

Uma Abordagem Aplicável de Deep Learning para Sistemas Recomendadores baseado em Geolocalização

Gabriel Silva Nascimento¹, Luiz Alberto Ferreira Gomes¹

¹Departamento de Ciência da Computação
Pontifícia Universidade Católica de Minas Gerais (PUC Minas)
37.701-355 – Poços de Caldas – MG – Brasil

{gomes.luiz, gabrielgsn30}@gmail.com

Abstract. *Recommendation systems represent the interest, the preference of certain users in something, be it a clothing, a song, a book, a movie. In this sense, the recommendation systems, seek to suggest, guide users to certain options of their interest, thus making the search easier. For working with large amounts of data, Recommendation Systems are better used when they are built with AI, using Machine Learning and its sub areas, such as Deep Learning. Today we may not realize it, but the Recommendation Systems are already in various activities of our day to day, whether to watch a movie on Netflix, or listen to a song on Spotify, on these and so many other platforms already make recommendations based on your preferences and preferences other similar users.*

Resumo. *Sistemas de Recomendação representam o interesse, a preferência de determinados usuários em algo, seja uma roupa, uma música, um livro, um filme. Nesse sentido os sistemas de recomendação, buscam sugerir, orientar os usuários para determinadas opções de seu interesse, fazendo assim que a busca seja facilitada. Por trabalhar com grandes quantidades de dados os Sistemas de Recomendação são melhores aproveitados quando são construídos com IA, utilizando Machine Learning e suas sub áreas, como por exemplo Deep Learning. Hoje podemos não perceber, mas os Sistemas de Recomendação já estão em várias atividades do nosso dia a dia, seja para assistir um filme na Netflix, ou escutar uma música no Spotify, nessas e tantas outras plataformas já fazem recomendação baseadas nas suas preferencias e preferencias de outros usuários similares.*

Palavras-chave Deep Learning, Sistemas de Recomendação, Filtragem Colaborativa, Geolocalização, Cold Start

1. Introdução e contextualização

Com a crescente disseminação de informações e dados, muito em decorrência do avanço tecnológico que estamos passando, surge a necessidade de sugerir conteúdos para seus usuários, levando isso para um contexto de geolocalização, os motoristas, por exemplo, tem a necessidade de saber onde se encontra e onde se localiza postos de abastecimentos próximos a sua geolocalização.

Neste contexto surge o problema de motoristas terem um sistema recomendador, baseado na sua localização e outras preferências. Na cidade de Poços de Caldas, está inserido o projeto Poços + Inteligente, onde, um dos seus pontos, será a implantação

de postos de abastecimentos, para carros elétricos, e assim surgiu a necessidade da recomendação para motoristas de carros elétricos, que por meio de um aplicativo haverá um Sistema Recomendador baseado em determinadas regras e preferências para se fazer a recomendação.

Levando em conta determinadas regras de projeto, será construído um Sistema Recomendador em *Python*, utilizando o *framework Pytorch*, que é utilizada para aprendizado profundo (*Deep Learning*). Assim com esse sistema, os motoristas de carros elétricos terão uma recomendação de posto de abastecimento para suprir suas necessidades.

2. Sistemas de Recomendação

Sistemas de Recomendação são modelos que buscam a redução de informações que um usuário recebe ao realizar uma busca por intermédio de recomendações personalizadas baseadas no seu perfil (Adomavicius Tuzhilin, 2005). São utilizadas técnicas de modelagem de conhecimento para produzir recomendações personalizadas, mostrando sua utilidade para os usuários ao guiá-los na busca de informações e proporcionar indicações de produtos baseadas no seu perfil (Burke, 2002).

De acordo com (Domavicius e Tuzhilin 2005) os Sistemas de Recomendação são classificados da seguinte maneira:

- Sistemas de recomendação baseado no conteúdo: Os usuários recebem recomendações de itens semelhantes às de outros itens adquiridos no passado, esse sistema se baseia sempre em cima dos próprios dados daquele usuário em específico.
- Sistemas de recomendação baseado na filtragem colaborativa: os usuários recebem recomendações baseados nas preferências de outros usuários que possuem perfis semelhantes.
- Sistemas de recomendação híbridos: Utiliza técnicas de recomendação baseada no conteúdo com técnicas de recomendação baseada na filtragem colaborativa. É uma junção das duas técnicas anteriores.

2.1. Sistemas de Recomendação Baseado no Conteúdo

Na recomendação baseada por conteúdo é analisado os dados entre o usuário e sistema, ou seja seu histórico de interação, tendo como característica obter informações por exemplo de quais produtos foram comprados e suas características, ou levando para a música qual estilo de música mais ouvido, sempre utilizado o histórico o passado daquele usuário em específico.(Felfernig Burke, 2008).

Dessa maneira o sistema vai recomendar sempre levando em consideração seu histórico com a interação e nada mais.

Neste modelo a dificuldade ou o que podemos dizer os problemas, são as classificações que devem ser feitas através das características, ou seja, o que pode ser relevante ou não de acordo com o problema. Por exemplo, para compra de livros, o que deve ser levado como característica para classificação, Número de páginas, por exemplo, pode não ser tão relevante, o autor pode ser uma característica relevante, este problema se encontra em como vamos avaliar corretamente as características para organizar o Sistema Recomendador Baseado no Conteúdo.

2.2. Sistemas de Recomendação Baseado em Filtragem Colaborativa

A Filtragem Colaborativa tem sido amplamente utilizada no mundo de hoje, e segundo Berkonsky et al (2008) esse tipo de recomendação são utilizados para premeditar e gerar as recomendações.

A Recomendação baseada em Filtragem colaborativa de acordo com (HERLOCKER, 1999) pode ser dividido em três passos principais:

- Calcular os pesos – Realizar o cálculo de peso para cada usuário em relação ao usuário alvo utilizando uma métrica de similaridade.
- Buscar os vizinhos – Seleciona-se um subconjunto de usuários com maiores níveis de similaridades (vizinhos) para considerar na predição.
- Realizar a predição – Normalizam-se as avaliações fornecidas pelos usuários para o item analisado na recomendação e calcula-se a predição, ponderando-se as avaliações dos vizinhos com seus respectivos pesos de similaridades.

Para a filtragem colaborativa nos deparamos com alguns problemas. Segundo (FILHO, 2004) e (BALABANOVIC, 1997):

- Novo Item – Para um novo item cadastrado no sistema, não tem avaliações sobre ele, histórico, neste caso não há como fazer uma recomendação.
- Novo Usuário – Para um usuário novo no sistema também não há histórico de avaliações de nenhum item, objeto dentro do sistema, ou seja, sem avaliações torna-se impossível encontrar usuários similares para recomendar algo, nessa questão o problema é chamado de *Cold Start*, traduzindo para o português partida a frio.
- Escalabilidade – Volume enorme de dados, seja por muitos cadastros de itens, ou usuários, ou avaliações, e para realizar os cálculos se tornam inviável.
- Dados Esparsos – Quando há muitas informações no sistema não sendo possível realizar adequadamente a similaridade entre usuários.
- Superespecialização – Quando usuário tem muitos itens semelhantes também causa dificuldade na recomendação.

2.3. Sistemas de Recomendação Híbridos

Essa filtragem é a junção dos dois modelos citados anteriormente. Existem diversas estratégias para combinar os dois modelos, abaixo são citadas 2 técnicas:

- Ponderada – Onde os dois modelos, filtragem colaborativa e filtragem baseada no conteúdo são implementados de maneiras separadas, e ao possuir o resultado de cada uma delas, é utilizado uma técnica de combinação linear.
- Mista – Onde novamente os dois modelos geram recomendações, mas na exibição são mostradas juntas para o usuário.

2.4. Partida a Frio (*Cold Start*)

A partida a frio, problema citado no Modelo de Filtragem colaborativo, quando há um usuário novo no sistema e não similaridade com nenhum usuário ou interação com nenhum item do sistema. Podemos resolver esse problema, por exemplo, dentro do mundo da *NetFlix*, para um usuário recém cadastrado, não há histórico e muito menos similaridade com outros usuários, nesse caso pode ser resolvidos com dados da própria base dados, como por exemplo o Filme mais visto, ou o filme mais bem avaliado, ou seja para a partida a frio de um usuário no sistema a solução está na própria base de dados.

3. Modelo

Para o nosso Sistema de Recomendador, baseado em geolocalização a primeira etapa definida é o pré-processamento das informações, que chega através de um arquivo de formato csv, onde dentro deste arquivo está as informações de todos os veículos, nele o sistema faz três pré-processamentos.

- Pré-processamento de abastecimento – Que irá verificar se o veículo fez o abastecimento, caso fez é adicionado 'True' na coluna 'Abastecido'.
- Pré-processamento de localidade do abastecimento – toda coluna 'abastecido' 'True' através deste pré-processamento irá verificar onde foi feito o abastecimento e adicionará na coluna 'idposto' a 'ID' do posto correspondente, utilizando a fórmula de Haversine e os campos de latitude e longitude para conseguir realizar os cálculos.
- Pré-processamento de Afinidade – É verificado veículo por veículo e calculado baseado na quantidade de vezes que abasteceu em determinado posto.

Após todos os pré-processamentos dos dados o arquivo está pronto para ser feito as recomendações, que no final teremos três tipos de recomendações:

- A primeira é baseada na afinidade do motorista com determinado posto, ou seja, aquele posto que ele mais abasteceu.
- A segunda recomendação foi desenvolvida em Deep Learning e utilizando a filtragem colaborativa para a recomendação, baseado na similaridade de usuários com as avaliações dadas a cada posto.
- A terceira é baseada na fórmula de Haversine onde se calcula a distância entre pontos no globo Terrestre, essa fórmula me traz o posto mais próximo do motorista, essa recomendação também resolve um dos problemas da filtragem colaborativa que é o *Cold Start*. Através dela é possível recomendar algum posto para um usuário novo e também através dela é possível recomendar um posto que ainda não possui recomendação.

A recomendação utilizando a filtragem colaborativa, por ela pertencer a uma rede neural, essa rede precisa de um pré-processamento próprio onde, é feito a matriz de fatoração e no nosso caso as linhas dentro dessa matriz são os usuários e as colunas são os postos, dentro da célula é o valor da avaliação.

Após essa montagem da matriz a rede neural também atribui pesos a essas notas para conseguir encontrar a similaridade, após isso a rede faz o treinamento e teste, então o modelo fica pronto, somente após o modelo estar pronto que é possível recomendar os postos para determinado usuário.

A (Figura 1) representa todas as etapas da Rede Neural:

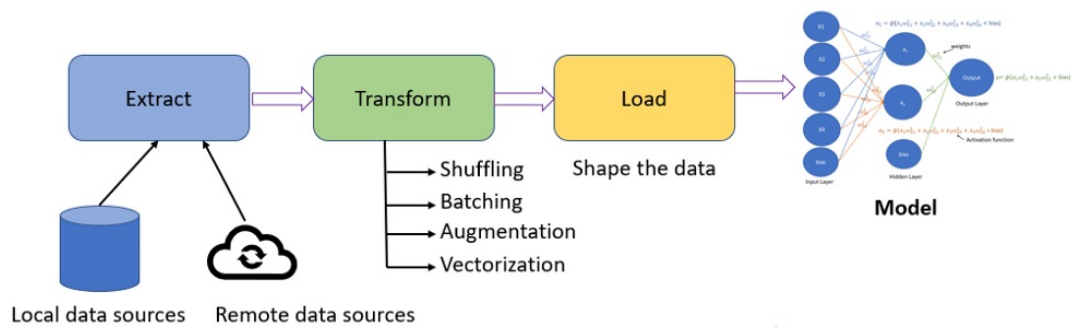


Figura 1. Demonstração das etapas dentro do modelo Deep Learning

A (Figura 2) representa todas as etapas do Sistema Recomendador:

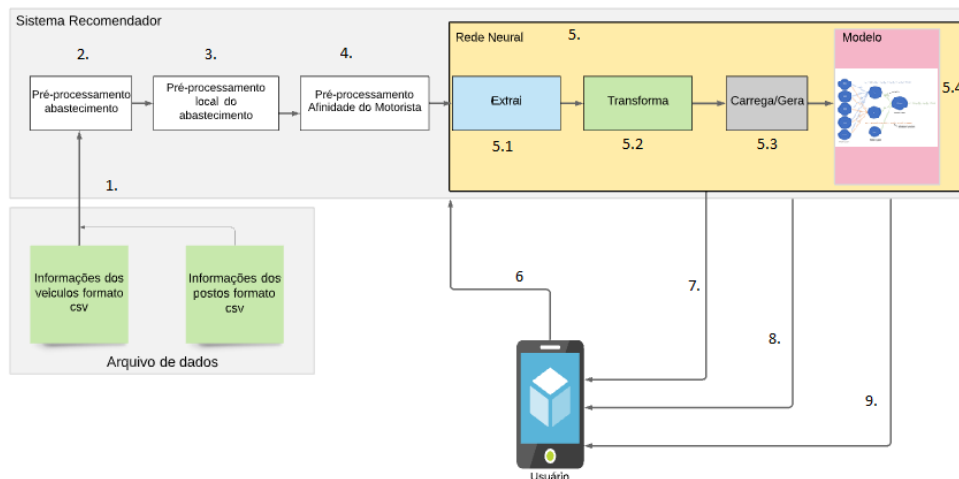


Figura 2. Demonstração de todas as etapas do Sistema Recomendador

Todas as etapas da (Figura 2):

1. Leitura dos arquivos csv
2. Pré-processamento de abastecimento
3. Pré-processamento da localização do abastecimento
4. Pré-processamento da afinidade do motorista
5. Entrando na Rede Neural
 - 5.1. Pré-processamento da Rede Neural
 - 5.2. Transformação dos Dados
 - 5.3. Geração do Modelo
 - 5.4. Modelo pronto (treinado e testado)
6. Requisição do Usuário
7. Recomendação baseada em afinidade
8. Recomendação baseada na Localização Atual
9. Recomendação Baseada na filtragem colaborativa (Rede Neural)

Para usuários novos o sistema só irá fazer a recomendação baseada na localização atual, pois não terá as informações e dados necessários para se fazer as outras recomendações.

4. Resultados

Para obtermos um resultado com maior volume de dados foi feito uma simulação com dados randômicos, onde há 25 postos espalhados na cidade de Poços de Caldas e 40 motoristas que possuem históricos de abastecimentos.(Figura 3):

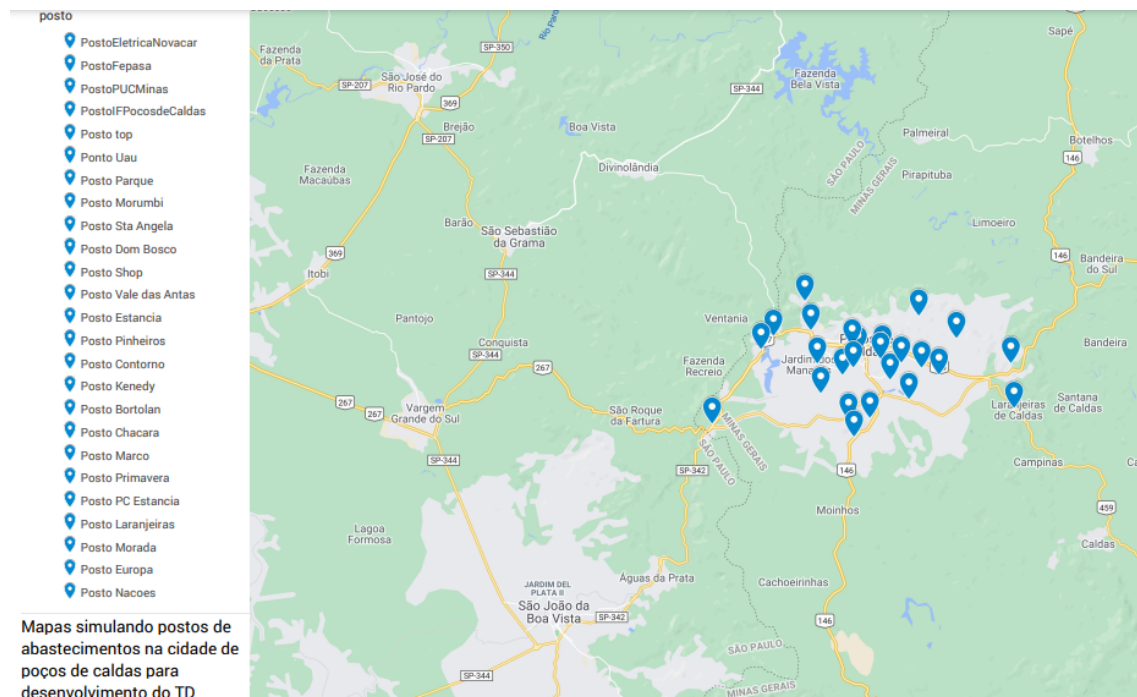


Figura 3. Mapa de Poços de Caldas com todos os postos simulados

Acompanhando a representação do modelo (Figura 2) o resultado será três recomendações ou uma, dependendo da situação do usuário teste imagem.

Para usuários com históricos de dados a saída será da seguinte maneira:(Figura 4):

```
Usuário ID: 3
Postos Recomendados por filtragem colaborativa ID: [3, 2, 19, 18, 15, 20, 14, 11, 17, 1]
Posto mais proximo é: 3
Posto com maior afinidade: 11.0

Usuário ID: 27
Postos Recomendados por filtragem colaborativa ID: [2, 19, 3, 17, 15, 18, 20, 13, 11, 14]
Posto mais proximo é: 3
Posto com maior afinidade: 5.0
```

Figura 4. Saída do Sistema Recomendador para os Usuários de ID 3 e 27

Para usuários sem históricos, a (Figura 5) representa a única recomendação que irá receber, a recomendação baseada na localização atual:

Para postos de abastecimentos que também não possuem históricos de avaliação a maneira que foi encontrada para recomendar estes postos foi também trazendo o posto mais

```
Usuário ID: 41
Usuário sem histórico para recomendacao por filtragem colaborativa e por afinidade, será recomendado o posto mais próximo
Posto mais proximo é: 3
```

Figura 5. Saída do Sistema Recomendador para os Usuários de ID 41 sem histórico de informação recomendação baseada na localização atual

próximo do usuário a mesma estratégia utilizada anteriormente, baseada na localização atual, então quando o motorista acessar o sistema caso esteja perto deste posto ele será recomendado, a a (Figura 6) representa de um posto novo, que não possui avaliações no histórico, o posto de ID 25:

```
Usuário ID: 41
Usuário sem histórico para recomendacao por filtragem colaborativa e por afinidade, será recomendado o posto mais próximo
Posto mais proximo é: 25
```

Figura 6. Saída do Sistema Recomendador para posto novo ID posto 25

Como citado na sessão Modelo deste artigo, a recomendação do posto mais próximo resolve dois problemas básicos que temos na recomendação baseada em filtragem colaborativa, o problema de usuários novos e o problema de postos novos e na sessão resultados ficou nítido que foi a maneira ideal para se resolver tais problemas.

5. Considerações Finais

Neste artigo, juntamente com o modelo e o protótipo conseguiu tratar com sucesso uma abordagem aplicável de *Deep Learning* para Sistemas Recomendadores Baseado em Geolocalização, desde a ideia inicial até um protótipo funcional com resultados. As dificuldades encontradas foram exatamente na construção do modelo em Redes Neurais, como seriam feitos os tratamentos das variáveis e qual era o framework com mais material, fóruns de pesquisa e ajuda para a construção.

Partindo da ideia até o presente momento, os resultados são bem promissores, com toda a construção do sistema. Foram empregadas diversas técnicas, como o cálculo da distância, utilizando a fórmula de Haversine, a técnica de redes neurais, implementando a filtragem colaborativa em um dos resultados do Recomendador.

Para trabalhos futuros, no entanto devem focar no aperfeiçoamento das funções juntamente com mais variáveis no modelo, por exemplo, variação no preço da energia durante o dia, análise de horários de pico nos postos, similaridade de usuários baseados em outras variáveis.

6. Referências

(ADOMAVICIUS, G.; TUZHILIN, A.) Toward the next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering – TKDE*, v. 17, n. 6, pp. 734-749, 2005.

Um Método de Filtragem Híbrida Baseado em Perfis Simbólicos Colaborativos disponível em www.cin.ufpe.br/tg/2013-2/cemb.pdf

(PAZZANI, MICHAEL J.) A Framework for collaborative, content-based and demographic filtering. Artificial Intelligence Review, Vol. 13, Issue 5. Kluwer Academic Publishers, Hingham, MA, USA, (1999), pp. 393-408.

(SANDVIG, JEFF J. et al) A Survey of Collaborative Recommendation and the Robustness of Model-Based Algorithms. IEEE Data Eng. Bull. 31(2): 3-13 (2008).

(HERLOCKER, J., KONSTAN, J., BORCHERS, A., RIEDL, J.) An algorithmic framework for performing collaborative filtering. In Conference on Research and Development in Information Retrieval, 1999.

(HERLOCKER, J., KONSTAN, J. e RIEDL, J.) Explaining Collaborative Filtering Recommendations. In Proceedings of ACM 2000 Conference on Computer Supported Cooperative Work, 2000.

<https://laptrinhx.com/how-to-reduce-training-time-for-a-deep-learning-model-using-tf-data-2669795218/>