

UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

BACHARELADO EM CIÊNCIAS DA COMPUTAÇÃO

Estudo comparativo de Algoritmos de Recomendação

Aluno:

Marcos Masanobu
TAKAHASHI

Orientador:

Prof. Roberto HIRATA
JUNIOR

2 de outubro de 2014

Sumário

1	Resumo	3
2	Introdução	4
2.1	Contextualização	4
2.2	Motivação	4
2.2.1	Amazon	4
2.2.2	Netflix	5
2.2.3	Last.fm	5
2.3	Objetivo	5
3	Conceitos	6
3.1	Algoritmos de Recomendação	6
3.1.1	Content Based Filtering	6
3.1.2	Collaborative Filtering	6
3.1.3	Hybrid Recommender Systems	6
3.2	Mineração de Dados	7
3.2.1	Text Mining	7
3.3	Social Network Analysis	7
4	Metodologia	8
4.1	Avaliação e escolha dos algoritmos a serem utilizados	8
4.1.1	Content based	8
4.1.2	Collaborative Filtering	8
4.1.3	Context based	8
4.2	Tratamento dos dados	8
4.2.1	Definição dos contextos	9
4.2.2	Implementação dos códigos para tratamento	9
4.3	Implementação dos algoritmos de recomendação	9
4.3.1	Large-scale Parallel Collaborative Filtering	9
4.3.2	Fast Context-aware Recommendations with Factori- zation Machines	9

5	Resultados	10
6	Conclusão	11
7	Parte Subjetiva	12
7.1	Desafios e Frustrações	12
7.2	Disciplinas e Conceitos Relevantes	12
7.3	Agradecimentos	12

Capítulo 1

Resumo

Breve resumo do trabalho desenvolvido

Capítulo 2

Introdução

2.1 Contextualização

Com o advento das mídias sociais, tecnologia RFID, computação em nuvem entre outras, é possível ter acesso a dados digitais em um volume, velocidade e variedade inéditos e extrair informações desses dados são um grande diferencial para as empresas. Com base nisso surgiu o Big Data, que é o conjunto de soluções tecnológicas capazes de lidar com essa imensa gama de dados estruturados e não-estruturados.

Atualmente muito se fala sobre Big Data, tanto que é considerado uma das principais tecnologias do momento, mas poucos são os que se aprofundam em algum assunto com exemplos práticos de uso do mesmo no mundo empresarial.

Um dos poucos casos de uso relatados são os Sistemas de Recomendação utilizados na internet e os mesmos são amplamente estudados e possuem uma extensa literatura acadêmica.

2.2 Motivação

Grandes empresas utilizam sistemas de recomendação com diversos algoritmos para melhorar a performance de seu site e outros tipos de mídias, tendo como objetivo exibir somente o produto que o cliente precisa naquele momento.

2.2.1 Amazon

A Amazon é referência em se tratando de serviços de internet e possui um sistema de recomendação de produtos muito poderoso, considerado um dos

mais efetivos. Esse sistema é um dos responsáveis pela alta Conversão do site (transações por visitas) que já chegou a ser de 10% e se baseia em diversas fontes de dados tais como: produtos comprados pelo usuário no passado, quais itens estão no carrinho de compras virtual, itens curtidos ou avaliados, o que outros consumidores viram ou compraram, entre outros. [1] O site da Amazon é completamente personalizado de acordo com o cliente e conforme navega e interage com o site, os produtos exibidos se alteram.

2.2.2 Netflix

Explicação dos trabalhos da netflix, falando do Netflix Prize

2.2.3 Last.fm

Explicação dos trabalhos relacionados a recomendação desenvolvidos pela Last.fm

2.3 Objetivo

Este trabalho tem como objetivo além de analisar e entender os diferentes tipos de algoritmos dos Sistemas de Recomendação, aplicar e testá-los em um contexto do mundo real (e-commerce) para comparar os resultados e retornos financeiros de cada um.

Para tanto, pretende-se utilizar dados de Mídias Sociais como Facebook, dados de buscas e interações dos clientes de forma que aproveite o máximo de informação para melhorar a acurácia da recomendação.

Capítulo 3

Conceitos

3.1 Algoritmos de Recomendação

Diante da cada vez maior exposição a produtos e sites, faz-se necessário ajudar as pessoas a encontrarem aquilo que procuram e nesse contexto é que surgiram os sistemas de recomendação. Estes combinam diversas técnicas computacionais para selecionar os itens que melhor se adequem aos interesses, compras passadas e contexto do usuário. [2]

Os sistemas de recomendação fazem uso de algoritmos de recomendação que são divididos em 3 grupos: Collaborative filtering, Content-based filtering e Hybrid Recommender Systems.

3.1.1 Content Based Filtering

Utiliza características do usuário e do item para ofertar produtos com características semelhantes

3.1.2 Collaborative Filtering

Se baseia em dados passados, tanto de itens comprados pelo usuário ou de notas atribuídas aos mesmos

3.1.3 Hybrid Recommender Systems

São algoritmos que combinam tanto Collaborative quanto Content-based e podem ser feitos de diversas formas diferentes

3.2 Mineração de Dados

3.2.1 Text Mining

Breve explicação sobre text-mining, focando na técnica TF-IDF

3.3 Social Network Analysis

Breve explicação sobre Social Network Analysis

Capítulo 4

Metodologia

4.1 Avaliação e escolha dos algoritmos a serem utilizados

Breve explicação dos algoritmos pesquisados e os motivos que levaram na escolha dos que foram implementados

4.1.1 Content based

4.1.2 Collaborative Filtering

4.1.3 Context based

4.2 Tratamento dos dados

Explicação dos tratamentos realizados para obter os dados utilizados pelos algoritmos

- 4.2.1 Definição dos contextos
- 4.2.2 Implementação dos códigos para tratamento
- 4.3 Implementação dos algoritmos de recomendação
 - 4.3.1 Large-scale Parallel Collaborative Filtering
 - 4.3.2 Fast Context-aware Recommendations with Factorization Machines

Capítulo 5

Resultados

Descrição dos resultados obtidos, explicando a metodologia utilizada na avaliação e também as métricas relevantes que foram levadas em conta

Capítulo 6

Conclusão

A conclusão após análise dos resultados. Qual tipo de algoritmo teve o melhor desempenho, porquê e quais os próximos passos.

Capítulo 7

Parte Subjetiva

7.1 Desafios e Frustrações

Parte destinada para falas dos desafios e frustrações encontrados tanto no desenvolvimento do TCC quanto durante o curso do BCC

7.2 Disciplinas e Conceitos Relevantes

Disciplinas relevantes para a elaboração do trabalho

7.3 Agradecimentos

Referências Bibliográficas

- [1] Amazon's recommendation secret. <http://fortune.com/2012/07/30/amazons-recommendation-secret/>. [Online; acessado em 01-outubro-2014].
- [2] Sistema de recomendação. http://pt.wikipedia.org/wiki/Sistema_de_recomenda%C3%A7%C3%A3o. [Online; acessado em 01-outubro-2014].