

Métodos de Geração de Recomendações Usando Filtragem Colaborativa

Aluno: Eduardo Delgado Coloma Bier

Universidade de São Paulo - Instituto de Matemática e Estatística

Informações para contato:

Instituto de Matemática e Estatística

Universidade de São Paulo

Rua do Matão, 1010, São Paulo - SP, Brasil

Email: eduardo.bier@usp.br

https://linux.ime.usp.br/~dududcbier/mac0499



IME-USP

Introdução

Com o grande crescimento da internet e de suas plataformas digitais nas últimas décadas, a tarefa de quais produtos mostrar para um usuário tornou-se algo extremamente não trivial. Sistemas de Recomendação foram criados com o principal objetivo de fazer com que usuários encontrem com maior facilidade aquilo que lhes possa agradar em um site. Isso é especialmente importante para grandes plataformas como a Amazon, que dispõe de um catálogo de milhões de produtos.

Uma das classes de algoritmos mais difundidas no mercado para a geração de recomendações é a filtragem colaborativa. Este trabalho propõe-se a estudar algumas das técnicas de filtragem colaborativa e comparar suas performances. Para isso, foram utilizados dados públicos da plataforma MovieLens, disponibilizados pelo GroupLens Research, para gerar recomendações de filmes para um grupo de usuários da base.

Filtragem Colaborativa

Algoritmos baseados em filtragem colaborativa baseiam-se na ideia de que se dois usuários se comportaram de forma semelhante no passado, então eles devem se comportar de forma semelhante no futuro. Assim, ao contrário do que acontece em outros métodos de geração de recomendações, estes algoritmos exploram somente dados transacionais (compra, avaliações, tags, etc) para gerá-las, usando apenas uma matriz relacionando itens a usuários como entrada. Dessa forma, nem dados demográficos (idade, gênero, etc), nem características dos próprios itens (título, ano de lançamento, entre outros) são explorados em métodos puramente colaborativos e usuários e itens são representados apenas por preferências indicadas através de avaliações.

Neste trabalho, três algoritmos de filtragem colaborativa foram implementados. Seja u o usuário que está recebendo recomendações. O primeiro dos algoritmos, baseado em usuários, busca os usuários similares a u e calcula uma estimativa de avaliação para cada um dos filmes avaliados por eles. O segundo algoritmo é análogo ao anterior mas, ao invés de olhar para os usuários, ele olha para os itens e, por isso, é chamado de filtragem colaborativa baseada em itens. Assim, a partir de cada um dos itens avaliados por u , o algoritmo busca itens similares a eles e estima a avaliação de u . Com essa lista de estimativas em mãos, ambos os algoritmos devolvem a lista de k filmes ainda não avaliados por u com maior estimativa.

O terceiro algoritmo, baseado em passeios aleatórios, se diferencia dos outros dois não só por não tentar estimar a nota que u daria para um filme, mas também por ser baseado em grafos. Nele, um grafo é construído onde usuários e filmes formam os nós e as avaliações, as arestas entre eles. Também são inseridas arestas entre dois filmes e entre dois usuários quando eles são considerados semelhantes entre si. O algoritmo então faz uma série de passeios aleatórios de tamanho fixo a

partir do nó que representa u e devolve a lista dos k filmes que atingiu o maior número de vezes. Foi também implementada uma variante desse algoritmo onde os pesos das arestas (valor das similaridades e avaliações) são levados em conta na hora de fazer os passeios, dando prioridade a arestas com maiores notas e similaridades.

Para os três algoritmos estudados foi necessário fazer o cálculo das similaridades entre usuários e itens. Dentre as várias formas de calcular-las, a Similaridade de Cossenos (1) e o Coeficiente de Pearson (2) foram escolhidos para suas performances serem comparadas.

$$\text{sim}(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} \quad (1)$$

$$\rho_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

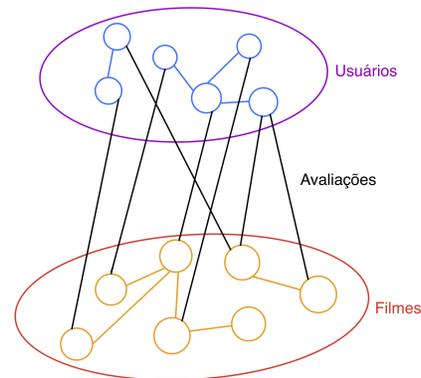


Figure 1: Grafo representando usuários, filmes, avaliações e similaridades

Métricas Utilizadas

• **Mean Average Error (MAE):** Erro médio das estimativas de avaliações feitas pelos algoritmos.

$$MAE = \frac{\sum_{u \in U} \sum_{i \in \text{testset}_u} |\text{rec}(u, i) - r_{ui}|}{\sum_{u \in U} |\text{testset}_u|} \quad (3)$$

• **Precisão (P):** Relação entre o número de itens corretamente recomendados e o número total de itens recomendados

$$P_u = \frac{|\text{hits}_u|}{|\text{recset}_u|} \quad (4)$$

• **Recall (R):** Relação entre o número de itens corretamente recomendados e o número total de itens que poderiam ter sido recomendados

$$R_u = \frac{|\text{hits}_u|}{|\text{testset}_u|} \quad (5)$$

• **F1:** Medida harmônica entre Recall e Precisão

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \quad (6)$$

Resultados

Os resultados dos experimentos foram surpreendentemente ruins. Os algoritmos em geral tiveram uma performance muito ruim em todos os cenários testados, que variaram tanto com relação às medidas de similaridade testadas quanto ao tamanho da população separada para o treinamento e número de recomendações geradas.

| Algoritmo | MAE | P | R | F1 |
|-----------|--------|--------|--------|--------|
| RW BI | 2.8183 | 0.0757 | 0.2347 | 0.0904 |
| BU | 3.2013 | 0.0567 | 0.0848 | 0.0508 |
| PA | N/A | 0.1049 | 0.0012 | 0.0023 |
| PAE | N/A | 0.0393 | 0.0007 | 0.0014 |

Table 1: Melhor performance atingida pelos algoritmos nos testes realizados

O algoritmo baseado em itens conseguiu atingir a melhor performance de forma consistente na grande maioria dos cenários de teste. Já o algoritmo baseado em usuários, conseguiu gerar recomendações melhores do que os outros dois algoritmos. Não foi possível observar um aumento de performance usando o passeio enviesado.

É importante notar que a péssima performance dos algoritmos baseados nos passeios aleatórios pode ser explicada pelo fato de que o grafo gerado pela base de dados utilizada é denso. Assim, como muitos nós estão conectados, o passeio aleatório pode facilmente alcançar a maioria dos filmes.

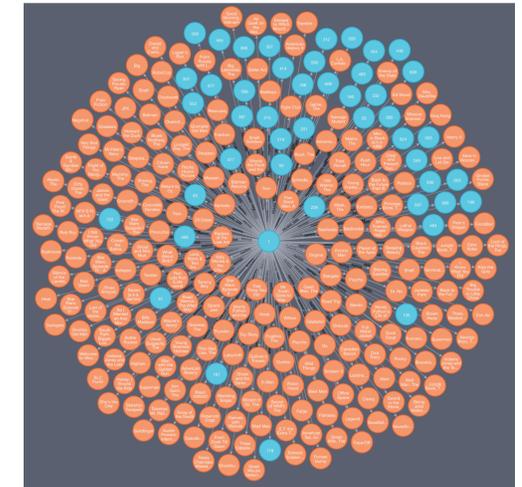


Figure 2: O grafo gerado é extremamente denso. A figura mostra todos os filmes e usuários que estão diretamente conectados com um único usuário.

Conclusões

- Mais testes precisam ser realizados para que se possa de fato avaliar a performance dos algoritmos, sendo necessário rodar os experimentos com uma outra base de dados, preferencialmente uma não manipulada.
- Os testes mostraram que os algoritmos testados tiveram melhor performance com uma população de treinamento maior.

References

- [1] Jonathan L. Herlocker, Joseph A. Konstan, Loren G. Terveen, and John T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22(1):5–53, January 2004.
- [2] Dietmar Jannach, Markus Zanker, Alexander Felfernig, and Gerhard Friedrich. *Recommender Systems: An Introduction*. Cambridge University Press, New York, NY, USA, 1st edition, 2010.
- [3] Francesco Ricci, Lior Rokach, and Bracha Shapira. *Introduction to Recommender Systems Handbook*, pages 1–35. Springer US, Boston, MA, 2011.
- [4] J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen. *Collaborative Filtering Recommender Systems*, pages 291–324. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.