

RELATÓRIO IMPLEMENTAÇÃO KNN

Autores: Paulo Roberto Xavier Júnior, Gabriel Segobia, Fellipe Bravo

E-mails: contato.pauloxavier@gmail.com, segobia.gos@gmail.com, fellipe.bravo@gmail.com

1 Sobre

Entre as abordagens de recomendação colaborativas, métodos baseados nos vizinhos mais próximos ainda possuem uma grande popularidade, devido a sua simplicidade, eficiência e sua capacidade de produzir recomendações personalizadas precisamente.

Por ser tão popular e de implementação *trivial*, o KNN foi implementado e seus resultados analisados para maior entendimento desse algoritmo.

2 Introdução

Em sua forma original, o KNN é compartilhado com virtualmente todos os CF em seu estado inicial é user-user based. Devido as seguintes características abaixo, essa abordagem é extremamente popular:

- **Simplicidade:** São intuitivos e relativamente simples de implementar. Em sua forma mais simples, apenas um parametro (número de vizinhos usados na predição) necessita de ajuste.
- **Justificabilidade:** Fornecem justificativas concisas e intuitivas para as predições computadas. Para alguns sistemas é de extrema importância que se explique as recomendações ao usuário.
- **Eficiência:** Um dos pontos fortes dos métodos de vizinhança é sua eficiência. Diferentemente dos sistemas baseados em modelos, eles não precisam de fases de treinamento custosas, o que precisaria ser feito em intervalos frequentes em aplicações comerciais grandes.
- **Estabilidade:** Outra propriedade útil nesse tipo de abordagem é que ele é pouco afetado pela adição constante de usuários, itens e avaliações.

3 Definição

Sistemas de recomendação baseados em Vizinhos Mais Próximos “automatizam” o princípio de “word-of-mouth”(boca-a-boca), onde um depende da opinião de outros indivíduos semelhantes ou outras fontes confiáveis para estimar o valor de um item de acordo com suas preferências.

Existem 2 tipos de abordagens ao se trabalhar com vizinhos mais próximos, são elas: *User-Based* e *Item-Based*. A user-based usa usuários semelhantes para prever avaliações de um usuário u e a item-based utiliza *rating* de itens parecidos para prever avaliações do usuário sobre um determinado item i .

4 User-Based

Métodos de recomendação baseados no usuário predizem a avaliação r_{ui} de um usuário u para um novo item i usando as avaliações dadas para i pelos usuários mais semelhantes para u , chamado de vizinhos mais próximos (K-nearest neighbours). Supondo que nós temos para cada usuário $v \neq u$ um valor w_{uv} representando as similaridades das preferências entre u e v .

Os k -vizinhos-mais-próximos de u , denotados por $N(u)$, são os k usuários v com a maior similaridade w_{uv} para u . Entretanto, somente os usuários que avaliaram o item i podem ser usados na predição de r_{ui} , e nós temos que ao invés de considerar os k usuários mais similares para u que *avaliaram* i .

5 Item-Based

Enquanto métodos user-based se baseiam na opinião de usuários com pensamentos semelhantes para prever uma avaliação, abordagens item-based olham as avaliações dadas para itens similares.

6 Resultados

O dataset utilizado foi o do movielens 100k, e foi usado como treinamento 20% da base e 80% para ..

6.1 Item-Based

6.2 User-Based

7 Conclusão

O KNN é um algoritmo que apesar da sua simplicidade é altamente competitivo, o que é corroborado pelos resultados obtidos em sua implementação. Foi notável também a diferença entre os resultados usando abordagens diferentes.

No item-based foram encontrados... CONCLUSAO ITEM-BASED

No user-based foram encontrados.. CONCLUSAO USER-BASED