

ANÁLISE DA CRIAÇÃO DE PERFIS DE USUÁRIO UTILIZANDO APRENDIZADO DE MÁQUINA

Israel Leandro Cardoso*, José Alexandre Ducatti* e Carlos Magnus Carlson Filho*
*Tecnologia em Informática para Negócios, Faculdade de Tecnologia de São José do Rio Preto

e-mail: ilcnanao@gmail.com; jose.ducatti@fatec.sp.gov.br e carlos.carlson@fatec.sp.gov.br

Resumo: Este artigo buscou, através da análise qualitativa de trabalhos similares, a fundamentação teórica observável e necessária para corroborar as ideias propostas, tentando analisar a possibilidade de uma empresa de vendas de peças e ferramentas, no ambiente *online*, se utilizar de tecnologia computacional de aprendizado de máquina, e com isso, seja capaz de criar perfis dos clientes, com o intuito de investigar preferências pessoais para sugestão acurada e dimensionamento da qualidade dos serviços prestados por eles. Para isso, investigamos possíveis arquiteturas de sistemas disponíveis e algumas das técnicas mais usadas pelos desenvolvedores na atualidade, visando analisar a existência da possibilidade de se obter um projeto que seja capaz de atender as demandas de um ingressante no comércio *online*.

Palavras-chave: Análise. Aprendizado de Máquina. Perfis. Sugestão. Arquiteturas. Técnicas. Comércio Online.

Abstract: *This article sought, through qualitative analysis of similar works, the observable and necessary theoretical foundation to corroborate the proposed ideas, trying to analyze the possibility of a company that sells parts and tools, in the online environment, to use computational technology of machine learning, and with this, to be able to create customer profiles, in order to investigate personal preferences for accurate suggestion and dimensioning of the quality of services provided by them. For this, we investigate possible architectures of available systems and some of the techniques most used by developers today, aiming to analyze the possibility of obtaining a project that is able to meet the demands of a newcomer to online commerce.*

Keywords: *Analysis. Machine Learning. Profiles. Suggestion. Architectures. Techniques. Online Commerce.*

1 Introdução

A tecnologia da informação e a automatização computacional disponíveis, nos dias de hoje, estão, a cada dia, melhorando a capacidade de se alcançar um melhor gerenciamento de processos dentro das empresas. Todas as empresas que tentam se estabelecer no mercado, cada vez mais, buscam maneiras de obter o diferencial que consiga ancorar e alavancar o sucesso para si.

Os recursos computacionais vêm possibilitando uma nova abordagem no que tange o aproveitamento das informações produzidas por uma empresa durante seu funcionamento diário. Esses processos tornam possíveis análises matemáticas automáticas que colaboram para a previsão de problemas com uma margem segura de tempo para se precaver.

Com o constante crescimento das parcerias entre as grandes lojas que já atuam no meio digital, uma imensa quantidade de novos produtos passou a ser ofertado em um mesmo espaço de venda *online* e isso tornou muito mais difícil a busca ou até mesmo o conhecimento da

existência de determinado produto quando não se sabia o nome no momento da pesquisa e isso começou a gerar grandes problemas. O cerne do problema não se encontrava na popularidade do produto anunciado ou na sua qualidade e sim na capacidade do cliente, que não o conhecia, saber de sua existência.

A fim de lidar com o excesso de novos produtos cadastrados a cada dia nos lugares comerciais eletrônicos começaram a ser desenvolvidos sistemas capazes de filtrar os dados com base nas preferências dos utilizadores. Um sistema de recomendação é capaz de filtrar a informação mais importante de acordo com o conjunto de preferencias para um determinado contexto específico (KONSTAN *et al.*, 2012).

Devido à grande importância que estes sistemas de recomendação desempenharam nos mais diversificados nichos de produtos do comércio eletrônico, empresas como Amazon e Netflix começaram a investir pesado no desenvolvimento dessa tecnologia. Como foi na grande competição criada pela Netflix, “The Netflix Prize” em 2006. Essa competição buscava o desenvolvedor de um sistema de recomendação que conseguisse superar, o Cinemath, em 10% nos testes *offline*. Os vencedores, após publicar e explicar os benefícios de sua abordagem, receberiam um prêmio de um milhão de dólares (BENNETT *et al.*, 2007).

Quanto mais as plataformas de comercio eletrônico crescem e ampliam a sua gama de produtos oferecidos, mais se requer a evolução dos sistemas de recomendação, pois é imperativo a sua integração nestas plataformas, com o intuito de se criar uma experiencia personalizada ao cliente (GIANG, 2019; STATISTA, 2022).

A necessidade de se conhecer a possibilidade da tomada de decisão assertiva em um ambiente empresarial de forma otimizada é o que, muitas vezes, garantem a perpetuação do negócio no mercado, pois a facilidade de aperfeiçoamento nas artes de gestão e administração estão acessíveis a todos que a buscam.

A ideia que conduz à sugestão da viabilidade de se usar cálculos estatísticos automáticos dentro de um algoritmo, em que, este seja capaz de sugerir decisões que melhoram os processos dentro da empresa fundamentado nos históricos produzidos pela mesma se mostra como uma razão solida para a investigação proposta por esse trabalho.

Verificar se o Aprendizado de Máquina e toda a ciência de dados que o envolve são apropriados na resolução dos problemas apresentados até o momento, norteiam a razão de ser do presente artigo.

O alvo principal é tentar descobrir um mecanismo computacional que seja capaz de realizar recomendação de produtos aos consumidores de uma loja virtual de venda de peças e ferramentas aos mais variados tipos de prestadores de serviços, mesmo se a loja não possuir histórico de vendas ou se encontrar em pleno projeto de desenvolvimento.

Ao analisar, observar e qualificar diversos fatos garimpados diretamente de trabalhos similares, promoveremos assim, a concatenação das ideias dentro de um mesmo raciocínio, com o intuito de alicerçar a sugestão verossímil que seja congruente para induzir a viabilidade do desenvolvimento do mecanismo computacional proposto. A concepção a partir de técnicas de Aprendizado de Máquina, estará potencialmente habilitada a gerar benefícios para qualquer tipo de negócio independente do nicho escolhido.

Se for possível redarguir a necessidade apresentada na introdução teremos atingido o objetivo principal deste trabalho, pois isso significa que enfaticamente ocorreu a transliteração de fatores comuns do universo computacional para qualquer tipo de empreendimento que busca o crescimento sólido no meio digital, uma vez que o comercio digital no Brasil se encontra, hoje, com uma plataforma concreta que favorece seu crescimento dentro dos próximos anos.

2 Fundamentação Teórica

2.1 Armazenamento de dados

Para se tornar possível analisar, comparar e mensurar a relevância de um conjunto de dados produzidos por uma empresa é necessário que esses dados sejam armazenados em um local onde todas as informações sejam registradas e mantidas em sua integridade podendo ser acessadas instantaneamente e com máxima segurança.

Um Sistema de Gerência de Banco de Dados (SGBD) consiste em um conjunto de dados inter-relacionados e de programas capazes de viabilizar o acesso a esses dados, a fim de recuperar as informações inseridas no banco de dados com acentuadas usabilidade e desempenho (SILBERSCHATZ, 2006).

2.2 Aprendizado de Máquina

O aprendizado de máquina estuda métodos computacionais que um algoritmo computacional se utiliza para desenvolver novas habilidades de organizar o conhecimento já existente e otimizar as inter-relações dentro da base de dados a fim de otimizar cálculos estatísticos extensos e demorados quando realizados manualmente (MITCHELL, 1997).

A capacidade das máquinas de se tornar efetivamente inteligentes, depende diretamente da habilidade de aprender novos conceitos e se adaptar a novas situações, em vez de apenas fazer o que lhes foi mandado (BATISTA, 2003).

2.3 Criação de Perfil de Usuário

A capacidade do ser humano de gerar perfis de usuário em ambiente digital afim de produzir avatares com características individuais com o intuito de promover inter-relações entre os usuários proprietários de perfis teve seu início com as primeiras redes sociais norte-americanas no decorrer da expansão da internet (FRANÇA, 2008). Segundo França (2008), a criação de perfis digitais norteia as interações entre usuários que não se conhecem, pois, as informações dos perfis de uma pessoa real viabilizam a apuração das características, afinidades, gostos e interesses em comum.

2.4 CAGR

O CAGR é um acrônimo da palavra inglesa *Compound Annual Growth Rate* que traduzida para o português do Brasil fica Taxa de Crescimento Anual Composta. Isso é definido como a taxa de retorno anual que é necessária para que um investimento parta de um valor inicial para um valor final esperado em um período estipulado. A fim de verificar se um determinado investimento é viável ou não (REIS, 2019).

2.5 Regras associativas

As regras de associação permitem que elementos em um conjunto de dados sejam expressos como $X \rightarrow Y$, e lê-se X implica em Y , desde que X e Y sejam subconjuntos da base de dados em questão e os conjuntos de itens não tenham interceptação entre si $X \cap Y$ (NOGARE, 2020).

2.6 Clusterização

O método de identificação de grupos de dados semelhantes em um conjunto de dados é chamado de cluster. As entidades de cada grupo são comparativamente mais semelhantes às entidades desse grupo do que as dos outros grupos (ILEOH, 2018).

2.7 Rede Bayesiana

Matematicamente, é uma representação compacta de uma tabela de conjunção de probabilidades do universo do problema. Com as redes bayesianas, máquinas podem fazer previsões como “se o evento X acontecer, tem alta probabilidade de aparecer Y” (STRONSKI, 2018).

2.8 *K-Nearest-Neighbor*

Vizinhos próximos, como é traduzido literalmente, ou KNN é um algoritmo não paramétrico, aonde a estrutura do modelo será determinada pela base de dados utilizada. Este algoritmo também é conhecido como de aprendizado lento. Os algoritmos do desse tipo, não necessitam de dados de treinamento para se gerar o modelo, o que diminui em partes o processo inicial, mas em contrapartida gerará uma necessidade de análise posterior mais apurada (LUZ, 2019).

2.9 *Vector Space Model*

É baseado em noção de similaridade. O modelo assume que a relevância de um documento de consulta é aproximadamente igual à similaridade entre a consulta de documentos. Tanto os documentos quanto as consultas são representados usando o modelo bag-of-words. Para uma coleção de documentos, determinamos primeiro um conjunto de termos, ou seja, vocabulário e em seguida, ordenamos os termos. Logo após, os documentos são representados como vetores n-dimensionais, em que cada dimensão corresponde a um termo. Os termos são ponderados utilizando os métodos tf-idf ou BM25. As consultas são também representadas como vetores n-dimensionais (GUDIVADA, 2018).

3 Trabalhos Similares

3.1 Sistema de Recomendação em tempo real para E-commerce

Segundo Lima (2020), cada vez mais as empresas vêm se utilizando das soluções que o Aprendizado de Máquina provê, quando se trata de recomendação de produtos aos consumidores. Todas as informações produzidas pelos seus próprios históricos de atividades conseguem gerar um sistema para recomendar produtos aos seus consumidores. Esse utilizou de diferentes métodos tendo como base várias bases de dados para o treino de modelos de *Machine Learning* a fim de encontrar uma solução do melhor método de recomendação de produtos para clientes de *e-commerce* de forma automática.

3.2 Aplicação da Clusterização para criação de sistema de recomendação de produtos baseado em perfis de compra

O uso de modelos preditivos de aprendizado de máquina é uma ferramenta indispensável para todas as empresas que desejam destaque ou apenas um diferencial

competitivo. Cria e realiza testes com modelos computacionais a fim de descobrir a eficácia de um sistema de recomendação de produtos com base nos perfis de compra de usuários utilizando-se do algoritmo de clusterização. Por fim, Santana e Pontes (2020) concluíram que é possível se obter alta taxa de assertividade dentro das empresas do ramo de distribuição de produtos a partir da utilização de métodos computacionais de inteligência artificial ou algoritmos preditivos, como o Machine Learning, e salientou que isso cada vez mais fara parte da realidade gerencial das empresas pautadas pelo advento da revolução 4.0.

3.3 Impacto do Aprendizado de Máquina na Detecção de Padrões de Compras de Consumidores para Generalizar Dados Futuros

Segundo Iylin (2019) as empresas não mais se limitam a pesquisas, questionários e outras formas tradicionais de coleta de dados para obter informações sobre o comportamento dos consumidores. Smartphones, cookies, drones sensores e câmeras são a forma moderna de coleta de dados. O Aprendizado de Máquina é o conjunto de técnicas utilizadas para programar computadores e equipá-los para tomar decisões automáticas, essas decisões buscam detectar os padrões em dados passados generalizando tais padrões para a construção de dados futuros (IYLIN, 2019).

4 Metodologia

A realização da Análise Qualitativa em trabalhos similares com o processamento das pesquisas e descobertas que os integram, poderia viabilizar a construção de uma inferência provável entre as diversas soluções da Otimização Computacional. Dentre as fontes de pesquisas encontradas, tudo o que se referir ou se utilizar de dados e após o devido tratamento e processamento dos mesmos, conseguirem gerar o benefício prescritivo ou preditivo, serão capazes de contribuir de forma concreta para satisfazer as demandas buscadas por esse artigo.

O intuito de pesquisa e análise da metodologia empregada nos assuntos investigados não estão aprisionados na ideia da simples análise de caso mediante a comparação de trabalhos similares, e sim, na aspiração dos princípios e ideais discutidos com a provação empírica ou até mesmo sugestiva no que tange a possibilidade da eficácia do método.

A construção do rumo da pesquisa, mediante o uso dos recursos disponíveis, ansiou pela busca de entendimento e a possibilidade de se atingir o sucesso no uso do preceito. Portanto se for possível inculcar que a tecnologia computacional seja de alguma forma capaz de viabilizar sucessos mediante associações estatísticas derivativas dentre as matrizes de dados e com isso promover supostos ideais ou linhas de raciocínios capazes de entender padrões ocorrerá a materialização da ideia sugerida. Com este feito, estaremos satisfeitos com o resultado, pois toda a dedicação na busca frutificou em algo tangível ainda que, por hora, seja somente uma ideia apresentada, e contribuiu para a construção de novos caminhos para outras verificações ou linhas de raciocínios que podem até serem independentes das verificadas aqui.

Existe uma infinidade de técnicas de geração de modelos voltados para Aprendizado de Máquina, bem como, diversas formas de organização de Redes Neurais, com seu devido tratamento da informação, e seria impossível falar de todas elas dado nosso limite permitido na confecção desse trabalho. Mas dentre as técnicas mais utilizadas no sentido que procuramos, citaremos algumas para mero título de comparação e focaremos naquela que avaliarmos como a mais efetiva em sua eficácia independente de sua complexidade.

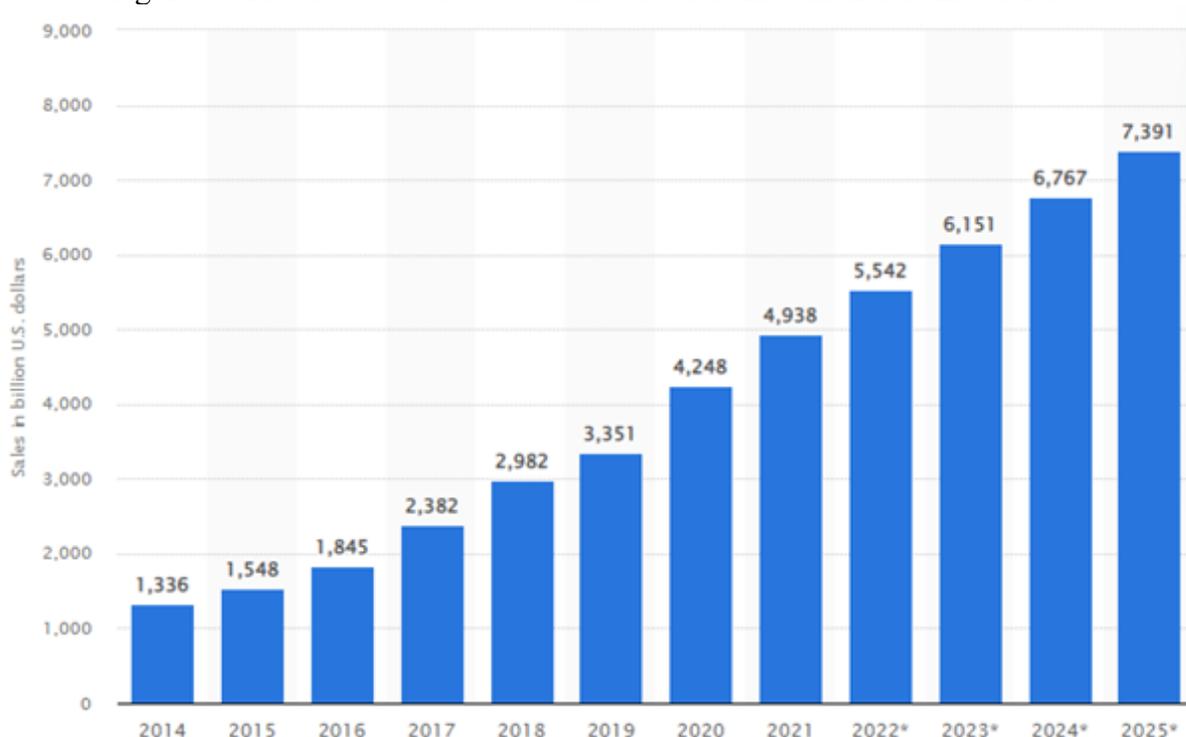
5 Desenvolvimento

Existem inúmeros fatores que estão contribuindo para a digitalização do comércio e isso vem ocorrendo a nível mundial. A recente pandemia do vírus Corona foi um dos impulsores que intensificaram bastante essa transformação. De repente, todos em casa e totalmente desabastecidos e a quarentena, que durou meses, foi a medida necessária mais básica e ao mesmo tempo mais eficaz que se utilizou na tentativa de fazer frente contra a disseminação do vírus, que espantosamente, se propagava muito rápido em locais com considerável aglomeração de pessoas.

É possível verificar, sem erros, que ocorreu o incremento da demanda no meio eletrônico, e isso já vem sendo percebido nos últimos anos. Encontramos o aumento acentuado das plataformas que buscam seu espaço na rede mundial de computadores. Estatísticas recentes demonstram que existe uma tendência sólida de crescimento que perdurará por tempo suficientemente capaz de inviabilizar muito negócios que resistem na não adaptação ou implementação dos novos paradigmas do comércio eletrônico (STATISTA, 2022a).

Em 2021 as atividades do setor de vendas, à nível mundial, atingiram o montante de 4,9 bilhões de dólares, como mostra a Figura 1. Para os próximos anos, os especialistas relatam que a previsão de desenvolvimento da categoria superará o crescimento dos anos anteriores em escala ascendente.

Figura 1 – Previsão de vendas do comércio eletrônico mundial em dólares



Fonte: STATISTA, 2022.

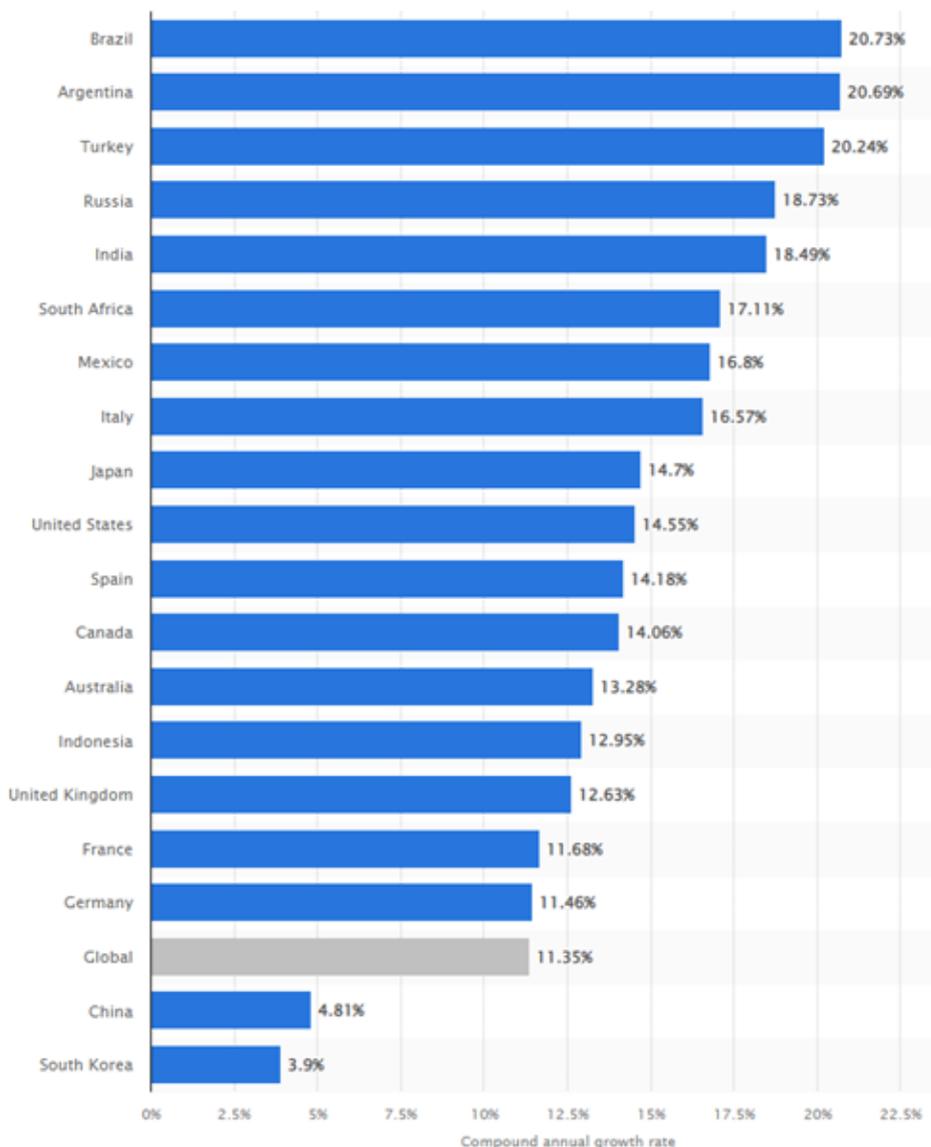
De acordo com cálculos recentes realizados pelos cientistas de dados do STATISTA (2022b), o Brasil ocupará o primeiro lugar em termos de desenvolvimento do comércio eletrônico a varejo com uma taxa de crescimento anual composta de 20,73 por cento entre os anos de 2022 e 2025. As compras online são uma das atividades mais populares na Internet a nível mundial - com algumas categorias de produtos a serem mais populares que outras. A CAGR (*Compound annual growth rate*) global do comércio eletrônico a varejo durante o mesmo período foi de 11,35 por cento.

Mediante a descoberta de tamanha capacidade de expansão do comércio eletrônico dos países, até então, menos favorecidos pela falta de pioneirismo nessa empreitada, encontramos diversos países componentes do grupo dos *BRICS* (*Brazil, Russia, India, China e South Africa*), entre outros com menos projeção no comércio na rede mundial de computadores, com uma seara totalmente capaz de beneficiar tanto o crescimento de sua economia bem como sua influência no cenário mundial.

Apesar da capacidade de desenvolvimento do comércio digital afetar diretamente o país no qual está sendo implementado, podemos contabilizar a certeza de que, após a concretização desse processo, o país em questão poderá colher benefícios capazes de sustentarem o desenvolvimento plausível de suas economias e tudo isso afetará diretamente a economia mundial quando se analisar o todo.

A seguir, podemos contemplar na Figura 2 os argumentos descritos, até o momento, com as projeções de crescimento de cada país dentro do período dos próximos cinco anos. Então podemos ver de forma gráfica como o crescimento do comércio eletrônico possui o cenário favorável aos países, até então, com menos projeção na lista dos países pioneiros na digitalização do comércio eletrônico.

Figura 2 - Previsão de expansão do comércio virtual



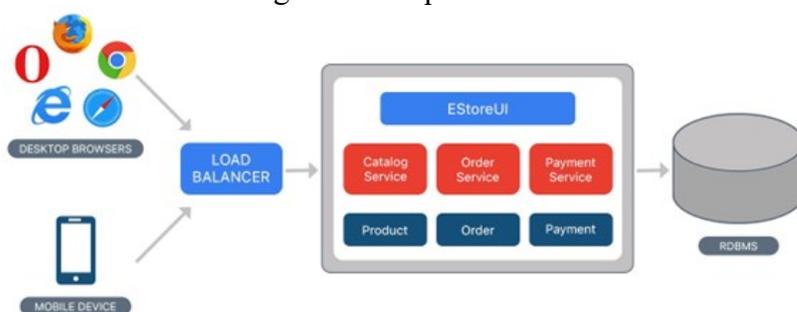
Fonte: STATISTA, 2022.

5.1 Escolha da arquitetura

Quando decidimos desenvolver uma plataforma de comércio eletrônico devemos sempre ter em mente qual o tipo de arquitetura a ser empregada. E essa questão só pode ser efetivamente respondida quando os requisitos do sistema já estão bem definidos e o tipo de software com sua abordagem alvo previamente elaborados.

Para o desenvolvimento da plataforma, muitas empresas, comumente utilizavam a arquitetura monolítica, como vemos na Figura 3, que contavam com a simples lógica de tudo residir apenas em um local. Isso permitia maior controle das regras de negócio, fácil testagem e, de certa forma, a segurança exigia menos requisitos uma vez que tudo se encontrava em um mesmo servidor. Por outro lado, à medida que a aplicação cresce torna-se mais complicada a manutenção, pois um erro em um dos módulos é capaz de comprometer o funcionamento do sistema como um todo (FOWLER et al. 2014).

Figura - 3 Arquitetura Monolítica



Fonte: HAQ, 2018.

Segundo Fowler (2014) outra forma estrutural de organização da arquitetura é denominada de micro serviços ou módulos. Mas a definição dessa arquitetura é muito mais abrangente pois está diretamente relacionada a um conjunto de serviços e cada um deles executados independente e ao mesmo tempo comunicando entre si através de interfaces bem definidas como vemos na Figura 4. Nesse tipo de arquitetura encontramos diversos benefícios como a independência de múltiplos serviços que mantém cada serviço relativamente pequeno e mais fácil de ser entendido pelos programadores bem como a facilidade de atualizações e implementação de cada serviço de forma independente.

De acordo com Haq (2018), assim como todas as abordagens possuem seus prós e contras assim também entre as arquiteturas vai depender muito da complexidade ou da lista de requisitos que se quer atender para fazer a melhor escolha dentre as tecnologias disponíveis. Quando se deseja produzir aplicações mais leves, arquitetura monolítica oferece melhor vantagem por sua simplicidade, para tanto, cada um deve escolher aquela que melhor se enquadra a cada um.

Figura - 4 Arquitetura Micro Serviços

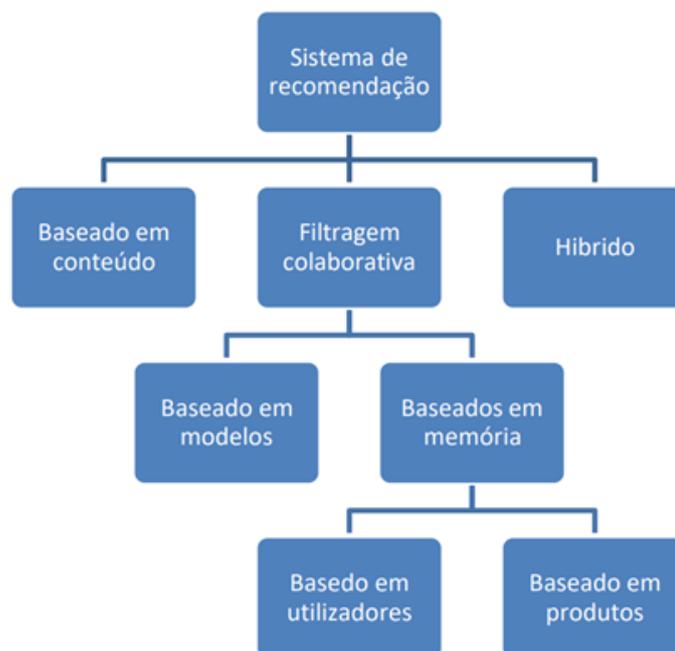


Fonte: HAQ, 2018.

5.2 Escolha da técnica de sistemas de recomendação

Existem muitas técnicas de sistemas de recomendação, mas conforme Isinkaye (2015) na Figura 5 conseguimos identificar três categorias principais que são: Baseado no conteúdo, Filtragem Colaborativa e Híbrido.

Figura - 5 Tipos de Sistema de Recomendação



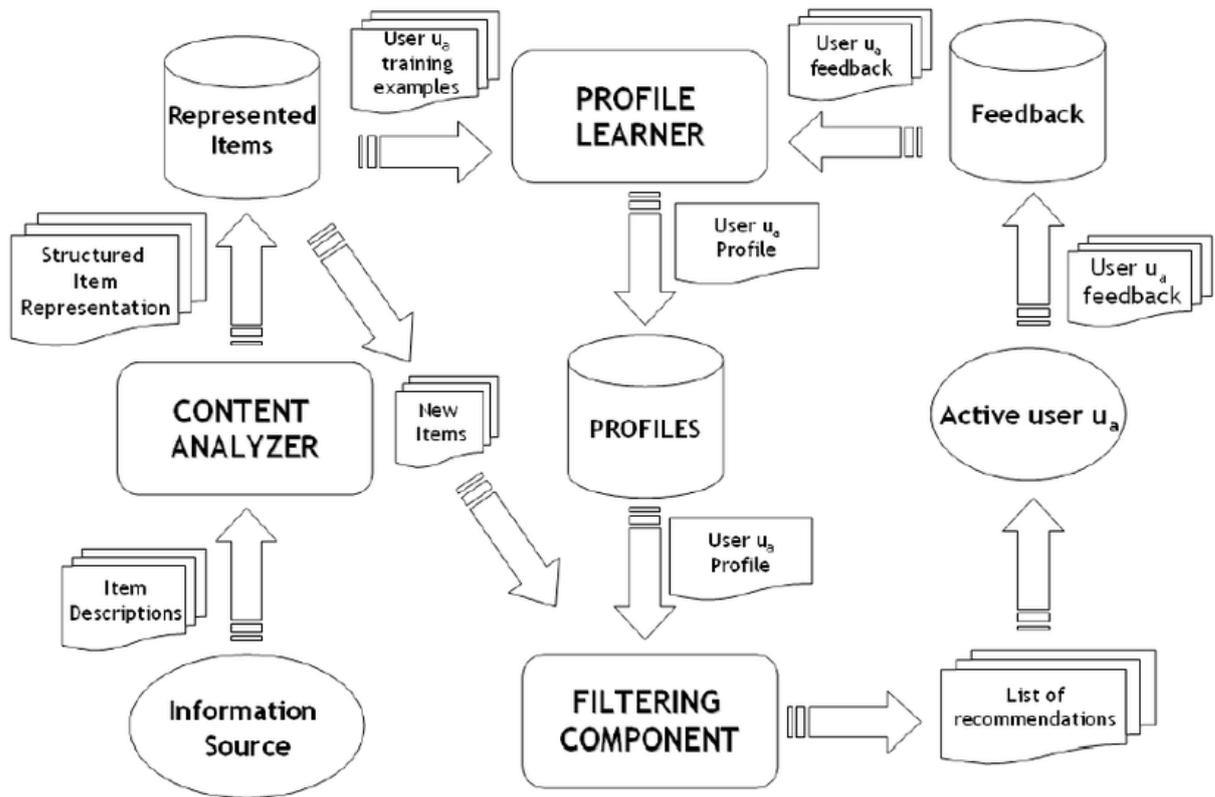
Fonte: ISINKAYE, 2015.

5.2.1 Baseado em Conteúdo

De acordo com Isinkaye (2015) a técnica Baseado em Conteúdo trata-se de um algoritmo dependente do domínio e enfatiza mais a análise dos atributos dos itens a fim de gerar as previsões. Por esse motivo, colhe grande benefício quando utilizada para documentos, notícias ou publicações a serem recomendadas, pois a recomendação é feita com base nos perfis dos utilizadores, mediante o aproveitamento das características extraídas do conteúdo dos itens que o utilizador avaliou no passado. A técnica em si, pode fazer uso de VSM (Vector Space Model), bem como a de modelos probabilísticos como Rede Bayesiana.

A arquitetura de um sistema Baseado em Conteúdo é constituída por três componentes principais como disposto na Figura 6 a seguir:

Figura - 6: Esquema gráfico da técnica Baseado em Conteúdo.



Fonte: KOENE, *et al.* 2015.

5.2.1.1 Profile Learner

Utiliza-se de algoritmos de aprendizagem em um modelo computacional que é capaz de aprender, mediante as características do item juntamente com o feedback do usuário para a conclusão da geração do perfil do comprador. Por esse motivo a técnica requer que as características dos produtos estejam bem definidas.

5.2.1.2 Content Analyser

Fica responsável pela análise e normalização dos dados, para tanto as descrições dos produtos no banco de dados não pode estar sucintamente incompreensível ao entendimento comum, pois sinonímia e polissemia podem causar algumas dificuldades na hora do processamento do texto.

5.2.1.3 Filtering Component

Como o próprio nome sugere é responsável pela filtragem dos produtos que probabilisticamente seriam mais interessantes ao perfil criado. Utiliza-se de algoritmos de correlação como a Similaridade do Cosseno. Nesse momento é possível priorizar produtos semelhantes, apresentando assim, uma sulista de produtos semelhantes (LOPS, *et al.* 2010)

5.2.2 Filtragem Colaborativa

As recomendações feitas por esse tipo de técnica estão diretamente relacionadas aos comportamentos e interesses dos utilizadores do sistema. A técnica em si parte do agrupamento dos utilizadores com base na sua preferência. Portanto se um conjunto de utilizadores partilham da mesma opinião sobre um conjunto de produtos é provável que partilhem da mesma opinião sobre outros produtos (SARWAR *et al.* 2001).

A técnica Filtragem Colaborativa se subdivide em duas categorias, uma é baseada em modelos matemáticos e a outra é baseada em memória, e cada uma possui suas particularidades (ISINKAYE *et al.* 2015).

5.2.2.1 Baseado em modelos

Segundo Isinkaye (2015), esta abordagem permite gerar recomendações rapidamente, com resultados semelhantes à abordagem baseada em memória e resolve os problemas associados com a esparcidade dos dados.

Para a construção do modelo nesse tipo de abordagem podem ser utilizadas as seguintes regras de aprendizado de máquina:

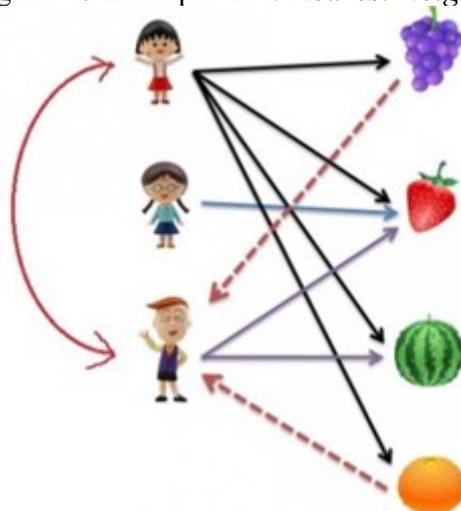
- Regras associativas – o foco principal é identificar padrões associativos dentro das características escolhidas, ou seja, a frequência de ocorrência dentro da lista de objetos (ADITYA *et al.* 2017).
- Clusterização – busca a identificação do conjunto de características entre grupos e a devida qualificação dos interesses dos grupos de utilizadores.
- Rede Bayesiana – realiza os cálculos para a construção das recomendações utilizando o teorema de Bayes. Devemos utilizar sempre em bases de dados constantes ou com pouca frequência de mudança (ISINKAYE *et al.* 2015).

5.2.2.2 Baseado em memória

Os dados disponíveis no sistema são correlacionados entre os clientes e produtos através da similaridade. Atualmente estão sendo amplamente usadas em diversos nichos que se utilizam recomendações ao usuário e se subdividem em duas abordagens.

- *K-Nearest-Neighbor* – vizinhos próximos possuem uma tendência muito forte na semelhança dos interesses como Exemplificado na Figura 7

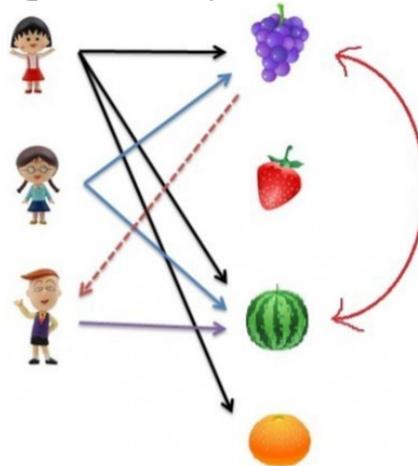
Figura - 7 Exemplo de *k-Nearest Neighbors*



Fonte: PINELA C. (2017)

- Item-to-item – recomendação item a item busca encontrar produtos semelhantes aos escolhidos ou visualizados pelo usuário a fim de recomendar os semelhantes como representado na Figura 8.

Figura - 8 Exemplo de Item-to-item



Fonte: PINELA (2017)

6.2.3 Híbrido

Nessa modalidade de sistema de recomendação ocorrem a combinação de mais de uma técnica a fim de minimizar os efeitos das limitações negativas que uma técnica possui. As vantagens e limitações desta técnica estão diretamente relacionadas com a forma com que as diferentes técnicas são combinadas (BURKE, 2002).

Estão subdivididas em sete categorias:

- Pesos: nessa categoria são combinadas diferentes técnicas de acordo com o peso de cada uma e todas são ajustadas pelo retorno do usuário.
- Alternativa: o sistema alterna entre as técnicas de acordo com o conteúdo.
- Mista: diferentes técnicas são convertidas ao mesmo sistema.
- Combinação de recursos: recursos de diferentes sistemas são combinados ao mesmo tempo.
- Cascata: sintetiza os resultados de outras técnicas construindo a lista de recomendação
- Implementação de recursos: utiliza a saída de uma técnica como entrada de outra técnica.
- *Meta-level*: o modelo gerado por uma técnica torna-se a entrada da segunda técnica.

6 Resultados e Discussões

Existem muitas formas de se criar sistemas de recomendação, como vimos nos trabalhos similares. A tentativa perspicaz, buscando tentar delinear um mecanismo que venha de encontro com a pendência, está atrelada proporcionalmente ao tipo de negócio que se quer estabelecer, juntamente com a disponibilidade de recursos.

Encontramos bastantes subsídios que agregam grande potencial suficiente e capaz de financiar o sucesso do comércio digital de qualquer nicho de negócio, pois o crescimento do comércio digital no Brasil é uma variável materializando-se a cada dia que passa.

A escolha da arquitetura corretamente é muito importante, pois devido aos vários tipos de mecanismos digitais de recomendação a arquitetura vai influenciar diretamente no desempenho do sistema.

Dentre os vários tipos de sistemas de recomendação que existem os que estudamos se mostraram eficazes com as devidas ressalvas detectadas em cada cenário analisado.

O sistema Baseado em Conteúdo é bastante eficaz, porém requer o feedback do usuário para iniciar o processo de recomendação pois o não fornecimento de dados aumenta o risco de esparcidade ou imprecisão.

No sistema Híbrido nos traz diretamente a possibilidade do aproveitamento do melhor de todas as técnicas em um único sistema, porém a harmonia do funcionamento das técnicas em conjunto requer alto nível de conhecimento em programação pela complexidade das técnicas em seu funcionamento cotidiano o que pode encarecer para as pequenas empresas que estão ingressando na digitalização.

Filtragem Colaborativa possui importantes subdivisões exploradas nesse artigo e identificamos que na modalidade Baseado em Modelos, a técnica é capaz de realizar recomendações apenas utilizando-se das regras de aprendizado de máquina para construção do modelo computacional, sanando o problema de esparcidade dos dados. Por outro lado, a técnica Baseado em Memória, se subdivide em Vizinhos Próximos e Item a Item. A primeira correlaciona escolhas de vizinhos e cria a sugestão da recomendação. Item a Item gera a recomendação de itens baseado na similaridade entre os itens escolhidos. Ambas as técnicas demandam nada ou uma quantidade muito pequena de dados para o funcionamento.

Todas as técnicas estudadas, neste artigo, se mostraram perfeitamente capazes de promover a recomendação de produtos independentemente do tipo ou do nicho comercial no qual estiver inserido. Todavia das três técnicas estudadas a que se mostrou mais promissora àqueles que não possuem dados para auxiliar na confecção das recomendações foi a Filtragem Colaborativa tanto na modalidade Vizinhos Próximo, bem como, Item a Item.

Existem muitos detalhes importantes sobre as peculiaridades de cada técnica que poderiam ser abordados por esse artigo, mas acreditamos que “O saber não ficará limitado ao tempo de produção de um único pensador”.

7 Conclusão

Conseguimos compreender que o assunto geração de recomendação de produtos em lojas instaladas no meio digital contam com uma infinidade de mecanismos aptos para sanar qualquer tipo de demanda que possa ser criada. Mesmo quando não se detém uma base de dados farta, existem saídas que viabilizam a criação do sistema como explorados nas técnicas abordadas.

Sempre que existir, na loja virtual de peças, o acúmulo de uma grande variedade de produtos ofertados, de acordo com as técnicas analisadas nos trabalhos similares, existe a possibilidade palpável de solução desse tipo de adversidade.

Portanto o desenvolvedor acredita que se continuarmos nessa linha haverá possibilidade de êxito na implantação desse tipo de atividade.

Referências

ADITYA P. H., BUDI I., MUNAJAT, Q. (2017) *A comparative analysis of memory-based and model-based collaborative filtering on the implementation of recommender system for Ecommerce in Indonesia: A case study PT X*, in 2016 International Conference on Advanced Computer Science and Information Systems, Disponível em: <https://ieeexplore.ieee.org/document/7872755> Acesso em 08 maio 2022.

BATISTA G. E. A. P. A. (2003) Tese de doutorado **Pré-processamento de dados em aprendizado de máquina supervisionado** pág. 39-40 Disponível em:

<https://www.teses.usp.br/teses/disponiveis/55/55134/tde-06102003-160219/en.php>. Acesso em: 20 mar. 2022.

BENNETT J., LANNING S. (2007) *The Netflix Prize, in Proceedings of KDD Cup and Workshop*. doi: 10.1145/1562764.1562769.

BURKE R. (2002) *Hybrid Recommender Systems: Survey and Experiments* Disponível em: <https://link.springer.com/article/10.1023/A:1021240730564>. Acesso em: 08 maio 2022.

CHEREDNICHENKO O., VOVK M. A., KANISHCHEVA O., GODLEVSKYI M. (2008) Tese de Doutorado *Studying items similarity for dependable buying on electronic marketplaces* 2018. Disponível em:

<http://repository.kpi.kharkov.ua/handle/KhPI-Press/46674>. Acesso em: 06 maio 2022.

FRANÇA A.C.P. (2008) Dissertação de Mestrado *Self Digital: explorações acerca da construção do eu na Internet*. Disponível em:

https://repositorio.ufpe.br/bitstream/123456789/8237/1/arquivo3881_1.pdf. Acesso em: 20 mar. 2022.

FOWLER M., LEWIS J. (2014) *Microservice* Disponível em:

<https://martinfowler.com/articles/microservices.html>. acesso em 08 maio 2022.

GIANG (2019) *Online Retail trends: 22 ecommerce predictions for 2019*. Disponível em:

<https://www.connectpos.com/online-retail-trends-2019/#.XEWNf1z7SUK> Acesso em: 08 maio 2022).

GUDIVADA V.N., GUDIVADA A. R. (2018) “*Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications*” Handbook of Statistics Disponível em:

<https://www.sciencedirect.com/topics/computer-science/vector-space-models> acesso em: 08 maio 2022.

HAQ S. U. (2018) *Introduction to Monolithic Architecture and MicroServices Architecture*. Disponível em:

<https://medium.com/koderlabs/introduction-to-monolithic-architecture-and-microservices-architecture-b211a5955c63>. Acesso em 08 maio 2022.

ILEOH (2018) *Introdução A Clusterização e os Diferentes Métodos*. Disponível em

<https://portaldatascience.com/introducao-a-clusterizacao-e-os-diferentes-metodos>. Acesso em 24 maio 2022.

ISINKAYE F.O., FOLAJIMI Y.O., OJOKOH B.A. (2015) *Recommendation systems: Principles, methods and evaluation*. Egyptian Informatics Journal v116 pag. 261-273

Disponível em: <https://www.sciencedirect.com/science/article/pii/S1110866515000341#!> Acesso em 08 maio 2022.

IYALIN D.F.A. (2019) *Impact of Machine Learning on Detecting Buying Patterns of Consumers to Generalize Future Data* THINK INDIA (Quarterly Journal). Disponível em:

<https://thinkindiaquarterly.org/index.php/think-india/article/view/9231>. Acesso em: 05 maio 2022.

KOENE A. R., PEREZ E., CARTER C. J. (2015) *Ethics of Personalized Information Filtering*. ResearchGate. Disponível em: https://www.researchgate.net/figure/High-level-architecture-of-a-Content-based-recommender-system_fig1_300884439.

Acesso em 24 maio 2022.

- KONSTAN, J. A., RIEDL J. (2012) *Recommender systems: From algorithms to user experience*, User Modeling and User-Adapted Interaction. doi: 10.1007/s11257-011-9112-x.
- LIMA R. J. C. (2020) Sistema de **Recomendação em tempo real para E-commerce**. Repositório Científico do Instituto Politécnico do Porto Disponível em <https://recipp.ipp.pt/handle/10400.22/17236> acesso em: 20 de mar. 2022.
- LOPS P., GEMMIS M., SEMERARO G. (2011) *Content-based Recommender Systems: State of the Art and Trends*, in Recommender Systems Handbook. doi: 10.1007/978-0-387-85820-3_3. Disponível em: https://link.springer.com/chapter/10.1007/978-0-387-85820-3_3 Acesso em 08 maio 2022.
- LUZ F. (2019) *Algoritmo Knn Para Classificação*. Disponível em <https://inferir.com.br/artigos/algoritimo-knn-para-classificacao/>. Acesso em 24 maio 2022.
- NOGARE D. (2020) **Explicando o algoritmo de Regra de Associação**. Disponível em: <https://diegonogare.net/2020/05/explicando-o-algoritmo-de-regra-de-associacao/>. Acesso em 24 maio 2022.
- PINELA C. (2017) *Recommender Systems – User-Based and Item-Based Collaborative Filtering* Disponível em: <https://medium.com/@cfpinela/recommender-systems-user-based-and-item-based-collaborative-filtering-5d5f375a127f>. Acesso em 24 maio 2022.
- REIS T. (2019) **CAGR: o que é? Qual sua importância nos investimentos?** Disponível em: <https://www.suno.com.br/artigos/cagr/>. Acesso em 24 maio 2022.
- SANTANA, R. V., PONTES, H. L. J. (2020). **Aplicação da Clusterização por K-means para Criação de Sistema de Recomendação de Produtos baseado em Perfis de Compra**. Navus: Revista de Gestão e Tecnologia, 10. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=7774814>. Acesso em: 20 mar. 2022.
- SARWAR B., KARYPIS G., KONSTAN J., RIEDL J. (2001) *Item-based collaborative filtering recommendation algorithms*. Disponível em <https://www.ra.ethz.ch/cdstore/www10/papers/pdf/p519.pdf> acesso em 08 maio 2022.
- SILBERSCHATZ A. (2006) **Sistema de Banco de Dados**. Haifa, Israel: Elsevier 781 p.
- SILVA J. F. G. (2019) Projeto de Graduação “**Efeitos da Esparcidade em Sistema de Recomendação**”. UFRJ (Escola Politécnica) Disponível em <https://monografias.poli.ufrj.br/monografias/monopoli10027591.pdf> Acesso em 24 maio 2022
- STATISTA (2022a) *Retail e-commerce sales worldwide from 2014 to 2025 (in billion U.S. dollars)*. Disponível em: <https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/>. acesso em: 24 maio 2022.
- STATISTA R. D. (2022b) *E-commerce retail sales CAGR 2022-2025, by country*. Disponível em: <https://www.statista.com/forecasts/220177/b2c-e-commerce-sales-cagr-forecast-for-selected-countries>. acesso em: 24 maio 2022.
- STRONSKI P. N., (2018) “**O que são redes bayesianas?**” Disponível em: <https://www.electricalibrary.com/2018/07/09/o-que-sao-redes-bayesianas/>. Acesso em 24 maio 2022.
- VELOSO M. J. P. S. A. (2003) Tese de Mestrado “**Regras de Associação Aplicadas a um Método de Apoio ao Planejamento de Recursos Humanos**” Disponível em: <https://repositorio-aberto.up.pt/bitstream/10216/7574/2/teseMiguelVeloFinal.pdf> Acesso em 24 maio 2022.