
COMPORTAMENTO DO CLIENTE: PROPOSTA DE UM SISTEMA DE INFORMAÇÃO PARA SUPORTE A TOMADA DE DECISÃO SOB A ÓTICA CONCEITUAL DA CIÊNCIA DE DADOS

Cláudio Sonáglio Albano
claudioalbano@unipampa.edu.br
Professor da Universidade Federal do Pampa/UNIPAMPA

Yura José Caldeira da Silva
yurasilva.aluno@unipampa.edu.br
Graduado em Engenharia de Produção, UNIPAMPA, Campus Bagé.

Resumo

Conhecer o perfil dos clientes pode ser um diferencial para uma organização obter vantagem competitiva. Este trabalho teve como objetivo propor a inserção de variáveis comportamentais dos clientes em uma base de dados, propiciando que este sistema possa apoiar o gestor, visando prever o comportamento dos clientes por meio da utilização de recursos e ferramentas tecnológicas aliadas a ciência de dados. Foi realizada uma pesquisa bibliográfica que resultou em um conjunto de variáveis comportamentais de clientes, estas variáveis foram validadas por entrevistas, posteriormente inseridas em um sistema de informação. Executados algoritmos sob a ótica conceitual da Ciência de Dados. Destacam-se como resultados: a viabilidade do uso de ferramentas livres para tarefas relacionadas à análise de dados; a pertinência da análise de dados para a descoberta de conhecimento para pequenas e médias empresas; e finalmente a pertinência dos resultados para apoio a tomada de decisões referentes ao perfil dos clientes.

Palavras-chaves: análise de dados; clientes; perfil; sistemas de informação; tomada de decisão.

Abstract

Knowing the customer profile can be an differential for an organization to gain competitive advantage. This work aimed to propose the inclusion of behavioral variables of customers in a database, providing that this system can better support the manager, aiming to predict the behavior of customers through the use of resources and technological tools allied to data science. A bibliographic research was carried out that resulted in a set of behavioral variables of clients, these variables were validated by interviews, later inserted in an information system. Executed algorithms under the conceptual optics of Data Science. The following results stand out: the feasibility of using free tools for tasks related to data analysis; the pertinence of data analysis for knowledge discovery for small and medium-sized companies; and finally the pertinence of the results to support decision-making regarding the clients' profile.

Keywords: data analysis; customers; profile; information systems; decision making.

1. Introdução e Justificativa

Com os avanços tecnológicos, a Tecnologia da Informação e Comunicação (TIC) é uma grande aliada dos gestores organizacionais em diversas tarefas na tentativa de obter vantagem competitiva. Neste sentido, em um mercado cada vez mais competitivo, ter conhecimento sobre o cliente é essencial para que uma empresa obtenha vantagem competitiva em relação aos seus concorrentes.

Neste cenário a TIC assume um papel preponderante pelas atuais capacidades de acesso, armazenamento e recuperação de dados (e/ou informações), em geral, mas no caso deste trabalho com foco sobre os clientes. No entanto, ter um grande volume de dados (e/ou informações), apenas, não é suficiente. Uma organização deve dispor de informações certas no momento certo, com qualidade e precisão, visando melhorar seus processos de tomada de decisão, (GUIMARÃES JUNIOR et. al, 2021).

Para Mendonça (2017) são necessários sistemas de informação (SI) e ferramentas tecnológicas, que possam manipular este grande volume de dados, em busca de informações e conhecimento que possam apoiar os gestores. O autor cita como exemplo, a utilização de SI específicos, tais como o *Customer Relationship Management* (CRM) ou Gestão de Relacionamento com o Cliente. Este sistema visa dotar a organização de maior conhecimento sobre seus clientes, a partir dos dados destes clientes em suas interações anteriores com a organização.

Assim, para que uma organização consiga atender as necessidades e expectativas dos seus clientes de maneira satisfatória, e obter a vantagem competitiva, ela precisa compreender quais as variáveis, e fatores comportamentais, podem influenciar o comportamento dos clientes, em especial para direcionar suas compras e/ou intenção de consumo.

Neste sentido, e usando recursos da TIC, nos últimos anos tem ganhado ênfase a área da “ciência de dados”. Para Provost & Fawcett (2016), ciência de dados é um ramo multidisciplinar que envolve técnicas de computação, inteligência artificial, estatística e matemática, com o intuito de resolver problemas analiticamente complexos. Possibilita a descoberta de conhecimento a partir de dados e compreende um conjunto de etapas que envolvem desde a organização e limpeza dos dados até a validação e o uso da informação (e/ou conhecimento) dela obtida. Ainda segundo estes autores esta área pode ser extremamente útil em tarefas que buscam identificar padrões de comportamentos em grandes bases de dados.

Entretanto, segundo alguns autores, ainda existem lacunas com relação a esta área (ciência de dados), em especial quando se trata de sua utilização junto a pequenas e médias empresas. Isto se deve, pois estas organizações, embora extremamente importantes para o cenário econômico e social, em especial no Brasil, tem algumas características que muitas vezes impedem o acesso delas ao uso de recursos de TIC e de profissionais especialistas em determinadas áreas, (ALONSO & KOK, 2020; e MEDEIROS, et. al. 2021).

Não restam dúvidas da importância das pequenas médias empresa para o cenário econômico brasileiro. Para o Sebrae (2020), o Brasil tem aproximadamente 20 milhões de empresas, deste total a soma das micro, pequenas e médias empresas totalizam aproximadamente quase 90% do total de empresas formalmente estabelecidas e representam aproximadamente um terço do Produto Interno Bruto brasileiro (PIB).

Assim, diante da representatividade destas organizações no cenário econômico brasileiro, justifica-se lançar o olhar sobre algumas dificuldades e desafios para estas organizações, tais como o uso dos potenciais recursos advindos da análise de dados. Visando contribuir para a mitigação dos problemas de gestão destas organizações, autores como Palhares Júnior et. al., (2019) e Moura et. al., (2019), propõem algumas ações, entre estas, um melhor conhecimento do mercado, em especial dos clientes.

Autores como Alonso & Kok (2020), reforçam a importância das pequenas e médias organizações desenvolverem um maior e melhor conhecimento sobre seus clientes de forma a obterem maior competitividade.

Desta forma, se justifica o desenvolvimento deste trabalho, atuando sobre um Sistema de Informação (SI) com o objetivo de propor a inserção de variáveis comportamentais dos clientes em uma base de dados, propiciando que este sistema possa melhor apoiar o gestor, visando

predizer o comportamento dos clientes por meio da utilização de recursos e ferramentas tecnológicas aliadas a ciência de dados.

Ao identificar métodos e modelos de análise de dados que possam ser utilizados sobre uma base de dados, e que sejam acessíveis tecnicamente, financeiramente e operacionalmente a pequenas e médias empresas o trabalho deverá contribuir para atender a lacuna exposta anteriormente. A viabilidade técnica, financeira e operacional será oportunizada pela utilização de ferramentas livres, quando da aplicação (execução) dos algoritmos. A utilização do software livre como ferramenta técnica em tarefas de análises de dados é visualizada como um das grandes potencialidades atualmente (COSTA et. al., 2017).

Para atender ao objetivo proposto o trabalho está estruturado da seguinte forma: esta seção apresentou a introdução e justificativa. Na seção seguinte apresentamos os conceitos e teorias que sustentam o desenvolvimento do mesmo, após os procedimentos metodológicos desenvolvidos, os resultados e respectivas análises e finalmente as conclusões.

2. Referencial Teórico

Neste tópico são abordados conceitos e teorias que suportam e justificam o desenvolvimento do trabalho. Primeiro apresenta-se o atual panorama competitivo entre as organizações e as potenciais contribuições de um Sistema de Informação. Na sequência, apresenta-se a importância da utilização dos dados como potencial para contribuir para uma melhor gestão, em busca de vantagem competitiva, para as organizações; e finalmente aborda-se as questões sobre o comportamento do cliente.

2.1 Sistemas de Informação (SI) – suporte a tomada de decisão.

No novo cenário competitivo das organizações um eficaz e eficiente Sistema de Informação (SI) torna-se cada vez mais importante. Podemos conceituar SI como um aglomerado de elementos que se relacionam, coletando, manipulando, armazenando e disseminando dados, e informações, que proporcionam mecanismos para suporte a um melhor processo de tomada de decisão que pode colaborar para que as organizações possam atingir seus objetivos (STAIR & REYNOLDS, 2016). Ainda para estes autores, um SI deve favorecer a interpretação e utilização dos dados, e informações, com o objetivo de tornar a organização mais competitiva, com melhor visão sobre si, sobre seus clientes e sobre seus concorrentes.

Os Sistemas de Informações (SI) vêm ganhando cada vez mais espaço no ambiente organizacional, sendo de fundamental importância para otimização dos processos e aumentando o nível de confiabilidade das informações. Por meio do SI, é possível gerenciar, tratar e filtrar todos os dados disponíveis, descartando os que não forem úteis e dando prioridade para aqueles que podem ser importantes para auxiliar e servir de base para a empresa tomar as melhores decisões em relação ao negócio. Com este diferencial, a empresa pode se destacar em um mercado de constantes mudanças e cada vez mais competitivo (MACHADO & NOGUEIRA, 2018).

Para Rodrigues (2021), a tomada de decisão é um dos papéis mais importantes de um gestor. Conforme a autora, a tomada de decisão pode ser conceituada como um processo de escolha a respeito de alguma coisa, envolvendo a decisão de um conjunto de ações a serem seguidas entre duas ou mais possibilidades disponíveis, objetivando chegar à resolução de um determinado problema e melhorar a *performance* da organização.

De acordo com Prado (2020), a tomada de decisão constitui-se no processo decisório de definir entre as alternativas possíveis e já existentes. Para a autora, o responsável pelas decisões

são os gestores, nos mais diversos níveis, incumbidos de relacionar a organização com seu planejamento estratégico.

O processo de tomada de decisão é importante pelo fato de que neste momento é definida as ações que serão realizadas no futuro. Toda decisão deve ter embasamento em informações pertinentes, sendo a principal fonte de informações um eficaz e eficiente Sistema de Informação (MACHADO & NOGUEIRA, 2018).

Atualmente, com as novas possibilidades tecnológicas, a utilização dos dados, e informações, armazenadas no SI de uma organização, ganham uma nova dimensão, pela abordagem destes por intermédio de algoritmos que tem como objetivos promoverem novas possibilidades de entendimento e aprendizado com estes dados, e informações, visando um melhor suporte a tomada de decisão, com objetivo de possibilitar vantagem competitiva a uma organização.

2.2 Ciência de Dados: dados e algoritmos para suporte a tomada de decisão.

Nos últimos anos uma nova área do conhecimento ganha importância no contexto da gestão das organizações, fazendo uso dos dados armazenados nos Sistemas de Informação e na utilização de algoritmos específicos para determinadas atividades, esta área é a Ciência de Dados.

Para Provost & Fawcett (2016) Ciência de Dados é campo de conhecimento multidisciplinar que envolve computação, inteligência artificial, estatística, matemática, além de conhecimentos de gestão, para desenvolver um processo de descoberta de conhecimentos a partir dos dados. É composto por uma série de etapas que envolvem desde a organização e limpeza dos dados, transformação destes em informação, e conhecimento, e finalmente validar o uso dos resultados obtidos.

Ciência de dados é implementada ou viabilizada, por intermédio de diversas ferramentas tecnológicas, entre estas a mais importante são os algoritmos que viabilizam algumas implementações, tais como: mineração de dados, aprendizado de máquina, entre outras.

Assim, os algoritmos têm se tornado uma importante ferramenta para apontar sugestões e realizar previsões a partir de uma base de dados, auxiliando os gestores no processo de tomada de decisão. Um algoritmo de apoio à decisão realiza um prognóstico com base em avaliações estatísticas dos dados de um Sistema de Informação (SI), (ENGEMAN, 2022).

De acordo com Pierro (2018), um algoritmo define-se como uma sucessão de passos escritos em linguagem de programação de computador, que tem como finalidade a resolução de um problema ou a realização de uma atividade de maneira automatizada, independentemente da quantidade de linhas de programação que ele contemple.

Para Engeman (2022), com a correta utilização de um algoritmo, é possível minimizar os riscos e maximizar as perspectivas de atingir os resultados desejados. Lopes (2021) complementa afirmando que, como benefícios, pode-se mencionar o processo de tomada de decisão mais assertivo, exploração e melhoria dos processos, correções de sistemas, entre outros.

Existem dois grandes grupos de algoritmos, os supervisionados e os não supervisionados. Neste trabalho utilizaremos algoritmos supervisionados, estes segundo Pierson (2019) atuam sobre dados de entrada rotulados, com características conhecidas que possuam valores históricos, que possibilitem previsões com boa margem de acerto para eventos futuros. Ainda segundo Pierson (2019) existem algoritmos adequados e/ou desenvolvidos para cada função. Neste trabalho serão utilizados algoritmos coerentes com o método *Naive Bayes*, que são indicados

quando desejamos obter a probabilidade de um evento ocorrer conforme evidências em seus dados históricos.

2.3 Comportamento do cliente.

O campo do comportamento do cliente é amplo, podendo ser considerado como o estudo dos processos envolvidos quando uma pessoa escolhe, adquire, utiliza ou até mesmo se desfaz de produtos e/ou serviços, com a finalidade de satisfazer suas exigências e aspirações. Uma empresa somente conseguirá satisfazer estas necessidades quando compreender os fatores que influenciam as escolhas dos clientes (SOLOMON, 2016).

O comportamento de compra pode estar relacionado aos gostos e preferências pessoais de um cliente, influenciando na maneira como ele consome os mais variados tipos de produtos ou serviços. Diretamente relacionadas a essas preferências, estão as condições financeiras e o potencial de compra do cliente. Quanto mais caro o produto, mais longo tende a ser o tempo de tomada de decisão por parte do cliente. Além disso, a interação do cliente com uma determinada marca, campanhas de *marketing* eficazes por parte das empresas e a influência de um determinado grupo de pessoas, sejam amigos ou outros clientes anunciados, também podem interferir no comportamento e gerar incentivos para decisões de compra (SILVA, 2020).

Ainda de acordo com Silva (2020), a fim de efetuar um diagnóstico em relação ao comportamento e hábito de compra do cliente, tem de se considerar qual o sentimento dos clientes em relação às opções de marcas, mercadorias ou serviços; quais fatores influenciam o cliente a optar por uma ou outra opção de produtos ou serviços; quais os hábitos do cliente ao pesquisar produtos e realizar a compra; e de que forma o ambiente de consumo pode motivar a decisão de compra.

Com os avanços tecnológicos e as facilidades proporcionadas por este fato, é cada vez mais imprescindível que as empresas estejam preparadas para oferecer um serviço eficiente, de qualidade e que transmita confiança aos seus clientes. As mídias digitais proporcionam às empresas um canal de comunicação que, se bem utilizado, permite antecipar as necessidades dos clientes e a realização de um atendimento mais assertivo (COSTA; SANTANA; TRIGO, 2015).

O comportamento dos clientes vem sofrendo alterações em função dos avanços comportamentais e tecnológicos ocorridos nos últimos anos. O novo perfil de cliente é digital, que utiliza a internet com facilidade, obtendo informações rápidas, comparando marcas e preços, verificando pontos de venda próximos e, por isso, sendo menos fiel a uma marca. Assim, os gestores compreender este cenário e conhecer este novo perfil de clientes (SHIMOYAMA, 2017).

Conforme a Agência Sebrae de Notícias (2020), com o passar do tempo, os clientes vêm se tornando cada vez mais conectados e adaptáveis aos serviços de compras *online*, buscando por serviços que sejam personalizados e inovadores. Com o período de pandemia, as alterações no comportamento do cliente são ainda mais visíveis, principalmente com a intensificação do uso de ferramentas digitais. Além disso, também estão mais exigentes e preocupados, principalmente no que diz respeito à limpeza e higienização dos estabelecimentos e a adesão aos protocolos de prevenção à contaminação por COVID-19.

Pelo fato de os clientes estarem criando um novo perfil, que está diretamente relacionado com fatores tecnológicos e impactos comportamentais, o presente trabalho se faz importante. Conhecer, compreender e atender as necessidades deste novo perfil de clientes é a base para que qualquer empresa obtenha uma estratégia de marketing bem-sucedida.

Para Coser (2019), perante uma era digital altamente competitiva, é essencial que as empresas consigam identificar o perfil de seus clientes, antecipando e satisfazendo

eficientemente suas necessidades, com o objetivo de conquistar o seu público-alvo e fidelizá-los como clientes. Baseado no conhecimento em relação aos clientes, torna-se possível elaborar e implantar estratégias de *marketing* eficientes para obtenção de melhores resultados comerciais para a organização, além de possibilitar a otimização de atendimentos. Identificar as informações de seus clientes facilita a determinação de quem são seus potenciais compradores e quem pode estar interessado em seus produtos ou serviços. Informações como sexo, idade, renda, localização e *hobbies* são dados interessantes que um estabelecimento tenha conhecimento a respeito dos seus clientes.

2.4 Sistemas Informações; Ciência de Dados e Comportamento dos Clientes

A partir do referencial teórico apresentado, podemos observar que um procedimento importante para uma organização obter sucesso é saber identificar as características de seus clientes, bem como seus padrões de compra e comportamento, para criar um perfil sobre os mesmos. Para isto, é importante realizar uma coleta de dados correta e precisa em relação aos seus clientes.

Para Guimaraes Junior et. al., (2021), ao utilizar um SI para atuar sobre o comportamento do cliente, as organizações podem criar vantagens competitivas em seus mercados. Estes sistemas contribuem para que as organizações consigam identificar o perfil e os hábitos dos seus clientes, se tornando mais assertivas no atendimento de suas necessidades e tomando decisões mais inteligentes.

Assim, possuir um Sistema de Informação atualizado, organizado e que se adapte às mudanças do mercado é um aspecto fundamental para as organizações, pois além de colaborar para o crescimento dos negócios, contribui para a construção de uma estratégia centrada no cliente, convertendo os dados em informações úteis aos gestores. Estas informações podem ser obtidas através dos algoritmos, que são especialmente úteis quando grandes quantidades de dados devem ser processadas para se tomar uma decisão com base em informações, de maneira mais rápida.

Desse modo, é possível afirmar que compreender sobre o comportamento do cliente é importante para construção de um plano de marketing preciso, enquanto o Sistema de Informação (SI) reúne todo um conjunto de elementos sobre os clientes, que podem ser tratados e processados por algoritmos.

Os algoritmos são capazes de realizar agrupamentos de dados de acordo com características compartilhadas, ou ainda, realizar a associação destes dados para descobrir padrões e fazer previsões sobre eventos futuros, fornecendo os parâmetros para as tomadas de decisões. A tomada de decisão envolve a ponderação de múltiplas considerações, onde a escolha da melhor opção é de responsabilidade final de cada gestor (CAETANO, 2006).

Relacionando os algoritmos com o contexto do trabalho, eles podem contribuir para a identificação de padrões nos hábitos de compra dos clientes. Com o uso de um algoritmo é possível que uma organização aprenda sobre as preferências dos clientes e quais os produtos e/ou serviços eles mais consomem, sendo possível ajustar suas estratégias de marketing.

3. Procedimentos Metodológicos.

Neste tópico estão descritos os procedimentos metodológicos desenvolvidos e que suportaram o atendimento ao objetivo do trabalho, estes são os seguintes: a) pesquisa bibliográfica, que consistiu em uma revisão da literatura, visando conhecer fatores e/ou variáveis comportamentais dos clientes; b) posteriormente, foi realizada uma pesquisa utilizando o método

de coleta de dados *Delphi*, com gestores de pequenas e médias empresas; c) após foram realizadas entrevistas com consumidores visando identificar as variáveis que mais impactam seu comportamento de compra; e d) finalmente estas variáveis foram adicionadas a uma base de dados, de um sistema de informação, sobre esta base de dados foram executados algoritmos, sob a ótica conceitual da Ciência de Dados, visando prever o comportamento dos clientes.

Ainda dentro do tópico procedimentos metodológicos é realizada uma breve descrição da plataforma tecnológica utilizada, o software *KINME*, e a forma de utilização deste, que é um software de código aberto, gratuito e contém diversas técnicas para aprendizado de máquina. Para maiores informações consultar: <https://www.knime.com/>.

3.1 Pesquisa bibliográfica

Na primeira etapa da pesquisa, foi realizada uma busca no *site* da Associação Brasileira de Engenharia de Produção (ABEPRO) de todos os Encontros Nacionais de Engenharia de Produção (ENEGEP). Esta pesquisa atendeu ao período de 1996 até o ano de 2021, inserindo como único critério de pesquisa a palavra clientes, que é identificada no título e no corpo do texto dos artigos. Posteriormente foi realizada uma pré-seleção destes artigos encontrados na pesquisa e, com base no tema e título do trabalho. O critério adotado para a seleção dos artigos foi que esses abordassem o comportamento e/ou satisfação dos clientes, bem como os fatores e/ou variáveis que influenciam o comportamento de compra dos clientes. Outro critério estabelecido foi que o artigo deveria trabalhar no contexto de empresas do ramo do comércio.

3.2 Método Delphi

A segunda etapa da pesquisa ocorreu com a realização de uma coleta de dados, a partir dos resultados da pesquisa bibliográfica. Para esta coleta de dados, foi utilizado o método de coleta de dados *Delphi*. Esta metodologia é composta de um conjunto de questionários com o objetivo de coletar e organizar opiniões e respostas de um determinado grupo de pessoas, composto por especialistas no campo em estudo, sendo bastante empregada como ferramenta para associar julgamentos individuais. Após o tratamento das respostas, os resultados são devolvidos aos participantes para reavaliar suas respostas e compará-las com o restante do grupo (CUTRIM *et. al.*, 2016). Esta etapa foi desenvolvida em dois momentos, descritos a seguir.

3.2.1 Primeira rodada – coleta de dados, método *Delphi*.

Realizada para solucionar uma possível similaridade entre os fatores e/ou variáveis comportamentais, resultantes da pesquisa bibliográfica. Dessa forma, procurou-se reduzir esta quantidade de fatores e/ou variáveis comportamentais iniciando a primeira rodada da *Delphi*, contando com a participação de dez respondentes com experiência e conhecimento no contexto do trabalho.

Esta coleta de dados ocorreu por intermédio da visita presencial, quando foi aplicado um questionário com o objetivo de utilizar a experiência e o conhecimento dos entrevistados para verificar quais variáveis, dentre todas encontradas na pesquisa bibliográfica, são semelhantes. Com isto, foi possível reduzir alguma possível redundância de variáveis e diminuir a quantidade de questões para as rodadas posteriores.

Os respondentes escolhidos utilizando os métodos de amostragem por conveniência, para Hair *et. al.*, (2005) este tipo de amostra pode ser utilizada dependendo das condições de acesso aos elementos da população, a representatividade desejada, a oportunidade de contar com esses

elementos e a disponibilidade de recursos financeiros, humanos, técnicos, entre outros, para compor sua amostra. A primeira rodada da *Delphi* teve caráter qualitativo para análise de dados.

3.2.2 Segunda rodada – coleta de dados, técnica *Delphi*

Coleta de dados, também realizada de forma presencial com objetivo de aplicar um questionário com objetivo de agrupar as variáveis resultantes da primeira rodada, descrita na etapa 3.21. O respondente deveria indicar em uma escala *Likert* de cinco pontos, onde 1 significava não uso/não indico (pouco importante) e 5 significava uso/indico (muito importante), qual a importância de cada variável no comportamento do cliente durante o processo de compra.

3.3 Pesquisa com clientes

Com o objetivo de avaliar o ponto de vista do cliente e verificar quais os fatores que influenciam o seu comportamento de compra, foi realizada uma pesquisa com cinquenta clientes, utilizando a técnica de amostra por conveniência. Para isto, foi elaborado um questionário através da plataforma de formulários do *Google*. Neste formulário, foi solicitado que o participante indicasse, dentre as quinze variáveis comportamentais validadas com o método *Delphi*, três variáveis que mais influenciam a sua decisão de compra. Posteriormente, estas variáveis indicadas foram inseridas em uma base de dados, do Sistema de Informação (SI).

3.4 Inserção de atributos na base de dados

Em parceria com uma empresa desenvolvedora de *software*, esta disponibilizou uma base de dados para o desenvolvimento do trabalho. Inicialmente, esta base estava no formato “.*csv*” e continha mais de um milhão de registros, dividida em quinze colunas de informações.

O primeiro passo para tornar possível a manipulação, foi importar esta base para o formato de uma planilha eletrônica. Após a conversão de formatos, com o objetivo de facilitar o trabalho e manuseio dos dados, a base foi reduzida para as primeiras mil e quinhentas linhas de compras. Este fato foi necessário para viabilizar o tratamento da planilha no software utilizado para a execução dos algoritmos.

Em relação às colunas, foram mantidas as quinze que já existiam e criadas quatro novas. A primeira coluna adicionada teve como objetivo identificar o código do cliente, que era uma informação ausente até então. Estes códigos foram inseridos de 1 até 50, de forma aleatória, a partir de um gerador de números da planilha eletrônica, de acordo com a coluna *NroCupom*. Dessa forma, cada número de cupom recebeu o código de um cliente aleatoriamente. As outras três colunas foram inseridas visando identificar as três variáveis comportamentais indicadas pelos clientes.

Após atribuir os códigos dos clientes foram inseridas as respostas obtidas na pesquisa com clientes, descrita no tópico 3.3. Exemplificando o processo: o cliente de código um na base de dados, recebeu as três variáveis indicadas pelo primeiro participante da pesquisa com clientes, sendo estas variáveis registradas em cada uma das três colunas criadas nesta base. Após, o cliente de código dois recebeu as três variáveis indicadas pelo segundo participante da pesquisa. Esta inserção de dados na planilha aconteceu de forma sucessiva, até o cliente de código cinquenta. Em caso de clientes com mais de uma compra, repetiu-se o código do cliente com suas respectivas variáveis comportamentais.

3.5 Plataforma *KNIME*

A plataforma *KNIME Analytics* é um *software* livre de código aberto utilizado para a integração e análises de dados, além possibilitar a construção de relatórios. É um software gratuito que abrange as principais técnicas de preparação e transformação de dados, somados ao aprendizado de máquina baseado (HUPDATA, 2022). Para Ribeiro (2016), o *KNIME* é um programa livre que propicia um ambiente acessível e intuitivo aos seus usuários, oferecendo técnicas avançadas de ciência dos dados.

Após utilizar os recursos do *KNIME* para preparar os dados da base, foi aplicado um algoritmo de análise preditiva chamado classificador bayesiano. Conforme Oliveira e Pereira (2017), o classificador bayesiano é um algoritmo estatístico que classifica objetos em classes e analisa a probabilidade de sua pertinência, predizendo qual a mais provável. Já para Gomes (2019), este algoritmo é bastante utilizado em aprendizado de máquina, justamente por se tratar de um classificador probabilístico baseado no Teorema de *Bayes*, sendo capaz de categorizar um texto de acordo com a frequência das palavras escritas.

Quando temos atributos qualitativos, nominais ou ordinais, este algoritmo pode ser usado para classificação destes objetos por meio das probabilidades de ocorrência ou frequências relativas (OLIVEIRA; PEREIRA, 2017). Para Sacramento (2021), o funcionamento do modelo bayesiano pode ser detalhado em termos estatísticos. Segundo o autor, para realizar as previsões, o algoritmo gera uma tabela de probabilidades com a frequência dos preditores em relação às variáveis de saída. Dessa forma, o cálculo final para a predição considera a probabilidade maior para oferecer uma solução.

Durante os testes, foram realizadas algumas manipulações na base de dados utilizando o recurso de filtrar as colunas, visando verificar se haveria discrepância entre resultados. O primeiro teste foi realizado mantendo as colunas das variáveis comportamentais, juntamente com o código do cliente, número do cupom, valor total e descrição do produto. No segundo teste, foram mantidas apenas as colunas referentes às variáveis e a descrição do produto. Para o terceiro teste, manteve-se a descrição do produto e analisaram-se as colunas das variáveis comportamentais separadamente.

Além do fluxo citado acima, também foi elaborado um modelo para verificação dos conjuntos de variáveis que foram indicadas de forma repetida pelos clientes na base de dados. Para isto, foi realizada a leitura da base e a filtragem das colunas, mantendo apenas as três colunas que continham as variáveis comportamentais mencionadas por estes clientes.

Na sequência, foi aplicado um nó responsável por combinar colunas e gerar uma única, contendo todos os dados combinados. Esta combinação foi realizada da seguinte forma: a) combinação da coluna 1, coluna 2 e coluna 3; b) combinação da coluna 1 e coluna 2; c) combinação da coluna 1 e coluna 3; e finalmente d) combinação da coluna 2 e coluna 3.

Após, cada combinação foi ligada a um nó *statistics*, objetivando a criação de uma tabela de ocorrências e de um histograma. Para melhor visualização dos resultados, foi aplicado um nó *bar chart* para reproduzir graficamente os conjuntos que se repetiram, juntamente com o número de ocorrência de repetição.

4 Resultados e Respectivas Análises

Este tópico tem como objetivo apresentar os resultados e respectivas análises. Este serão apresentados na mesma sequência da descrição dos procedimentos metodológicos, entretanto os resultados dos tópicos: 3.1 pesquisa bibliográfica; 3.2 método Delphi; e 3.3 pesquisa com clientes, serão apresentados juntos e de forma resumida, pois o objetivo principal do trabalho é o de propor a inserção de variáveis comportamentais dos clientes em uma base de dados,

propiciando que este sistema possa melhor apoiar o gestor, visando prever o comportamento dos clientes por meio da utilização de recursos e ferramentas tecnológicas aliadas a ciência de dados, sendo estes resultados descritos com maiores detalhes a partir do tópico 4.2.

4.1 Pesquisa bibliográfica, método Delphi e pesquisa com clientes.

Como resultados da pesquisa bibliográfica, obteve-se um total de 4085 resultados retornados. Deste total, 77 artigos foram pré-selecionados para *download* e leitura posterior, com base no tema e no título. Ao final desta etapa, restaram 51 artigos. Assim, foram encontradas cinquenta e oito variáveis comportamentais que podem influenciar a decisão de compra dos clientes.

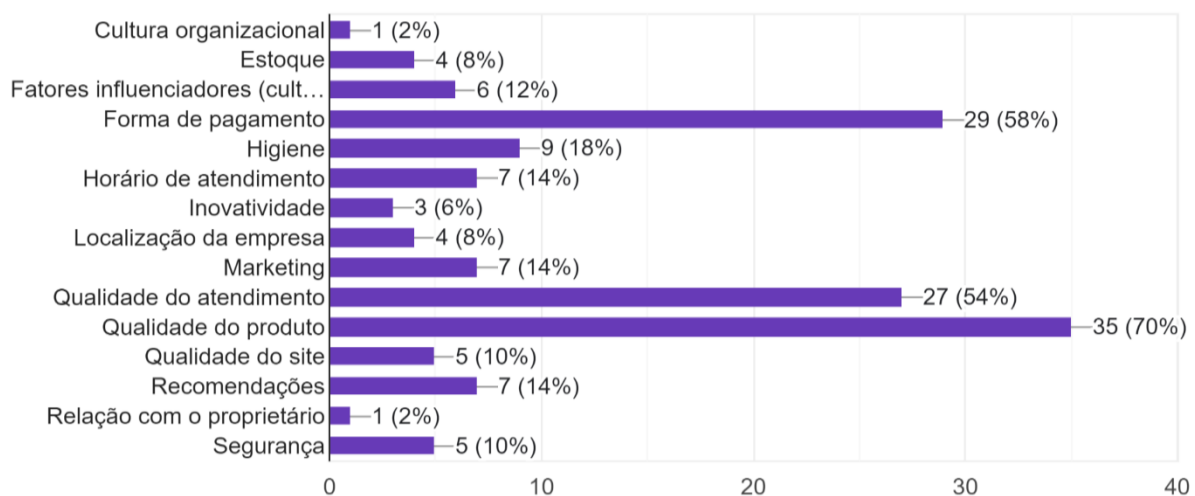
Mesmo após eliminar as repetidas, foi possível observar que poderia haver algum tipo de semelhança entre as variáveis encontradas. Dessa forma, optou-se pela realização de uma pesquisa com especialistas utilizando o método *Delphi*, visando verificar, por meio de consenso, quais variáveis podem ser consideradas semelhantes.

Após as duas rodadas de aplicação do método Delphi, restaram as seguintes variáveis comportamentais: cultura organizacional; estoque; forma de pagamento; higiene; horário de atendimento; inovatividade; localização da empresa; marketing; qualidade do atendimento; qualidade do produto; qualidade do site; recomendações; relação com o proprietário da empresa e segurança.

Neste ponto entendeu-se que foi alcançado um nível de estabilidade e consenso entre as respostas dos participantes das entrevistas, com base na literatura, optou-se por encerrar a pesquisa *Delphi* após a segunda rodada.

A próxima etapa do trabalho foi submeter estas variáveis a uma pesquisa com 50 clientes. Nesta pesquisa, os clientes, deveriam indicar três variáveis que mais impactam seu comportamento de compra, obtivemos como resultado as variáveis consideradas mais importantes no ponto de vista do cliente. A Figura 1 demonstra como foram os resultados desta pesquisa.

Figura 1 – Resultado da pesquisa com clientes



Fonte: Autores do trabalho (2023).

Os resultados desta pesquisa com os clientes foram coerentes com os resultados obtidos após a aplicação do método *Delphi*, com especialistas. Em ambas as pesquisas, as três variáveis citadas como mais importante, foram as seguintes: qualidade do produto; forma de pagamento e qualidade do atendimento.

4.2 Alterações na base de dados do Sistema de Informação.

Com os resultados da pesquisa com os clientes, foram realizadas modificações em uma base de dados, oriunda de um Sistema de Informação (SI), para agregar a esta base de dados as variáveis comportamentais resultantes da *Delphi*. Assim, foram inseridas três colunas na base de dados para registrar as variáveis comportamentais citadas pelo cliente no momento da compra, também foi adicionada uma coluna na base de dados referente ao código (identificação) do cliente.

Após o preenchimento das colunas referentes às variáveis comportamentais indicadas por cada cliente, foi realizada a aplicação dos algoritmos de predição por meio da plataforma *KNIME*.

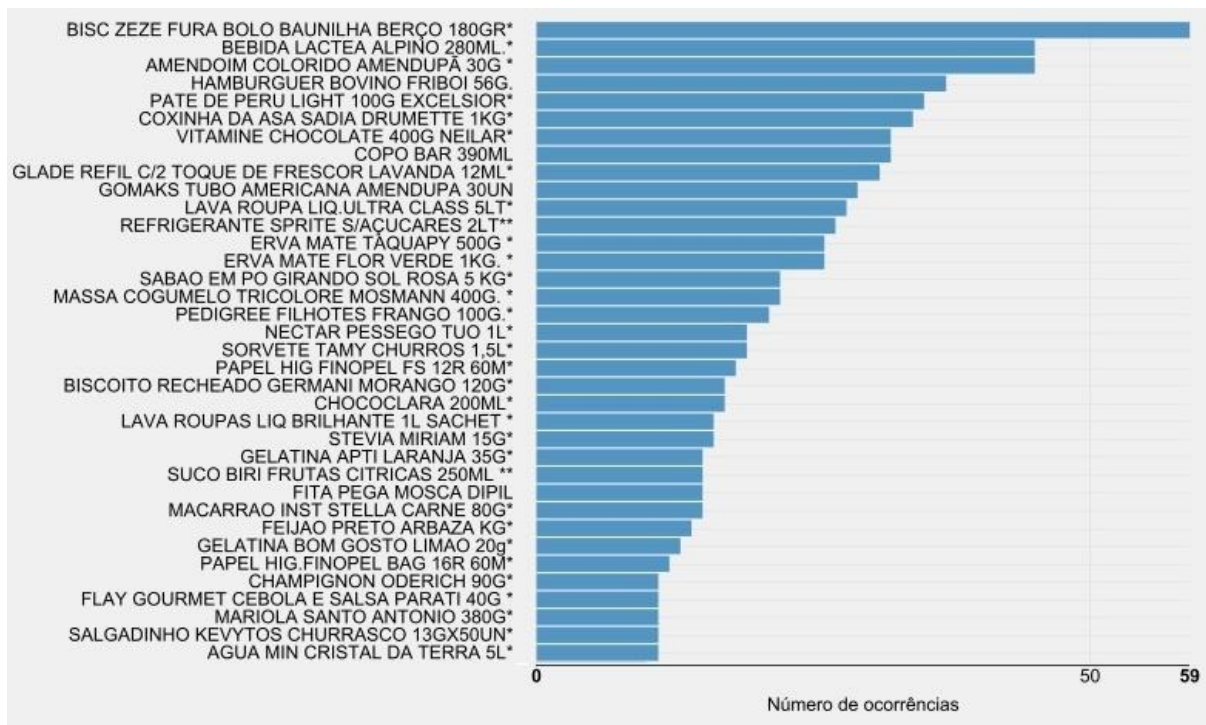
4.3 Plataforma *KNIME*: resultados e respectivas análises

Foram executadas algumas simulações na base de dados, visando prever o comportamento dos clientes, conforme suas variáveis (ou fatores) comportamentais e respectivas compras. Estas simulações, com respectivos resultados e análises estão demonstradas nos tópicos a seguir, quando foram realizados três testes com modelos distintos para análise das predições.

4.3.1 Modelo preditivo com as principais informações da base de dados

Para o primeiro teste, o modelo preditivo foi construído utilizando as principais informações da base de dados, quais sejam: código do cliente, valor da compra, produto comprado e as variáveis (fatores) comportamentais. Um recorte, com os produtos mais prováveis de serem comprados, está demonstrado na figura 2.

Figura 2 – Resultados do primeiro modelo.



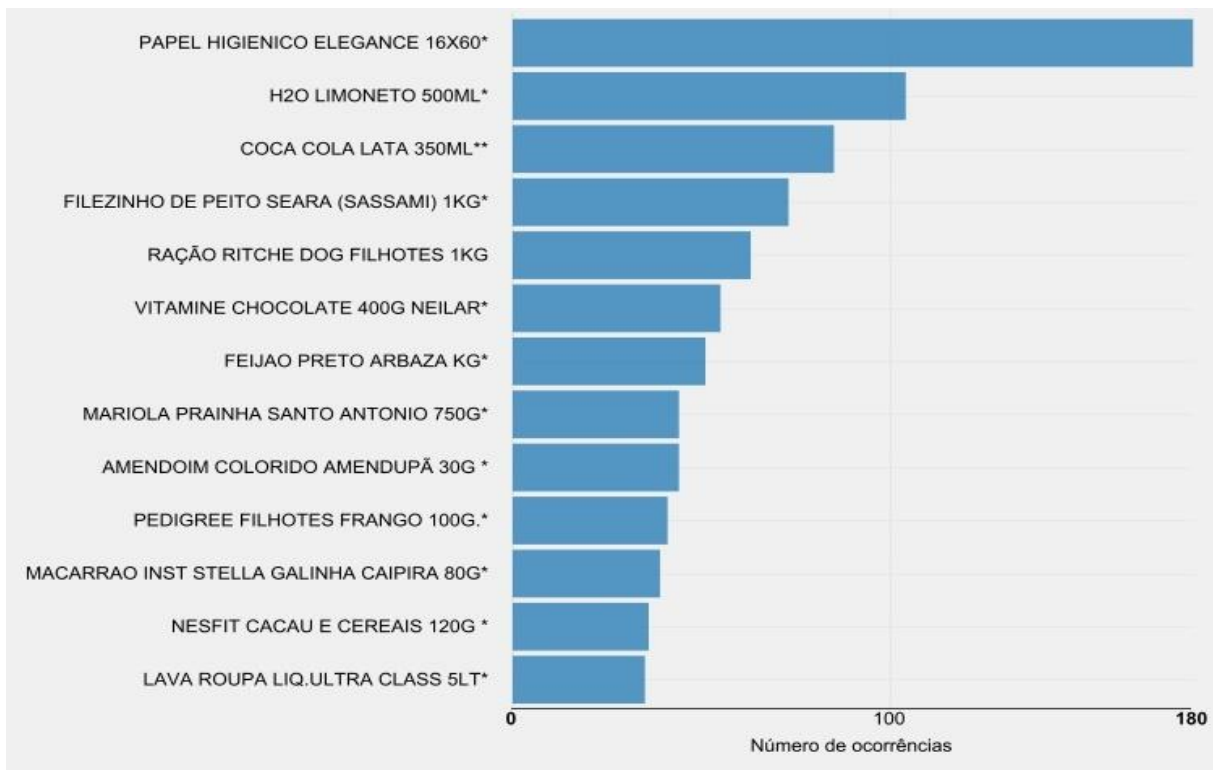
Fonte: Autores do trabalho (2023).

O gráfico demonstra os produtos com maior número de ocorrências segundo a análise preditiva. Podemos observar que, conforme as previsões deste modelo, o produto que tem a preferência dos clientes é o biscoito fura bolo de baunilha, com 59 ocorrências (em 1500 possibilidades).

4.3.2 Modelo preditivo usando apenas os produtos comprados e as variáveis

A segunda simulação foi realizada a partir de um modelo preditivo que utilizou apenas os produtos comprados e as três variáveis comportamentais. Os resultados obtidos são demonstrados na Figura 3.

Figura 3 – Resultados estatísticos do segundo modelo



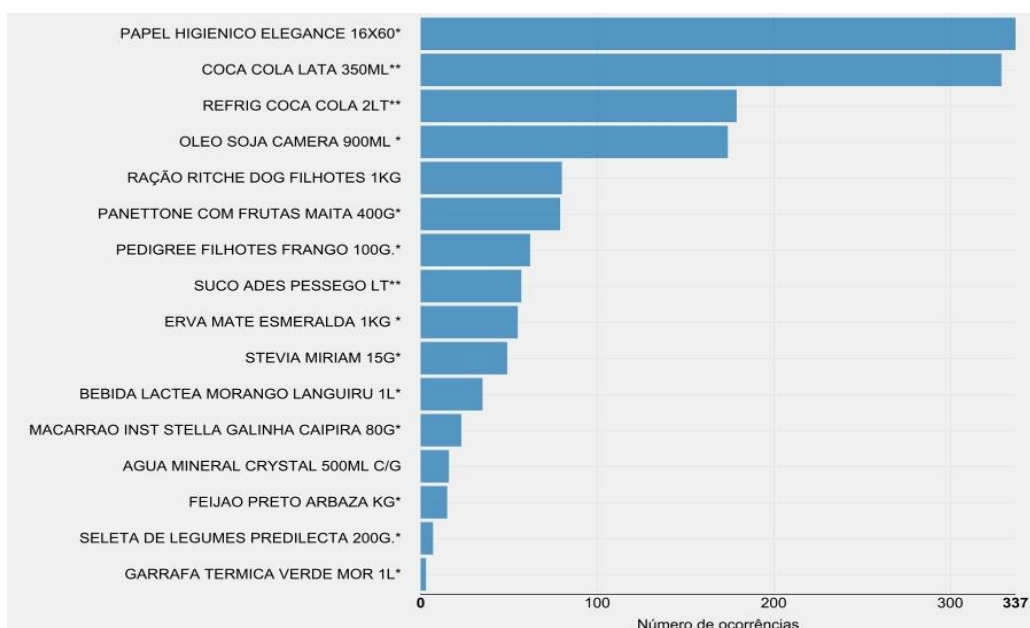
Fonte: Autores do trabalho (2023).

Semelhante ao teste anterior, o gráfico acima demonstra os produtos com maior ocorrência conforme a análise preditiva. Neste segundo modelo, podemos verificar alterações em relação aos produtos, onde o preferido pelos clientes passou a ser o papel higiênico 16x60, com 180 ocorrências (base de 1500 compras).

4.3.3 Modelo preditivo usando apenas as variáveis comportamentais

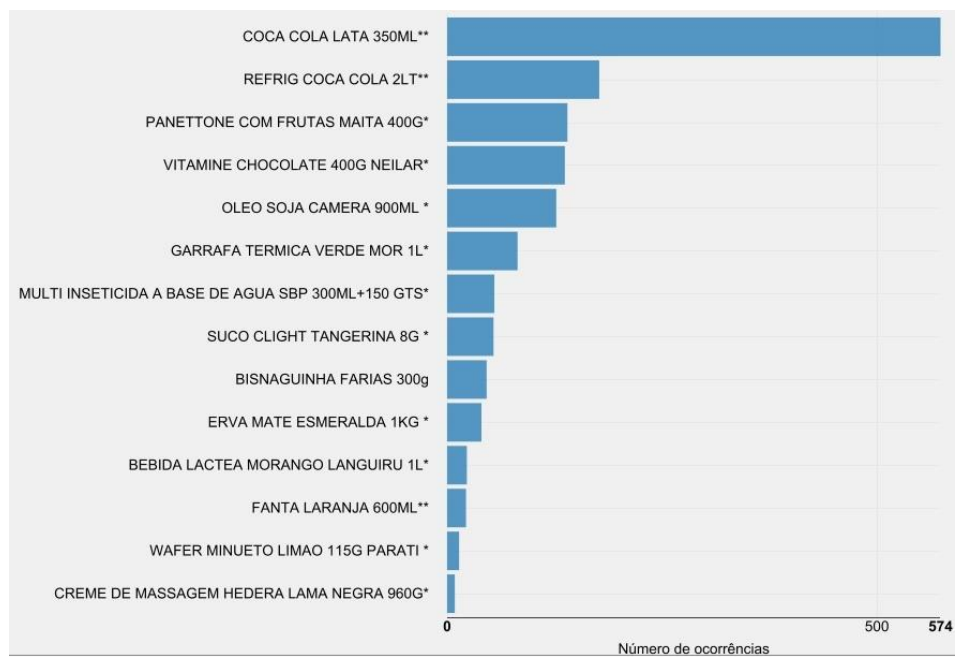
Nesta terceira rodada de simulação foram construídas três modelos de predição, cada um com uma variável comportamental. Nas figuras 4, 5 e 6, mostra-se um recorte dos resultados (produtos mais possíveis de serem comprados), quando o algoritmo foi respectivamente executado com a primeira, segunda e terceira variável comportamental.

Figura 4 – Resultados com a primeira variável comportamental.



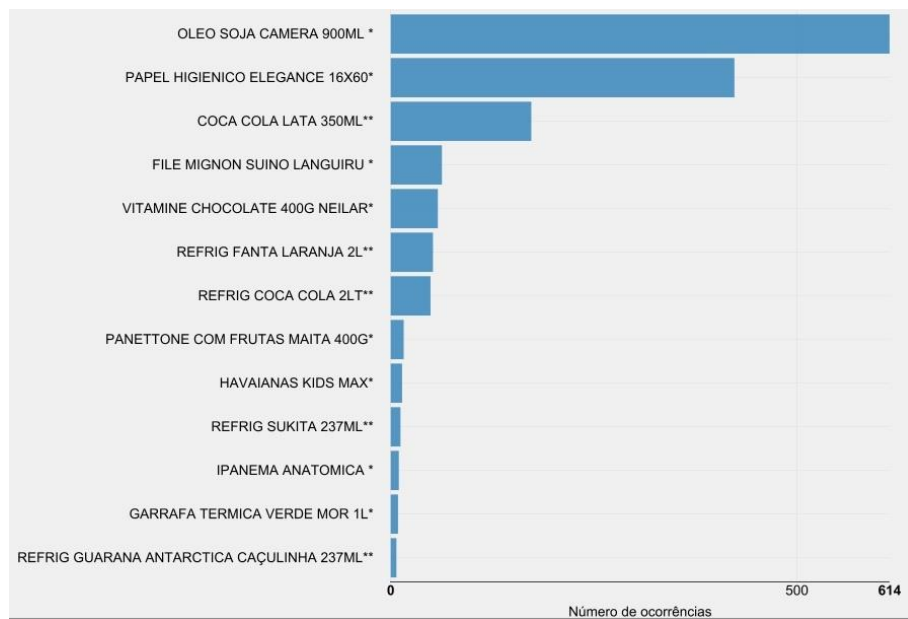
Fonte: Autores do trabalho (2023).

Figura 5 - Resultados com a segunda variável comportamental.



Fonte: Autores do trabalho (2023).

Figura 6 - Resultados com a terceira variável comportamental.



Fonte: Autores do trabalho (2023).

É possível verificar que, logicamente, conforme a variável comportamental, os resultados foram diferentes. Analisando os resultados obtidos com o auxílio do *software KNIME*, constata-se uma maior disparidade em relação às previsões de produtos realizadas no primeiro modelo. Além disso, este primeiro teste resultou em número maior de produtos gerados pela análise preditiva, com um número de ocorrência baixo.

Já em relação ao segundo e terceiro modelo, percebe-se a presença de uma maior quantidade de produtos semelhantes, principalmente no que diz respeito aos que possuem um número de ocorrência mais elevado. Estes testes resultaram em uma quantidade menor de previsão de produtos, porém o número de ocorrência foi superior, principalmente no modelo onde foram realizados testes com as colunas de variáveis separadas. O terceiro teste apresentou pouca divergência entre os produtos preditos, mudando apenas a sua ordem de acordo com o número de ocorrência.

Assim, o terceiro modelo se mostrou mais compatível com os objetivos do estudo, pois através dele foi possível analisar de forma independente todas as variáveis comportamentais citadas no momento da compra e qual sua relação com as preferências dos clientes por determinado produto.

5 Considerações Finais

Este trabalho atendeu aos seus objetivos, assim poderá contribuir para que seja observada a importância de conhecer o perfil do cliente, bem como seu comportamento e seus hábitos de consumo. O trabalho ainda tem potencial para proporcionar uma ampliação do conhecimento acerca dos temas abordados, aumentando a capacidade de realizar pesquisas mais aprofundadas, analisar bases de dados de empresas comerciais e propor ferramentas para solução de problemas reais, auxiliando gestores a tomarem as melhores decisões.

Este trabalho poderá contribuir com o surgimento de inovações a partir dos resultados alcançados, além de servir como fonte para pesquisas futuras a respeito do comportamento dos clientes e a utilização de Sistema de Informação como ferramenta de auxílio para a tomada de decisão.

Outra contribuição do trabalho é que gestores de empresas possam fazer uso de um Sistema de Informação que contenha variáveis comportamentais relacionadas com os clientes, dando suporte para a sua tomada de decisão. Assim, conhecendo melhor os hábitos de seus clientes, estas empresas poderão potencializar suas ações de *marketing*, proporcionar melhor experiência de compra, fidelizar clientes e alavancar as suas vendas.

Como sugestões para trabalhos futuros, propõe-se que sejam realizadas pesquisas com proprietários e gestores de outros ramos empresariais, buscando verificar outros fatores que podem influenciar no comportamento dos clientes. Além disso, sugere-se realizar uma pesquisa com clientes para verificar quais variáveis comportamentais eles consideram mais importantes, tendo em vista que neste estudo as variáveis foram classificadas por ordem alfabética.

Referências Bibliográficas

- ALONSO, A. D.; KOK, S.K. Understanding critical success factors and perceived future among micro and small firms through entrepreneurial action theory. 2020. *European Business Review*, 33 (2), 383-406. DOI: 10.1108/EBR-10-2019-0243.
- CAETANO, M. A. L. Lógica fuzzy para tomada de decisão em negócios e finanças. *Revista de Economia e Administração*, v. 5, n. 1, p. 12-39, 2006.
- COSER, Giulia. A Importância de conhecer o Perfil dos Clientes. Sorocaba, SP: [s. n.], 2019. Disponível em: <https://metajunior.com.br/blog/2019/07/14/a-importancia-de-conhecer-o-perfil-dos-clientes/>. Acesso em: 11 nov. 2021.
- COSTA, Ariana de Sousa Carvalho; SANTANA, Lídia Chagas de; TRIGO, Antônio Carrera. Qualidade do atendimento ao cliente: um grande diferencial competitivo para as organizações. *Cairu: Revista de Iniciação Científica*, v. 2, n. 2, p. 155-172, 2015. ISSN 2258-1166. Disponível em: https://cairu.br/riccairu/pdf/artigos/2/10_QUALIDADE_ATEND_CLIENTE.pdf. Acesso em: 28 jul. 2022.
- COSTA, Lucas Rodrigues; SILVEIRA, Lucas Ângelo; BRITO, Rennie Fagundes; SHINTAKU, Milton. Uso de software livre para disseminação e análise de dados abertos governamentais. *Informação & Tecnologia (ITEC)*, Marília/João Pessoa, v.4, n.1, p.03-20, jan./jun. 2017.
- CUTRIM, Sérgio Sampaio; TRISTÃO, José Américo Martelli; TRISTÃO, Virgínia Talaveira Valentini. Aplicação do método Delphi para identificação e avaliação dos fatores restritivos à realização de Parcerias Público-Privadas (PPP). [São Paulo]: *Revista Espacios*, v. 38, n. 22, 2017. ISSN 0798 1015. Disponível em: <https://www.revistaespacios.com/a17v38n22/a17v38n21p29.pdf>. Acesso em: 05 jan. 2022.
- ENGEMAN. O que é um algoritmo de decisão? Veja exemplos. [s. l.: s. n.], 2022. Disponível em: <https://blog.engeman.com.br/o-que-e-um-algoritmo-de-decisao-veja-exemplos/#:~:text=O%20algoritmo%20de%20decis%C3%A3o%2C%20tamb%C3%A9m,da%20categoria%20de%20aprendizagem%20supervisionada>. Acesso em: 09 mar. 2022.
- FOSTER PROVOST & TOM FAWCETT. *Data Science para Negócios*. Editora Alta Books. Rio de Janeiro/Brasil. Primeira edição, 2016.
- GUIMARÃES JUNIOR, D. S.; SILVA, R. J. A.; SANT'ANNA, C. H. M.; MARTINS JÚNIOR, J. E. V.; MELO, F. J. C. A percepção da inovação tecnológica em serviços e a fidelização de

clientes: uma análise do setor de serviços bancários. GESTÃO.Org - Revista Eletrônica de Gestão Organizacional, v. 19, n. 2, p. 232-249, 2021.

HAIR, JR. J.F., BABIN, B., MONEY, A.H., SAMOUEL, P. Fundamentos de métodos de pesquisa em administração. Porto Alegre: editora Bookman; 2005.

HUPDATA, Equipe. Por que usar a plataforma KNIME? [s. l.: s. n.], 2022. Disponível em: <https://hupdata.com/por-que-usar-a-plataforma-knime/>. Acesso em: 19 jul. 2022.

LOPES, Ahirton. A relevância dos dados e algoritmos para as pessoas e negócios. [s. l.: s. n.], 2021. Disponível em: <https://tiinside.com.br/14/03/2021/a-relevancia-dos-dados-e-algoritmos-para-as-pessoas-e-negocios/>. Acesso em: 09 mar. 2022.

COSTA, V. M. H. M. A utilização de tecnologia da informação por micro e pequenas empresas: estudo do setor de cerâmica artística e de decoração de município do Estado de São Paulo. Revista de Administração, Sociedade e Inovação, v. 7, n. 3, p. 104-122, 2021.

MACHADO, Isabela Gouvêa; NOGUEIRA, Sérgio Antônio de Paula. O uso dos sistemas de informação gerencial na tomada de decisão empresarial. [s. l.: s. n.], 2018. Disponível em: <https://dspace.doctum.edu.br/bitstream/123456789/2025/1/O%20USO%20DOS%20SISTEMAS%20DE%20INFORMA%C3%87%C3%83O%20GERENCIAL%20NA%20TOMADA%20DECIS%C3%83O.pdf>. Acesso em: 08 mar. 2022.

MENDONÇA, D. A.; TERRA, L. A. A. Vantagens da implementação do crm na indústria: uma análise multicase. Revista Inteligência Competitiva, v. 7, n. 2, p. 107-132, 2017.

MEDEIROS, M. M.; MAÇADA, A. C. G.; FREITAS JUNIOR, J. C. S. Estratégia e ciência de dados relacionadas à vantagem competitiva – um ensaio teórico. Future Studies Research Journal: Trends and Strategies, v. 13, n. 3, p. 325-355, 20.

MOURA, A. L. de; SANTOS, D. F. L. S.; CONCEIÇÃO, E. V. Proposal for a financial management model applied to a small business in the fertilizer segment. Revista de Empreendedorismo e Gestão de Pequenas Empresas, v. 8, n. 3, p. 36+, 2019.

OLIVEIRA, Gabriela Silva de; PEREIRA, Silvio do Lago. **Implementação de classificação bayesiana.** São Paulo: FATEC, 2017. Disponível em: [http://bt.fatecsp.br/system/articles/1036/original/BT%20-%20\(Oliveira%20&%20Pereira,%202017\)%20-%20Implementa%C3%A7%C3%A3o%20de%20Classifica%C3%A7%C3%A3o%20Bayesiana.pdf](http://bt.fatecsp.br/system/articles/1036/original/BT%20-%20(Oliveira%20&%20Pereira,%202017)%20-%20Implementa%C3%A7%C3%A3o%20de%20Classifica%C3%A7%C3%A3o%20Bayesiana.pdf). Acesso em: 19 jul. 2022.

PALHARES JÚNIOR, D.; SILVA, W. A. C.; SANTOS, A. O.; AMARAL, H. F. Ações de Sustentabilidade Financeira em Empresas da Microrregião de Pará de Minas e Cidades Circunvizinhas. Revista de Empreendedorismo e Gestão de Pequenas Empresas, v. 8, n. 3, p. 0-0, 2019.

PIERSON, L. Data Science para Leigos. Editora Alta Books. Rio de Janeiro. 2019.

PIERRO, Bruno de. O mundo mediado por algoritmos. São Paulo: FAPESP, 2018. Disponível em: <https://revistapesquisa.fapesp.br/o-mundo-mediado-por-algoritmos/>. Acesso em: 09 mar. 2022.

PRADO, Tatiana. O que é a tomada de decisão e como ela pode ser estratégica para uma organização?. [s. l.]: Voitto, 2020. Disponível em: <https://www.voitto.com.br/blog/artigo/tomada-de-decisao>. Acesso em: 08 mar. 2022.

RIBEIRO, Evandro Marcos Saidel. Estatística Aplicada à Administração com o software KNIME. São Paulo: Universidade de São Paulo, 2016. Disponível em: https://edisciplinas.usp.br/pluginfile.php/3381126/mod_resource/content/1/ApostilaEstatistica_Knime.pdf. Acesso em: 19 jul. 2022.

RODRIGUES, Viviane. O que é a tomada de decisão nas organizações?. [s. l.]: Siteware, 2021. Disponível em: <https://www.siteware.com.br/lideranca/tomada-decisao-nas-organizacoes/#:~:text=A%20tomada%20de%20decis%C3%A3o%20pode,solu%C3%A7%C3%A3o%20para%20um%20dado%20problema>. Acesso em: 08 mar. 2022.

SACRAMENTO, Gabriel. Naive bayes: como funciona esse algoritmo de classificação. [s. l.: s. n.], 2021. Disponível em: <https://blog.somostera.com/data-science/naive-bayes#:~:text=O%20classificador%20Naive%20Bayes%20%C3%A9,features%20s%C3%A3o%20independentes%20entre%20si..> Acesso em: 19 jul. 2022.

SEBRAE. Painel de empresas. SEBRAE, 2020. Disponível em: <https://datasebrae.com.br/totaldeempresas/>. Acesso em: 10 jan. 2022.

SHIMOYAMA, Claudio. A comunicação é a alma da marca. [s. l.: s. n.], 2017. Disponível em: <https://www.mundocoop.com.br/entrevista/comunicacao-e-alma-da-marca-segundo-claudio-shimoyama.html>. Acesso em: 27 jul. 2022.

SILVA, Douglas da. Comportamento do cliente: como analisar e melhorar as experiências dos consumidores? [s. l.: s. n.], 2020. Disponível em: <https://www.zendesk.com.br/blog/comportamento-cliente/>. Acesso em: 09 nov. 2021.

SOLOMON, Michael. R. O comportamento do consumidor comprando, possuindo e sendo. Porto Alegre: Bookman, 2016. ISBN 9788582603680. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788582603680/>. Acesso em: 31 jan. 2022.

STAIR, Ralph M.; REYNOLDS, George W. Princípios de Sistemas de Informação. Tradução Noveritis do Brasil. 11. ed. São Paulo: Cengage Learning Brasil, 2016. 9788522124107. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788522124107/>. Acesso em: 26 jun. 2022.