

Evento: Debates sobre Inteligência Artificial (d.I.A.)

## UTILIZAÇÃO DE IA EM SISTEMAS DE RECOMENDAÇÃO DE PRODUTOS<sup>1</sup>

### AI APPLIED TO PRODUCT RECOMMENDATION SYSTEMS

**Gabriel Cavalheiro Ullmann<sup>2</sup>, Cristiano Alex Künas<sup>3</sup>, Leandro Perius Heck<sup>4</sup>, Edson Luiz Padoin<sup>5</sup>**

<sup>1</sup> Trabalho de estágio desenvolvido no curso de Ciência da Computação da UNIJUÍ

<sup>2</sup> Aluno do Curso de Graduação em Ciência da Computação da UNIJUÍ, gabriel.cavalheiro@sou.unijui.edu.br

<sup>3</sup> Egresso do Curso de Graduação em Ciência da Computação da UNIJUÍ, cristiano.kunas@sou.unijui.edu.br

<sup>4</sup> Egresso do Curso de Graduação em Ciência da Computação da UNIJUÍ, leandro.h@sou.unijui.edu.br

<sup>5</sup> Professor Doutor do Departamento de Ciências Exatas e Engenharias da UNIJUÍ, Orientador, padoin@unijui.edu.br

## INTRODUÇÃO

No contexto de comércio eletrônico os sistemas de recomendação, por meio de diferentes técnicas, são utilizados com o objetivo de selecionar e indicar itens personalizados que um consumidor possa vir a se interessar, a partir de informações sobre suas compras anteriores e interesses. Para as empresas, o surgimento deste novo conceito representa uma maneira de direcionar novos produtos a potenciais compradores, agregando mais valor às vendas. Por outro lado, para os clientes, este sistema pode ser considerado como um tipo de consultor, com a missão de conquistar a confiança dos clientes e usuários ao longo do tempo com a qualidade das recomendações que são apresentadas (TAKAHASHI; HIRATA, 2015).

A abordagem de filtragem colaborativa, uma das mais frequentemente utilizadas em sistemas de recomendação, consiste na identificação de interesses dos utilizadores através da aprendizagem de relações com outros utilizadores (DELANNAY; VERLEYSSEN, 2008). De maneira similar, Chen et al. (2018, tradução nossa) define que a abordagem "faz recomendações para o usuário atualmente ativo utilizando vários históricos de avaliação de outros usuários sem analisar o conteúdo do recurso informacional". Quando cita "recurso informacional" o autor se refere ao produto, pois essa abordagem permite inferir sugestões sem que seja necessário conhecer características do item.

A importância dos sistemas de recomendação pode ser entendida quando levamos em consideração que atualmente tanto indivíduos quanto empresas desejam agilizar seus processos e fazer transações comerciais de forma simples, mas ao mesmo tempo abrangente o suficiente para que várias opções de itens possam ser exploradas e comparadas. Para alcançar esse objetivo, é fundamental que os sistemas de *e-commerce* analisem e entendam os hábitos de seus consumidores. Além das técnicas estatísticas convencionais, abordagens de Inteligência Artificial (IA) e *Machine Learning* (ML) têm se tornado populares recentemente para realizar esse tipo de inferência (RHEUDE, 2019).

Nesse contexto, o presente trabalho consiste no desenvolvimento por parte da empresa *GH Branding* de um sistema de recomendação de produtos e geração automática de pedidos para a empresa *Abradhi*, que atua no ramo de distribuição de produtos de limpeza e higiene. Este sistema, além de gerar recomendações personalizadas aos clientes, também permitirá que os mesmos efetuem pedidos através da *Web*, trabalho que antes era feito por atendentes que cadastravam

Evento: Debates sobre Inteligência Artificial (d.I.A.)

manualmente em um sistema ERP pedidos que chegavam por telefone, mensagem de texto ou solicitação presencial.

## METODOLOGIA

Para o desenvolvimento do projeto definiu-se as seguintes etapas:

**i) Obtenção do *dataset* de pedidos:** A empresa *Abradhi* compartilhou com a *GH Branding* os dados provenientes do ERP que utiliza para controle dos pedidos. Estes dados foram utilizados como base para desenvolvimento e teste do sistema de recomendação. Os *datasets*, organizados em arquivos .csv, contém dados referentes a clientes e seus segmentos de atuação, produtos, categorias de produto e pedidos.

O *dataset* de pedidos foi fundamental para o desenvolvimento do sistema de recomendação, visto que através dele foi possível entender a relação entre os clientes atendidos e produtos comercializados pela *Abradhi*. Os *datasets* de clientes, produtos e categorias serviram de auxílio à análise, pois enquanto o arquivo de pedidos continha apenas os códigos únicos referentes às entidades no banco de dados, eram nesses conjuntos que se encontravam dados mais detalhados sobre os registros (ex.: nome e preço dos produtos, nome e localização dos clientes, dentre outros).

**ii) Definição da lógica do sistema de recomendação:** A análise da relação direta entre clientes e produtos se mostrou insuficiente para produzir recomendações de qualidade por conta do problema conhecido como partida fria (*cold start*), que consiste na “incapacidade do sistema para recomendar novos produtos e gerar recomendações para novos usuários” (FRESSATO, 2019). Isso ocorre pois clientes novos nunca realizaram pedidos, então nada se pode inferir pelo seu histórico. Portanto, é preciso inferir os produtos que serão mais adequados a ele com base nas preferências de outros usuários similares.

Com intuito de buscar um parâmetro de similaridade entre os clientes de acordo com uma abordagem de filtragem colaborativa, os dados do *dataset* de pedidos e segmentos de clientes foram cruzados, criando um 3º conjunto no qual constasse: i) Segmento de cliente; ii) Lista de produtos mais comprados; iii) Mês analisado. Na Figura 1 são descritas as relações entre os dados utilizados nesta etapa.

Para atingir esse objetivo, filtrou-se o *dataset* de pedidos por segmento de cliente e então realizou-se uma contagem para definir qual produto foi comprado com maior frequência. Os 5 produtos com maior número de ocorrências foram armazenados em uma lista. Posteriormente, o conjunto de segmentos/meses/listas de produtos foi salvo em um arquivo no formato .json.

Evento: Debates sobre Inteligência Artificial (d.I.A.)

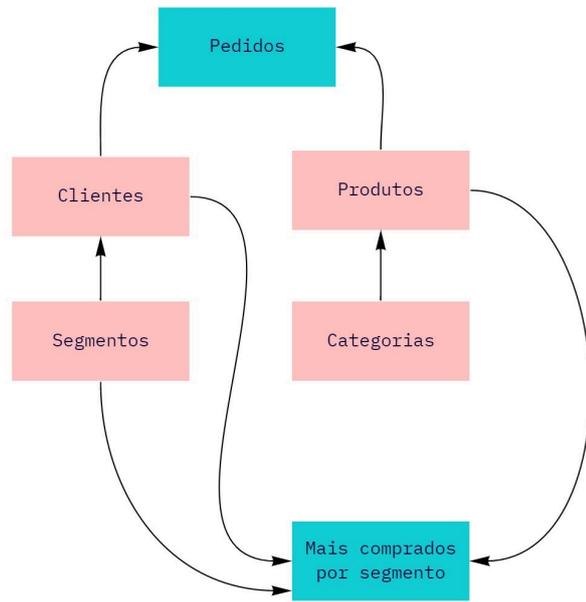


Figura 1 - Conjuntos de dados utilizados no trabalho e suas relações.

**iii) Prototipação e implementação do sistema:** Almejando permitir o consumo deste *dataset* de recomendações por outras aplicações, foi então criada uma API REST na linguagem Python. Para tanto foi utilizando a biblioteca Bottle, que possibilita a criação de um micro *web-framework* WSGI que é rápido, simples e leve (HELLKAMP, 2020).

## RESULTADOS E DISCUSSÃO

Após geradas as recomendações, essas foram analisadas a fim de buscar conclusões sobre o *dataset* estudado e analisar o comportamento do sistema. Foram enumerados o número de ocorrências de cada produto nas listas de mais vendidos de cada segmento e a relação desta contagem com o número total de ocorrências. O número máximo de ocorrências por segmento é 12, já que o algoritmo produz uma lista de mais vendidos por mês. Alguns segmentos não apresentam ocorrências em todos os meses do ano. Esses dados foram organizados em tabelas.

Analisando os produtos mais vendidos por cliente pode-se constatar que os artigos de papel são os mais populares. Considerando o topo da lista conforme a Figura 2, quatro dos cinco itens figuram com percentual acima de 30% independentemente do mês ou segmento analisados, ou seja, apresentam alta volume de vendas de forma recorrente.

Evento: Debates sobre Inteligência Artificial (d.I.A.)

GERAL (não filtrado por categoria)				
ID produto	Nome	Ocorrências	%da categoria	%do total
1805	PAPEL TOALHA A	160	-	45,45
3121	PAPEL TOALHA B	115	-	32,67
2	PAPEL HIGIENICO A	111	-	31,53
2706	PAPEL TOALHA ROLO C	108	-	30,68
1450	PAPEL HIGIENICO B	80	-	22,73

Figura 2 - Os 5 produtos mais vendidos, considerando todos segmentos

Tomando uma categoria específica como base de comparação, pode-se constatar que no segmento “Indústria”, dois dos cinco produtos mais vendidos constam também na lista geral. Os dois produtos no topo da lista (Figura 3) ocorrem em todos os meses analisados (100%). Em contraste com o *ranking* geral, o produto 3552, mais vendido no segmento em questão, foi a escolha mais popular em apenas 2,27% das vezes, considerando todos pares de segmento/mês possíveis.

INDUSTRIA				
ID produto	Nome	Ocorrências	%da categoria	%do total
3552	PAPEL HIGIENICO C	8	100	2,27
2706	PAPEL TOALHA ROLO C	8	100	2,27
1450	PAPEL HIGIENICO B	7	87,5	1,99
3553	PAPEL TOALHA ROLO G	7	87,5	1,99
3231	PAPEL HIGIENICO D	5	62,5	1,42

Figura 3 - Os 5 produtos mais vendidos, considerando o segmento Indústria

## CONSIDERAÇÕES FINAIS

Este trabalho apresentou o desenvolvimento de um sistema de recomendação de produtos para *e-commerce* com utilização de técnicas de IA aliadas com análise estatística. O sistema descrito sugere produtos com base no segmento de atuação de um cliente e um dado mês do ano. Embora a versão atual atenda a um cenário específico, pretende-se utilizá-la como base para a criação de novos modelos de IA. Também visando aprimorar esses sistema, serão utilizadas diferentes abordagens de *Machine Learning* e *Deep Learning*.

**Evento:** Debates sobre Inteligência Artificial (d.I.A.)

**Palavras-chave:** Sistemas de Recomendação. Big Data. Ciência de Dados. Análise de Dados. Análise Estatística. Inteligência Artificial.

## REFERÊNCIAS

CHEN et al. **A Survey of Collaborative Filtering-Based Recommender Systems: From Traditional Methods to Hybrid Methods Based on Social Networks.** 2018. Disponível em: <https://ieeexplore.ieee.org/document/8506344>. Acesso em: 06 out 2020.

DELANNAY, Nicolas; VERLEYSSEN, Michel. **Collaborative filtering with interlaced generalized linear models.** Neurocomputing, v. 71, n. 7-9, p. 1300-1310, 2008.

FRESSATO, E. P. **Incorporação de metadados semânticos para recomendação no cenário de partida fria.** 2019. Disponível em: <https://doi.org/10.11606/d.55.2019.tde-09082019-134753>. Acesso em: 06 out 2020.

HELLKAMP, Marcel. **Bottle: Python Web Framework.** 2020. Disponível em: <https://bottlepy.org/docs/dev/>. Acesso em: 06 out 2020.

RHEUDE, J. Will. **E-Commerce Benefit from Machine Learning or Face a New Threat?** 2019. Disponível em: <https://neilpatel.com/blog/will-e-commerce-benefit-from-machine-learning>. Acesso em: 06 out 2020.

TAKAHASHI, M. M.; HIRATA JR, R. **Estudo comparativo de Algoritmos de Recomendação.** 2015. Disponível em: [https://bcc.ime.usp.br/tccs/2014/marcost/monografia\\_final.pdf](https://bcc.ime.usp.br/tccs/2014/marcost/monografia_final.pdf). Acesso em: 06 out 2020.