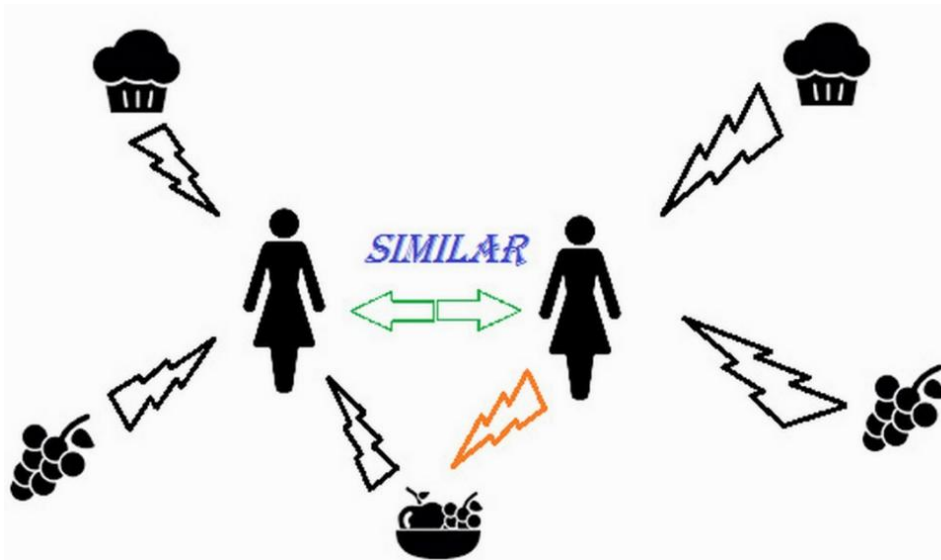


Machine Learning

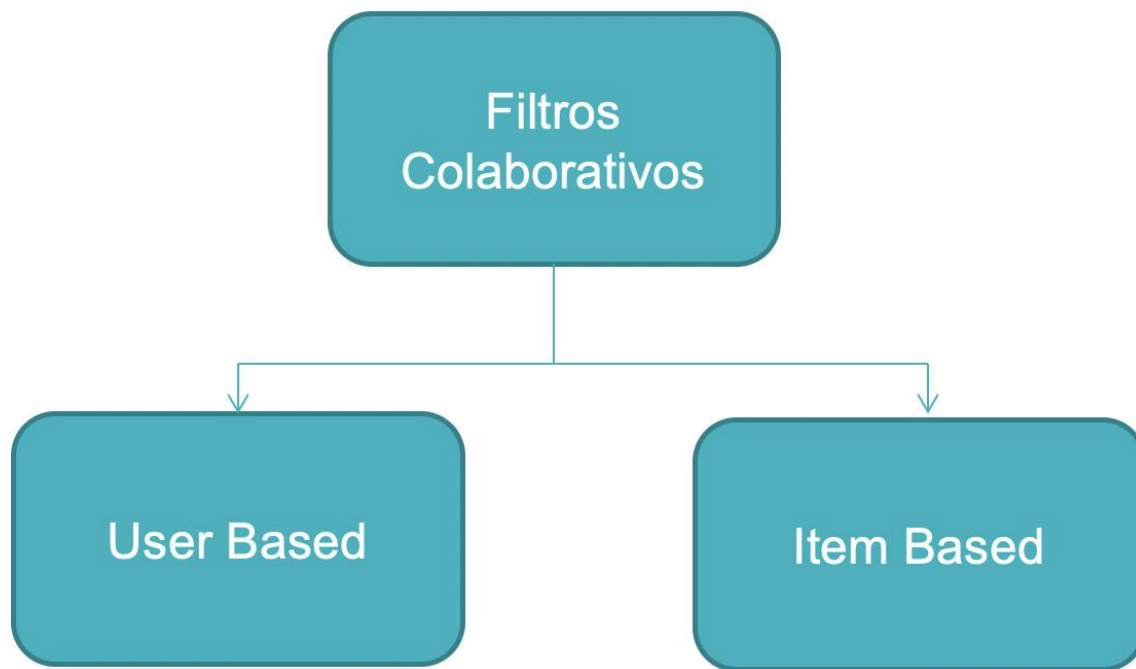
Machine Learning

Filtros Colaborativos

A filtragem colaborativa foca o usuário e outros usuários que são matematicamente similares. Em teoria, não são necessários atributos específicos para o conteúdo que os filtros colaborativos podem inferir. Mais tarde veremos que adicionar atributos de conteúdo pode melhorar o desempenho, mas não é tecnicamente necessário. A premissa é que se dois usuários tiveram uma forte semelhança de gostos e desgostos no passado, é provável que eles vão continuar a ter uma forte semelhança no futuro. Sistemas de recomendação baseados em Filtros Colaborativos irão recomendar às pessoas que gostam de filmes de romance, aqueles filmes que têm forte conteúdo romântico sem a exigência de definir necessariamente o que seja "romance". São apenas números sendo analisados. Uma vez estabelecida a semelhança, os itens consumidos por um usuário podem ser recomendados a outros usuários semelhantes. Essa é a ideia por trás dos filtros colaborativos.



Sistemas de recomendação de filtragem colaborativa são formas básicas de motores de recomendação. Neste tipo de mecanismo de recomendação, filtrar itens de um grande conjunto de alternativas é feito pela colaboração de preferências dos usuários. A suposição básica em um sistema de recomendação de filtragem colaborativa é que, se dois usuários compartilharam os mesmos interesses um do outro no passado, eles também terão gostos semelhantes no futuro. Se, por exemplo, o usuário A e o usuário B tiverem preferências de filme semelhantes, e o usuário A assistiu recentemente ao Titanic, que o usuário B ainda não viu, então a ideia é recomendar esse filme ao usuário B. As recomendações do filme no Netflix são um bom exemplo deste tipo de sistema de recomendação. Existem dois subtipos de sistemas de recomendação de filtragem colaborativa:



Na filtragem colaborativa baseada no usuário (User Based), as recomendações são geradas considerando as preferências na vizinhança do usuário. A filtragem colaborativa baseada em usuário é feita em duas etapas: Identificar usuários semelhantes com base em preferências de outros usuários semelhantes e recomendar novos itens a um usuário ativo com base na classificação dada por usuários semelhantes.

Nos sistemas baseados em itens (Item Based), as recomendações podem ser feitas com base na semelhança de produtos com outros produtos que o cliente comprou ou navegou.

A vantagem dos sistemas de filtragem colaborativa é que eles são simples de implementar e muito precisos. No entanto, eles têm seu próprio conjunto de limitações, como o problema do Cold Start (ou início frio), o que significa que os sistemas de filtragem colaborativa falham em recomendar aos usuários iniciantes cujas informações não estão disponíveis no sistema.

Ao construir sistemas de recomendação de filtragem colaborativa, vamos aprender sobre os seguintes aspectos: Como calcular a similaridade entre os usuários? Como calcular a similaridade entre itens? Como são geradas as recomendações? Como lidar com novos itens e novos usuários cujos dados não são conhecidos? Veremos isso em linguagem R e Python.