discutidos em [239]. Um excelente levantamento sobre explicações em sistemas de recomendação é fornecido em [598].

Diferentes sistemas de recomendação são mais eficazes em diferentes tipos de cenários. A avaliação [246] de sistemas de recomendação é importante para julgar a eficácia de diferentes algoritmos. Uma discussão detalhada dos métodos de avaliação também pode ser encontrada em [538]. Sistemas híbridos [117] podem combinar vários sistemas de recomendação para obter resultados mais eficazes. Além disso, métodos de conjunto também podem combinar algoritmos do mesmo tipo para obter resultados mais eficazes. Os principais trabalhos do concurso Netflix Prize, como "The Ensemble" [704] e "Bellkor's Pragmatic Chaos" [311], foram ambos métodos de conjunto.

Os sistemas de recomendação requerem métodos especializados para torná-los mais eficazes em uma ampla variedade de cenários. Um grande problema no uso eficaz de tais sistemas é o problema da inicialização a frio, no qual um número suficiente de classificações não está disponível no início do processo de recomendação. Portanto, métodos especializados são frequentemente utilizados para lidar com esse problema [533]. Em muitos casos, o contexto da recomendação, como localização, horário ou informações sociais, pode melhorar significativamente o processo de recomendação [7]. Cada um desses diferentes tipos de contexto também foi estudado individualmente como uma área separada de

sistemas de recomendação. Sistemas de recomendação com consciência temporal foram estudados em [310], enquanto os sistemas de recomendação com reconhecimento de localização foram discutidos em [26]. O contexto social é particularmente diverso porque permite uma ampla variedade de configurações de problemas. Pode-se ou recomendar nós ou links em redes sociais, ou pode-se recomendar produtos com a ajuda de pistas sociais. A primeira dessas configurações está intimamente relacionada ao domínio da comunicação social análise de rede [656]. Cada um dos problemas tradicionais de classificação, classificação de nós e A previsão de links [22, 656] pode ser vista como um problema de recomendação estrutural em redes sociais redes. Além disso, essas formas de recomendação são úteis além do contexto das redes sociais. Curiosamente, métodos como a predição de links também podem ser usados para fins tradicionais. recomendação, transformando as interações usuário-item em uma estrutura de grafo bipartido [261]. Uma forma diferente de recomendação social é o caso em que as dicas sociais são usadas para a realização de recomendações [588]. A estrutura da rede social também pode ser diretamente usado no contexto de aplicações de marketing viral [297].

Como os sistemas de recomendação geralmente auxiliam na venda de produtos, os vendedores desses produtos ou seus concorrentes têm motivações significativas para atacar sistemas de recomendação manipulando as classificações. Nesses casos, é improvável que as recomendações sejam de alta qualidade, e, portanto, não confiáveis. Nos últimos anos, uma quantidade significativa de esforço foi dedicada para o projeto de sistemas de recomendação confiáveis [444]. Vários sistemas de recomendação de grupo são discutidos em [45, 271, 272, 407, 408, 412, 413, 415, 653]. Recomendação multicritério sistemas são discutidos em [11, 398, 604]. Métodos de aprendizagem ativa são discutidos em [513]. Uma discussão geral sobre métodos de preservação de privacidade pode ser encontrada em [20]. Os primeiros estudos sobre o tema das recomendações de preservação da privacidade foi apresentado em [133, 451, 484, 485, 667]. A privacidade continua a ser um desafio significativo para tais sistemas devido à natureza altamente dimensional dos dados. Foi demonstrado em [30, 451] como a dimensionalidade pode ser aproveitado para fazer ataques de privacidade em diferentes tipos de conjuntos de dados.

## 1.8 Exercícios

---

1. Explique por que as classificações unárias são significativamente diferentes de outros tipos de classificações em o projeto de sistemas de recomendação.
2. Discuta os casos em que as recomendações baseadas em conteúdo não terão um desempenho tão bom quanto filtragem colaborativa baseada em classificações.
3. Suponha que você configure um sistema, onde uma interface visual guiada é usada para determinar o produto de interesse de um cliente. Qual categoria de recomendador sistema em que esse caso se enquadra?
4. Discuta um cenário em que a localização desempenha um papel importante na recomendação processo.
5. O capítulo menciona o fato de que a filtragem colaborativa pode ser vista como uma generalização do problema de classificação. Discuta um método simples para generalizar algoritmos de classificação para filtragem colaborativa. Explique por que é difícil usar tal método. métodos no contexto de matrizes de classificação esparsas.
6. Suponha que você tivesse um sistema de recomendação que pudesse prever classificações brutas. Como você o usaria para projetar um sistema de recomendação top-k? Discuta o computacional complexidade de tal sistema em termos do número de aplicações da base algoritmo de previsão. Em que circunstâncias tal abordagem se tornaria impraticável?