

PROJETO DE MACHINE LEARNING: COMPREENSÃO DA HOSPITALIDADE COMO DIFERENCIAL COMPETITIVO NA GESTÃO DE RESTAURANTES

*MACHINE LEARNING PROJECT:
UNDERSTANDING HOSPITALITY AS A COMPETITIVE DIFFERENTIAL IN RESTAURANT
MANAGEMENT*

 **Paulo Sergio Gonçalves de Oliveira**
Doutor em Engenharia de Produção
Universidade Anhembi Morumbi – UAM
São Paulo, São Paulo – Brasil
psgoliveira@hotmail.com

 **Thaís Goldbard Youshiura**
Mestranda em Hospitalidade
Universidade Anhembi Morumbi – UAM
São Paulo, São Paulo – Brasil
thais_gy@hotmail.com

 **Carlos Alberto Alves**
Doutor em Administração
Universidade Anhembi Morumbi – UAM
São Paulo, São Paulo – Brasil
calves761@gmail.com

Resumo

O objetivo deste artigo é apresentar o desenvolvimento de um projeto de Machine Learning para prever a classificação do cliente em relação ao restaurante, possibilitando dessa forma a utilização da Hospitalidade como um diferencial competitivo. Para atingir o objetivo foi desenvolvido um projeto de Machine Learning, o qual envolveu o desenvolvimento de um script na linguagem R, que possibilita a análise e aplicação em Restaurantes, de forma a apoiar os gestores na tomada de decisões e eventuais ações de mitigação dos problemas. De forma a capturar a experiência dos especialistas, um modelo foi desenvolvido por meio da aplicação do algoritmo Naïve Bayes, o qual foi treinado utilizando dados obtidos do Site TripAdvisor, atingindo uma taxa de acerto de cerca de 84% com os dados de teste. Esse valor é aceitável para novas análises com dados oriundos das opiniões dos clientes, demonstrando dessa forma que o projeto atingiu o seu objetivo.

Palavras-chave: Projeto de Software. Naïve Bayes. Machine Learning. Hospitalidade na Competitividade de Serviços. Gestão em Alimentos e Bebidas.

Abstract

The aim of this article is to present the development of a Machine Learning project to predict the classification of the customer in relation to the restaurant, thus enabling the use of Hospitality as a competitive differential. To achieve the objective, a Machine Learning project was developed, which involved the development of a script in the R language, which allows analysis and application in Restaurants, in order to support managers in decision-making and eventual actions to mitigate problems. In order to capture the experts' experience, a model was developed by applying the Naïve Bayes algorithm, which was trained using data obtained from the TripAdvisor Site, reaching a hit rate of around 84% with the test data. This value is acceptable for new analyzes with data from customer opinions, thus demonstrating that the project has achieved its objective.

Keyword: Software Design. Naïve Bayes. Machine Learning. Hospitality in Service Competitiveness. Food and Beverage Management.

Cite como - American Psychological Association (APA)

Oliveira, P. S. G., Youshiura, T. G., & Alves, C. A. (2020, set./dez.). Projeto de machine learning: compreensão da hospitalidade como diferencial competitivo na gestão de restaurantes. *Revista de Gestao e Projetos (GeP)*, 11(3), 26-45.
<https://doi.org/10.5585/gep.v11i3.18748>.

1 Introdução

Atualmente o mercado de alimentos e bebidas vem passando por profundas mudanças nos últimos 30 anos, sendo a principal delas o que se observa na cidade de São Paulo, onde os restaurantes que anteriormente não tinham grandes diferenciações entre si, sendo os menus geralmente *à la carte* apresentando poucas variedades com receitas estrangeiras (francesa, portuguesa e japonesa) ou ainda com uma culinária regional (mineira, baiana, comida dos pampas), passando para uma extensa variedade de opções e tipologias, tornando dessa forma o mercado bastante complexo e variado, no que se refere as opções de escolha do cliente, gerando diversos nichos de mercado, tornando o desafio gerencial ainda maior para os proprietários e gestores desses estabelecimentos, pois devem estar atentos a uma série de particularidades relativas aos seus clientes, em termos de variação dos cardápios, análise do ambiente do restaurante, atendimento, qualidade dos produtos além de outros atributos, sempre visando a satisfação do cliente.

Nesse contexto, a gestão da Hospitalidade por meio das opiniões dos clientes em Sites de avaliação tem se apresentado uma importante ferramenta para os gestores, já que possibilita o monitoramento da opinião dos clientes de forma quase que instantânea, favorecendo a rápida ação dos gestores, por meio de decisões mais assertivas. Sendo um diferencial competitivo em cidades

como é caso de São Paulo, que é focada em Turismo de Negócios, e que em muitos dos casos serve como ferramenta de tomada de decisão para os clientes, já que muitos deles estão visitando a cidade pela primeira vez e dessa forma não possuem conhecimento dos estabelecimentos, focando nessa forma em Hospitalidade no conceito comercial.

Segundo Tideman (1983, p. 1), as atividades interligadas à hospitalidade no contexto comercial, traz o enfoque, “As necessidades do hóspede são satisfeitas ao máximo, significando a oferta de bens e serviços em quantidade e qualidade desejada pelo hóspede e a um preço”. Devido aos motivos apresentados se pensou no projeto do sistema de análise e previsão das opiniões dos clientes, tendo por base a utilização de Sites Online de avaliações.

Sendo assim, se pensarmos nas inúmeras atividades de lazer, as que estão ligadas às áreas da gastronomia, como bares, restaurantes temáticos, casa de shows, ofertam a seus frequentadores muito mais do que a comida e a bebida, podem dar a eles, a oportunidade e momentos de descontração. Estes lugares, estão mais suscetíveis em criar uma interação entre o cliente e o ambiente, ou entre o colaborador e o cliente, proporcionando muitas vezes momentos de prazer a quem o frequenta, suprimindo de alguma forma suas necessidades e desejos.

Pensando nisso, surgem novas demandas de necessidades e tantas outras ainda

surgirão, e claro, junto a isso surgem uma infinidade de serviços e produtos para supri-las (king, 1995). Bares e restaurantes, almejam ter a preferência e ter a atenção do consumidor, estabelecendo mais que um simples contato de compra e entrega, estas empresas buscam criar uma conectividade com seu cliente, querem gerar contentamento e fidelização de longo prazo, querem ser lembradas em primeiro lugar por estes clientes.

Dentre os fatores que geram essa conectividade e fidelização, destaca-se a hospitalidade, que para Jones (2002, p.1), divide-se em dois casos, o primeiro, a provisão de acomodação para as pessoas que estão longe de suas casas, os viajantes; e o segundo, dar comida e bebida para as pessoas que comem fora de casa. Estas suas diferenciações surgem num momento pelo qual os hoteleiros e donos de tabernas, restaurantes gostariam que seu setor fosse percebido.

Investigar sobre a hospitalidade como diferencial competitivo na gestão de restaurantes japoneses, justifica-se pelo fato de ser um assunto que poderá interferir direta ou indiretamente no sucesso ou fracasso, especificamente na forma de gerir um empreendimento como restaurantes de comida japonesa na cidade de São Paulo. Restaurantes, empreendedores e gestores no setor de alimentos e bebidas, vivenciam essa realidade no setor, com base em experiências profissionais do dia a dia da operação desse estabelecimento, com base nas exigências dos consumidores por um serviço diferenciado,

necessidade do estabelecimento ser competitivo frente a outros concorrentes, a importância de conhecer a fundo a gestão de restaurantes e ter a hospitalidade como agregadora de valor para esse negócio.

Desta forma, o presente artigo partiu da necessidade de entender a hospitalidade no domínio comercial, conceituar competitividade na área de alimentos e bebidas, estabelecendo uma conexão com a hospitalidade, diferencial competitivo e gestão de negócios em restaurantes japoneses.

O presente artigo estabeleceu como problema de pesquisa: Quais as influências da hospitalidade como um diferencial competitivo na gestão de restaurantes japoneses? E como objetivo geral, apresentar o desenvolvimento de um projeto de Machine Learning para prever a classificação do cliente em relação ao restaurante, possibilitando dessa forma a utilização da Hospitalidade como um diferencial competitivo. Durante o desenvolvimento foram aplicadas técnicas de gestão de projetos a fim de aumentar o sucesso do projeto, já que historicamente nessa área em geral existe um baixo índice de sucesso (Prado et al., 2020). Sendo essa a razão para se apostar nos fatores críticos de sucesso do projeto (Link et al., 2020; Russo & Silva, 2019).

Para atingir o objetivo da pesquisa na segunda seção foi efetuada uma revisão teórica do tema Hospitalidade, os procedimentos metodológicos foram apresentados na seção 3, e análise e discussão dos resultados foi efetuada por meio da seção 4.

2 Fundamentação teórica

Nesse capítulo iremos apresentar os principais conceitos referentes a hospitalidade por meio do tópico 2.1, além da Gestão Administrativa de restaurantes por do tópico 2.2, de forma fornecer para o leitor um panorâma dos desafios enfrentados durante o desenvolvimento do projeto do software para auxiliar os gestores nas tomadas de decisão.

2.1 Hospitalidade

Antes do início da apresentação dos referenciais teóricos associados ao tema e conceitos do presente trabalho, é importante destacar que o estudo da hospitalidade inserida no contexto comercial, bem como o aprofundamento da matéria competitividade e ao conhecimento da gestão de restaurantes japoneses da cidade de São Paulo são ferramentas necessárias para o bom funcionamento de negócios de alimentos e bebidas, estendendo-se a áreas que impactam nos processos de decisão de uma empresa.

Sendo assim, serão apresentados conceitos de diferentes autores sobre a hospitalidade como diferencial competitivo na gestão de restaurantes japoneses.

Na abordagem de Lashley e Morrison (2010), a hospitalidade é um ato de estreitar as relações humanas, uma espécie de ritual com relação ao acolhimento e muitas vezes um símbolo de civilidade, que mostra o papel do anfitrião, aquele que acolhe, bem como o do hóspede, aquele que é recebido. Esses ritos

acontecem em diferentes aspectos, na perspectiva do autor, existem domínios diferentes em relação à hospitalidade, sendo estes, no âmbito social, privado e comercial.

Se tratando do âmbito comercial, da qual será o foco abordado ao longo do trabalho, a hospitalidade acontece na sociedade com uma atividade econômica atrelada, ocorre sempre uma troca monetária, uma movimentação financeira, nas palavras de Lashley (2004, p. 20), tem uma redução da obrigatoriedade de deveres da hospitalidade entre o anfitrião e o hóspede.

Para Pitt-Rivers (2017, p. 8) no conceito da hospitalidade social, as questões culturais, no modo de recepcionar, alimentar o outro, da acolhida, são tratadas desde as mais antigas civilizações, estas que se preocupavam com o código de conduta dessa hospitalidade que estaria enraizada culturalmente nesta sociedade. O comportamento diferenciado e solícito desse anfitrião para com o estrangeiro, de provisionar comida, bebida e lugar para dormir, trazia a convicção ao anfitrião de que este seria bem recebido quando estivesse viajando.

Segundo Godbout (2019, p. 40), no domínio privado, envolve a oferta e partilha da comida, bebida e moradia (acolhimento), esses laços estão ligados a atos genuínos, particulares, íntimos, assim como sugere Godbout (2019), que na condição de dádiva, o conceito de dar, receber e retribuir ao próximo seja presente no ato de fazer sem esperar nada em troca.

Segundo Jones (2002), exemplifica em uma perspectiva comercial, administrativa e econômica, em que o termo hospitalidade tenha dois lados diferentes, o primeiro o fornecimento de alimentação para as pessoas que estão fora de casa e o segundo à aqueles que necessitam de acomodação para passar a noite, pois estão longe de suas casas. Essas questões foram levantadas como o modo pelo qual os fornecedores de alimentos e bebidas e dos hoteleiros gostaria que sua classe fosse vista pelo setor.

Há também conceitos desenvolvidos por Tideman (1983) em que os desejos e necessidades dos hóspedes são supridos e suas expectativas superadas, significando que produtos e serviços são ofertados a um preço que os hóspedes paguem e sintam que estes tiveram valor agregado que valeu o preço pago, vale lembrar que a qualidade percebida e entregue nesses quesitos acima incluem-se nesse conceito abordado.

A hospitalidade uma vez estabelecida em locais comerciais, pode ser transformada em uma “indústria” rentável em diversos segmentos de negócios, desde que haja a permissão de fornecimento de comida, bebida e acomodação. Ainda segundo Lashley (2004), para a hospitalidade comercial acontecer, existe a necessidade de um hóspede (cliente) e o anfitrião (colaborador), envolve ainda uma junção de elementos tangíveis - bebida, comida, produto, e elementos intangíveis, os que não são difíceis de mensurar, pois são palpáveis e depende da percepção de cada hóspede

(cliente).

Levando em consideração os conceitos já apresentados sobre o assunto, pode-se também entender que o ato de receber e servir está diretamente ligado a qualidade dos serviços e produtos ofertados aos clientes. É o mesmo questionamento que leva em consideração se o restaurante está preparado e a altura para receber e servir seus comensais. Para Lashley e Morrison (2014) o chamado hospitaleiro é determinado pela capacidade e qualidade da receptividade, da acolhida ao outro, diga-se o estranho, incluindo consequentemente a comida e bebida.

Posto isto, os desejos, necessidades e vontades, seja de hóspedes ou clientes estão ligados ao fato se quiserem ser bem recebidos. O termo hospitalidade se faz presente em diversos momentos da ida a um restaurante, por exemplo, seja no momento da escolha do prato, no site de reserva ou até mesmo na despedida do local. Além disso, seja uma pesquisa de satisfação, um agrado ao cliente no final da refeição ou até mesmo um cartão de fidelização demonstram que o estabelecimento quer criar um laço de confiança e de boas memórias para os comensais, uma experiência vivenciada nesse local pode interferir na visão do cliente, deixando claro para este se o local é ou não hospitaleiro. tornando um momento de hospitalidade.

Walker (2002, p. 26) reforça ainda que a qualidade de receber bem, é percebida como um diferencial competitivo para os prestadores de serviços, seja bares, hotéis, restaurantes, por

isso empreendimentos que se enquadram nessa categoria devem estar atentos aos desejos e necessidade de seus clientes.

2.2 *Gestão administrativa de restaurantes*

O ato da hospitalidade engloba não só o território em que situa, abrange locais físicos, como se houvesse uma materialização, transformando os negócios como restaurantes e bares nessa materialidade, envolvendo também outros quesitos imateriais, como a gestão, planejamento financeiro, contratação dos recursos humanos, entre outras atividades.

Para Walker (2004), de forma unânime os empreendimentos comerciais têm como objetivo principal alcançar a lucratividade, compensando os esforços e o capital investido. Esse lucro surge através do controle e desempenho satisfatório do administrativo-operacional. Drucker (2000) afirma que “o único propósito de um negócio é criar um cliente”, ao que Walker (2004), complementam: “o único propósito de um negócio é fazer com que o cliente retorne”.

O ato de comer fora de casa faz parte da vida contemporânea. Nas cidades como São Paulo, é quase unânime almoçar ou jantar fora de casa, seja pela correria do dia a dia, a falta de um local apropriado dentro da empresa, seja a distância entre a residência e o trabalho, que entre outras questões diminuem o fato do trabalhador poder almoçar em casa, por exemplo. Em cidades do interior, o fator comer fora de casa acaba sendo uma questão de sociabilidade e não uma imposição pela

distância de casa-trabalho, como dito anteriormente. Essas e tantas outras questões; o sonho de abrir um restaurante ou bar, trabalhar com comida, investir dinheiro, faz com que pessoas projetem na gastronomia um setor rentável e uma oportunidade de negócio, deixando de levar em consideração todo o dinamismo e complexidade que permeia a gestão de um estabelecimento de alimentos e bebidas.

Para o SEBRAE SP (2019), a área de alimentos e bebidas abrange diferentes setores, dentre eles, comércio atacadista e varejista, bares, restaurantes, a produção agrícola, de pesca e pecuária, fábricas de alimentos, o comércio ambulante, incluindo por consequência e a atividade que será abordada no decorrer do presente trabalho, a gestão de restaurantes. Ainda dados do SEBRAE BR (2019), o setor de alimentação fora do lar cresce anualmente 10% frente a outros negócios, gerando 450 mil novos empregos formais, de acordo com a Associação de Bares e Restaurantes (ABRASEL, 2019).

Segundo dados da Prefeitura de São Paulo (2019), a quantidade de estabelecimentos formais dos serviços de alimentação cresceu 68% nos últimos 10 anos na capital paulista, o que equivale a 23.092 estabelecimentos, 7,7% do total de comércios da cidade. O setor tem faturamento anual médio de R\$ 31,9 bilhões por ano. Desses estabelecimentos, em média de 700 a 1.500 são restaurantes japoneses no Brasil, já esse número na Grande São Paulo é de 3 mil

estabelecimentos de comida oriental, considerando temakerias, restaurantes no sistema de rodízio, buffets e segmentos similares.

Para a Associação Brasileira de Bares e Restaurantes de São Paulo (ABRASEL, 2019), o gasto médio por refeição pelos clientes nesses locais gira em torno de R\$80,00, o que projeta um faturamento total de R\$36 milhões por dia, R\$13,4 bilhões anualmente, somente para a Grande São Paulo. Já para o Brasil inteiro os restaurantes dessa categoria movimentam algo de R\$19 bilhões anualmente.

De acordo com dados do SEBRAE SP (2018), os setores de alimentação, englobando vários tipos de bares e restaurantes japoneses, movimentam uma vasta cadeia produtiva, multiplicando uma diversidade de negócios varejistas, assim como prestadores de serviços, empresas de manutenção de equipamentos, feiras específicas do setor, entre outros tantos. Contudo, a falta de planejamento adequado e gestão estruturada, faz com que o negócio vá a falência com poucos meses de operação, em um prazo de 24 meses, 35% dos estabelecimentos encerram suas operações (ABRASEL, 2019).

O segmento de alimentos e bebidas, especificamente em restaurantes japoneses, deu-se em meados de 1995, com a facilidade de acesso aos produtos importados, imprescindíveis nesse tipo de culinária, ora considerada exótica (Santos, Mendes & Corte-Real, 2014). Teve um começo fragmentado, com incertezas pela falta de conhecimento da cultura alimentar japonesa, e hoje um vasto

conglomerado de redes de *fast food*, pequenas mercearias, restaurantes de alto padrão, conectando fornecedores de produtos locais, cadeias de suprimentos entre outros, se unem para a conexão de empresas com melhores práticas de gestão (Fonseca, 2018).

Santos, Mendes & Corte-Real (2014) afirma que as cidades como um todo estão se organizando para melhorar a gestão de negócios de alimentação, visto que é um mercado altamente competitivo. Se preparar com embasamento em indicadores de vendas, estudo do mercado e entender o comportamento do consumidor faz com que os estabelecimentos de alimentos e bebidas avance para a gestão eficaz do negócio.

Segundo Fonseca (2018 p. 30), o serviço é algo difícil de ser mensurado, é como se tivesse data de validade, pois muda conforme o mercado e a percepção do cliente, sua produção e o momento de consumo acontecem quase que simultaneamente. Quando fala da dificuldade em mensurar o serviço é porque o processo de avaliação do cliente para com o serviço ainda não ocorreu, somente depois da aquisição desde é que o cliente terá suas primeiras impressões, vezes positivas ou não. Outro ponto, o serviço é perecível, pois é impossível de ser armazenado, a intangibilidade dificulta o processo da avaliação o serviço pelo cliente (Mullins, 2001). A simultaneidade neste quesito acaba tornando o cliente partícipe do processo de produção/acontecimento do serviço, fazendo com que sua ida aos estabelecimentos comerciais, como restaurantes

e bares seja imprescindível para que o serviço aconteça. Todos os aspectos descritos acima, a gestão total e a validação do serviço, são pontos de difícil mensuração do próprio empreendimento juntamente por ser intangível. (Fonseca, 2018).

3 Procedimentos metodológicos

De forma a se buscar o atingimento dos objetivos da pesquisa, foram feitos levantamentos por meio da revisão da literatura de forma a apoiar o desenvolvimento do projeto do software por meio da aplicação da técnica de Machine Learning, procurando dessa forma desenvolver uma base conceitual que pudesse suportar o desenvolvimento da ferramenta.

Com esse foco a presente pesquisa pretende seguir os argumentos de Kerlinger (1988) onde explica que a aplicação do método científico se deve a necessidade de obtenção de conhecimento, além de sua compreensão de modo mais confiável, por meio de um método que provê controle, gerando conhecimento com base na análise e conhecimento dos fenômenos naturais.

O conhecimento científico segundo Marconi e Lakatos (2017) é mais confiável devido a sua característica principal de lidar com ocorrências e fatos por meio de proposições ou hipóteses que podem ser testadas e replicadas, sendo passíveis de confirmação ou não, gerando dessa forma um saber ordenado, o qual pode ser verificado, não se propondo a ser infalível, já que a ciência

reconhece sua mutabilidade, ou seja, é verdadeiro até que se prove o contrário.

Nesse sentido também vale o argumento de Kerlinger (1988) ao afirmar que a pesquisa científica obtém relevância ao buscar a objetividade, pois ela perde o caráter único e especial, reduzindo a complexidade por meio do caráter empírico, devido a sua facilidade de compreensão, pois tem como característica a experimentação sistemática no sentido de buscar o entendimento do fenômeno.

No caso da presente pesquisa o fenômeno estudado foi o desenvolvimento de um modelo de classificação de avaliações de restaurantes japoneses na cidade de São Paulo, de forma a permitir que os proprietários desses estabelecimentos consigam prever de que maneira serão avaliados por seus clientes na plataforma TripAdvisor.

A pesquisa irá buscar descrever de que maneira uma parcela da população avalia os restaurantes japoneses presentes na Cidade de São Paulo, por meio da participação das variáveis na avaliação dos clientes, pois segundo Gil (2008) um problema pode ser formulado para verificar se determinada ação pode gerar consequências variadas, permitindo a previsão das ações, por meio da análise dos padrões.

Os dados foram coletados a partir dos dados do TripAdvisor por meio dos valores contidos nas páginas, os quais segundo Skiena (2017) possuem dados valiosos para pesquisa tanto no formato texto, quanto numéricos.

Em relação aos métodos para a obtenção de dados em páginas, Skiena aponta que existem dois métodos “Spidering” e “Scraping”, a saber:

- Spidering: é o processo de fazer o download das páginas certas para análise;
- Scraping: é a fina arte de retirar o conteúdo de cada página de forma a prepara-las para análises computacionais.

O presente estudo efetuou a coleta por meio do segundo método, sendo que a coleta de dados foi efetuada por meio da obtenção da avaliação de restaurantes Japoneses da Cidade São Paulo, os quais foram coletados por meio de uma pesquisa do tipo levantamento, diretamente no site TripAdvisor, caracterizando dessa forma os dados como secundários, já que são oriundos de uma base dados de avaliações existentes e públicas (Babbie, 1999).

O Universo da pesquisa foi composta por 605 restaurantes da Cidade de São Paulo coletados de forma aleatória na Plataforma TripAdvisor, tendo os restaurantes sido classificados com notas de 1 a 5, segundo metodologia do próprio site, posteriormente esses dados foram classificados em duas categorias: “Bem Avaliado” e “Mal Avaliado” (Babbie, 1999).

Para a análise dos dados foi aplicado o conceito de aprendizagem de máquina, o qual tem sido uma assunto de grande interesse em diversos campos, principalmente nos ramos de

ciências da computação e informações, tendo como foco principal a busca por um algoritmo que possibilite a automatização e a melhoria de tarefas gerenciais, tornando-as mais efetivas e acuradas (Kou et al., 2019).

O crescente interesse em aprendizado de máquina vem aumentando desde os anos de 1983, quando houve um sensível aumento no número de artigos publicados com as mais diversas proposições, que visavam responder ao grande mistério que circundava essa área, fazendo com que surgissem alguns livros dedicados ao tema e também um jornal dedicado ao tema, além de diversas conferências (Kubat, 2017).

Nos últimos anos esse interesse só vem crescendo, motivado principalmente pelo aumento do uso da Internet e devido ao surgimento das redes sociais, o que aumentou sensivelmente o número de informações disponíveis, tornando-se um desafio para os analistas de dados, pesquisadores e empresas que precisavam analisar e tomar decisões com base nesses dados.

Segundo Sharma, Kumar e Chand (2019) o conceito de aprendizagem de máquina se tornou um campo de interesse, tanto em aplicações empresariais, quanto científicas, devido a sua capacidade da inteligência artificial de criar máquinas que imitem o intelecto humano, gerando dessa forma um sem número de possibilidades em termos de aplicações tais como: automóveis autônomos, a escrita e publicação de relatórios diversos, a comunicação com seres humanos e a prevenção

a atentados terroristas, apenas para citar algumas possibilidades.

O aprendizado de máquina se dá por meio do aprendizado a partir da experiência, em relação a uma classe tarefas, a qual é melhorada de com a experiência, sendo que no geral ela se dá pelo entendimento da tarefa, seleção do algoritmo e criação do modelo (Kou et al., 2019).

Por essa razão Jordan e Mitchel (2015) chamam a atenção de que o aprendizado de máquina é uma disciplina focada em duas questões interconectadas: De que maneira construir um computador que automaticamente melhora por meio da experiência? Quais são as leis estatísticas, computacionais e teóricas que governam todos os sistemas de aprendizagem, incluindo computadores e seres humanos? O que segundo esses autores faz com o estudo da aprendizagem de máquina se tornou interessante para responder as questões fundamentais da ciência e da engenharia, sendo de grande aplicabilidade prática por meio de muitas aplicações.

Esse dinamismo se demonstra por meio de sua aplicação em diversos campos do conhecimento humana, tendo encontrando um campo fértil na área de saúde, tais como por exemplo a pesquisa de Villani et al. (2018) que propuseram um sistema categorizar pesquisas relativas a prevenção de doenças financiadas pelo NIH (*National Institutes of Health*) por meio da proposição de uma categorização desses esforços.

A maior parte das técnicas de Machine Learning são baseadas em algoritmos que utilizam técnicas de inteligência artificial que se baseiam em um conjunto de exemplos para treinamento e testes, para verificar o quanto aprenderam. Os métodos de aprendizagem são divididos em métodos supervisionados os quais possuem uma classe na qual o algoritmo deve se basear e não supervisionados o quais não possuem classe para que o algoritmo se baseie (Kubat, 2017).

Segundo Skiena (2017) os problemas de aprendizagem de máquina recaem em duas áreas, a saber: classificações e regressões. Nesse trabalho foi aplicada uma técnica de classificação, por meio da qual é aplicada uma etiqueta para um item de um conjunto discreto de possibilidades, o que permite decidir qual é a classificação de um determinado filme segundo o seu gênero ou mesmo prever o vencedor de um determinado tipo de esporte (Skiena, 2017).

De forma para se fazer a análise dos dados se adotou o procedimento de utilizar um conjunto de dados já pré-classificado com exemplos pertencentes a duas partes, ou seja, estabelecimentos “bem avaliados” e estabelecimentos “mal avaliados”, obtidos a partir do site TripAdvisor. A partir disso conjunto de dados foi dividido em duas partes sendo 70% utilizados para o treinamento e 30% utilizados para teste, tendo sido essas amostras divididas de forma aleatória por meio do cálculo de probabilidade com remoção dos participantes da amostra restante, buscando se

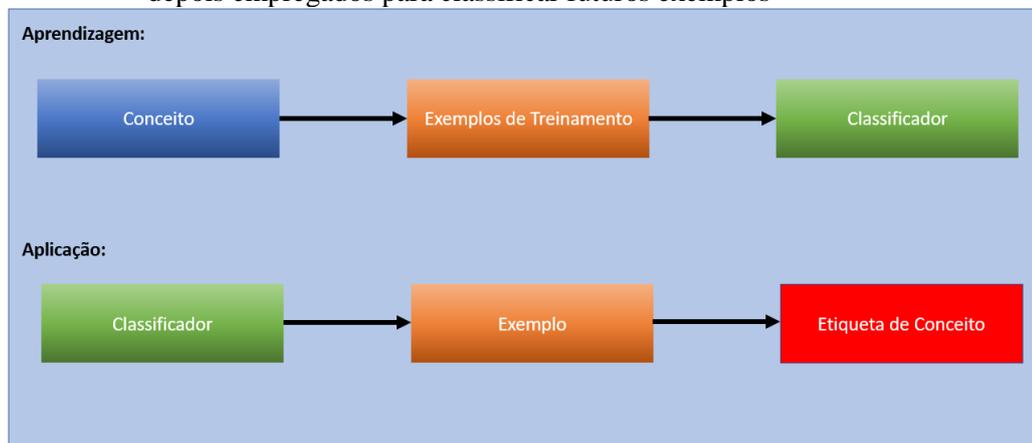
obter uma performance de acertos superior a 75% (Kubat, 2017).

Esses procedimentos visam aferir a acuracidade do modelo criado a partir dos dados de treinamento, permitindo se avaliar o quão assertivo ele é, por meio do aprendizado por meio da experiência (Jordan & Mitchell, 2015). Dessa forma, a partir do modelo treinado, pode-se verificar a taxa de acerto

quando o mesmo conjunto de dados é submetido ao modelo em termos de acerto da previsão de que classificação esse novo conjunto de dados irá receber.

A Figura 1 demonstra o processo de aprendizagem ao qual modelo é submetido e o processo de reconhecimento da categoria que um novo conjunto de dados irá receber ao ser apresentado para o modelo existente.

Figura 1 – Demonstração de como os exemplos são utilizados para induzir a classificação e depois empregados para classificar futuros exemplos



Fonte: Adaptado de Kubat (Kubat, 2017).

Tendo por base o modelo demonstrado na Figura 1, os dados obtidos foram tratados para se obter o modelo de aprendizagem proposto no estudo, o qual utilizou dados de 605 classificações de restaurantes japoneses obtidas no site TripAdvisor. Para análise dos dados foi utilizado o método Naïve Bayes o qual é baseado na teoria probabilística bayesiana, que calcula a probabilidade de pertencimento a uma determinada classe de

acordo com o objeto fornecido, gerando um valor de probabilidade para isso e então escolhendo a classe que recebeu o maior valor (Kubat, 2017).

O método Naïve Bayes é baseado no teorema de Bayes, dado pela fórmula de probabilidade, expresso matematicamente pela Equação 1 (Asafu-Adjei & Betensky, 2015; Chen et al., 2016; Sánchez-Franco et al., 2019; Skiena, 2017; Wang, Jiang, & Li, 2015):

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

Sendo A e B evento e $P(B) \neq 0$

$P(A)$ e $P(B)$ são probabilidades a priori A e B

$P(A|B)$ probabilidade a posteriori de A condicional a B

$P(B|A)$ probabilidade a posteriori de B condicional a A

O método é chamado Naïve devido ao fato de existir a probabilidade das variáveis preditoras sendo correlacionadas, fato que não é levado em conta pelo algoritmo, sendo essa a razão de ser chamado Naïve (Inocente) (Li et al., 2014; López-Cruz, Bielza, & Larrañaga, 2015; Shiri Harzevili & Alizadeh, 2018; Vural & Gök, 2017).

Esse aspecto é demonstrado por Skiena (2017) onde demonstra por meio da apresentação de dois eventos independentes

$$p(A \text{ e } B) = p(A) * p(B | A) = p(A) + P(B) - p(A \text{ ou } B) \quad (2)$$

Utilizando o algoritmo Naïve Bayes para classificação utilizando o teorema de Bayes acontece da seguinte maneira, suponha que queiramos classificar o vetor $X = (x_1 \dots x_n)$ dentro de um universo m de classes ($C_1 \dots C_m$). Dessa maneira é assumido que se proverá a etiqueta X para a classe com a maior

$$p(C_i | X) = \frac{p(C_i) \cdot p(X | C_i)}{p(X)} \quad (3)$$

O termo $p(C_i)$ é a probabilidade a priori de uma etiqueta de classe sem qualquer evidência específica, por exemplo a probabilidade de possuir um cachorro preto em casa, já o denominador $P(X)$ fornece a probabilidade de se verificar um dado vetor X sobre todas as possibilidades de entrada,

A e B se $p(A \text{ e } B) = p(A) * p(B)$, citando como exemplo os seguintes fatos: a probabilidade de que meu time de futebol vença hoje e o mercado de ações, os quais são presumivelmente independentes, não sendo verdadeiro em geral, existindo dessa maneira dependência entre eles, fato que é demonstrado pela Equação 2.

probabilidade, o que é demonstrado por meio da Equação 3, tendo por base o teorema de Bayes apresentado na Equação 1 (Chen et al., 2018; Karandikar et al., 2015; Li et al., 2014; Shiri Harzevili & Alizadeh, 2018; Skiena, 2017):

permitindo que possamos estabelecer a uma etiqueta para X . Sendo a classe selecionada a com a maior média de possibilidade (Asafu-Adjei & Betensky, 2015; Skiena, 2017; Wang et al., 2015; Xu, 2018). Esse aspecto é demonstrado pela Equação 4.

$$C(X) = \underset{i=1, \dots, m}{\operatorname{arg\,max}} \frac{p(C_i) \cdot p(X|C_i)}{p(X)} = \operatorname{arg\,max}_{i=1, \dots, m} p(C_i) \cdot p(X|C_i) \quad (4)$$

O termo resultante $p(X|C_i)$ é a probabilidade de apresentado um vetor X , existir a possibilidade de se conhecer a classe C_i . No entanto, vamos partir do princípio que os eventos são independentes sendo a

probabilidade dos eventos A e B serem sempre $p(A) * p(B)$ (Chen et al., 2016; Jiang et al., 2013, 2013; Shiri Harzevili & Alizadeh, 2018; Skiena, 2017; Wang Jiang, & Li, 2015). Então, temos a Equação 5.

$$p(X|C_i) = \prod_{j=1}^n p(X_j|C_i) \quad (5)$$

Partindo da crença de que o mundo é formado por probabilidades independentes, sendo, portanto, de certa forma ingênuo (daí o nome do algoritmo) com o objetivo de tornar o processamento computacional mais fácil. Isso é

colocado junto e demonstrado por meio da Equação 6 (Asafu-Adjei & Betensky, 2015; Skiena, 2017; Vural & Gök, 2017; Wang, Jiang, & Li., 2015; Wen et al., 2019).

$$C(X) = \underset{i=1, \dots, m}{\operatorname{arg\,max}} p(C_i) \cdot p(X|C_i) = \underset{i=1, \dots, m}{\operatorname{arg\,max}} p(C_i) \prod_{j=1}^n p(x_j|C_i) \quad (6)$$

Por fim, após chegar ao produto, devemos aplicar o logaritmo para então aplicar uma somatória, com o objetivo de se chegar a

uma melhor estabilidade numérica, permitindo se chegar à fórmula final do algoritmo Naïve Bayes, apresentado por meio da Equação 7.

$$\tilde{C}(X) = \underset{i=1, \dots, m}{\operatorname{arg\,max}} (\log(p(C_i)) + \sum_{j=1}^n \log(p(x_j|C_i))) \quad (7)$$

Para calcular a probabilidade $p(x_j|C_i)$, a partir de uma observação x_j , para encontrar a etiqueta $C(X)$. Isso se torna um processo fácil por meio da utilização dos dados de treinamento, principalmente se x_i é uma variável categórica. Para se chegar ao resultado simplesmente se seleciona todas as instâncias da classe i nos dados de treinamento e se computa uma fração deles os quais tem a

propriedade x_i , sendo essa fração razoavelmente estimada de $p(x_i|C_i)$ (Li et al., 2014; López-Cruz, Bielza, & Larrañaga, 2015; Sánchez-Franco et al., 2019; Skiena, 2017; Vural & Gök, 2017; Wang, Jiang, & Li, 2015; Wen et al., 2019).

Para o tratamento do banco de dados foi utilizado o software R versão 3.6.3, por meio da package e 1071 (R Foundation, 2019).

4 Análise dos resultados

Como mostrado ao longo do trabalho, a coleta dos dados foi por meio do site TripAdvisor, onde as informações dos 605 restaurantes foram inicialmente carregadas, depois de feita as análises descritivas, de modo a se verificar os dados coletados. A partir do banco de dados carregado no R, o valor da mediana foi determinado em 4, permitindo a decisão de que os restaurantes “bem avaliados” deveriam receber uma nota $>$ que 4, enquanto os restaurantes “mal avaliados” deveriam receber uma nota ≤ 4 .

Dessa forma os dados foram ajustados na coluna chamada “média avaliações” para “bem avaliado” $\Rightarrow 301$ e “mal avaliado” $\Rightarrow 304$, ficando a amostra de 605 restaurantes divididos dessa forma.

O método de utilização entre as probabilidades dessas informações que se relacionam entre si é chamado Naïve. Posto isso, a primeira análise foi a determinação das notas gerais recebidas pelos restaurantes, existindo a probabilidade desses restaurantes de serem mais “bem avaliados” ou “mal avaliados” pelos usuários do site.

O fato do banco de dados poder apresentar informações não completas e com objetivo de melhorar a previsibilidade do modelo testado, se observou o comportamento dos dados segundo esse preceito, onde 16% desses dados não preenchidos, o que pode causar uma redução na precisão e previsão no contexto geral da pesquisa, ou seja fará a diferença final, o fato desses 16% não terem

sido respondidos. De forma a se eliminar os dados não preenchidos foi utilizada a função “convert”, quando essas funções convertem uma expressão de um tipo de dados em outro.

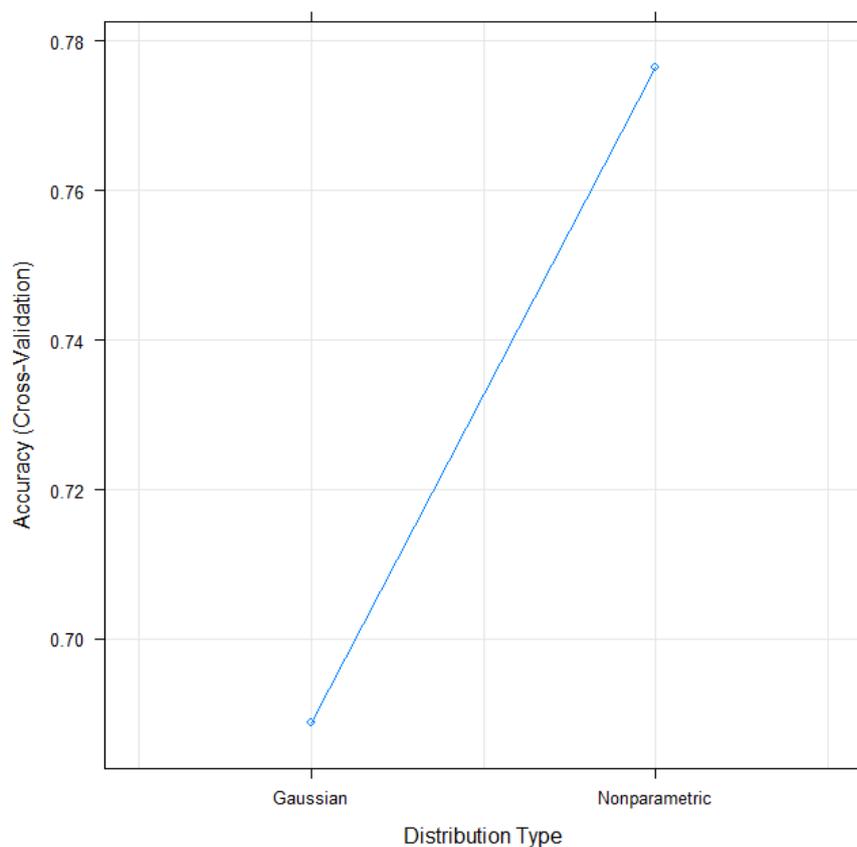
Logo após foi aplicada a função “mice” por meio do método “rf”, que utiliza a técnica “Random Forests” para a imputação dos dados, a qual é baseada em árvores de decisão para se tentar prever qual dado preencheria esse dado incompleto (Arora & Kaur, 2020; Aulia et al., 2019; Recknagel, 2001; Silver, Yang, & Li, 2013).

Tendo sido finalizado o ajuste das informações do banco de dados, passou-se então a análise descritiva das variáveis de forma independente, segundo a sua distribuição em relação a variável classe. Com base no número de avaliações recebidas do restaurante, pode observar que devido a sua distribuição irregular não permitindo estabelecer uma relação visual, o que pode indicar que não será um bom preditor para o modelo, porém decidiu-se mantê-la para fins de análise. Tendo sido finalizada a análise descritiva da variável “num_rev_recebidas” passou-se então para a demonstração da variável “aval_ambiente”, por meio da qual os clientes avaliam o ambiente do restaurante. Posteriormente passou-se então para a variável “aval_comida”, que é utilizada para que os clientes avaliem a comida servida no restaurante, na qual quanto menores a notas, maior é a probabilidade de que o restaurante seja “mal avaliado”, ou “bem avaliado” a partir daí, indicando ser uma boa variável preditora.

Após a variável da avaliação da comida ser finalizada, passou-se então a análise da variável “aval_preco”, na qual são avaliados os preços praticados pelos restaurantes segundo a percepção do cliente, posteriormente passou-se então para a visualização dos dados da variável “aval_servico”, na qual são avaliados os serviços prestados pelos restaurantes japoneses.

Com a finalização da análise descritiva das variáveis que compõem o modelo passou-se então a preparação dos dados para treinamento e teste, os quais foram separados de forma aleatória por meio do método sample, a proporção de 70% para treinamento e 30% para teste.

Figura 2 – Acuracidade do modelo treinado - Naïves Bayes



Fonte: Elaborada com dados da pesquisa.

A Figura 2 demonstra a evolução da acuracidade do modelo na medida em que ele é treinado, em seguida foi verificado a matriz de

confusão, que visa demonstra a taxa de acerto esperada para o modelo. Os valores são demonstrados por meio da Tabela 1.

Tabela 1 – Matriz de Confusão

	Bem Avaliado	Mal Avaliado
Bem Avaliado	64	10
Mal Avaliado	16	72

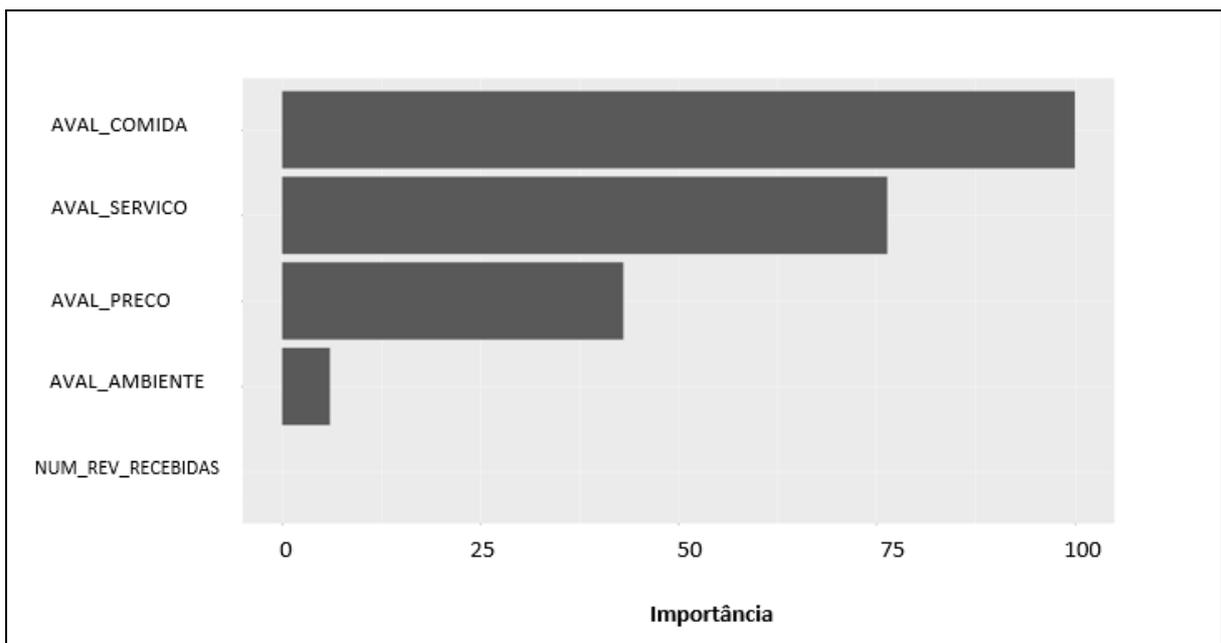
Fonte: Elaborada com dados da pesquisa (2020).

Por meio da diagonal da Tabela 1, chega-se a acurácia do modelo que é de 0,8395, com intervalo de confiança de 95% (0,7737, 0,8924), Teste de Macnemar = 0,3268 e Kappa de 0,6787, permitindo inferir uma taxa de acerto de cerca de 0,84% e uma taxa de erro de cerca de 0,16%, ou seja, para cada 10 previsões

o modelo apenas 1,6% será errada com margem de 5% para mais ou para menos.

A Figura 3 demonstra a variáveis preditoras mais importantes para o modelo, no sentido de determinar a classificação dos restaurantes japoneses.

Figura 3 – Importância das variáveis na previsão da Classificação do Restaurante



Fonte: Elaborada com dados da pesquisa (2020).

Com as análises demonstradas no decorrer do trabalho, a conexão de métodos de inteligência artificial usado a favor das análises de dados para medir a aceitação do cliente no desempenho desses 605 restaurantes japoneses auferidos, seja o olhar que o consumidor tem sobre o ambiente, em como é recepcionado,

seja na ideia do cliente ir fisicamente no local ou até mesmo no momento de fazer uma reserva; como a comida é avaliada por esses clientes, sua apresentação, o frescor ou até mesmo a variedade que está disponível a seus comensais; a questão do preço que o cliente dispense para ir até um desses

estabelecimentos, para ele compensa? Ele sairá satisfeito? Terá um custo benefício atrativo ou só o custo, o preço por si só? Além dessas observações, o fator primordial e talvez o que tenha maior representatividade no quesito de “bem avaliados” ou “mal avaliados”, está no serviço, que se tratando de restaurantes japoneses em São Paulo, existe uma diversidade de formatos de serviço nesse segmento de alimentação.

A ideia é prover uma ferramenta na qual os donos de restaurantes podem coletar essas opiniões a partir dos seus clientes de forma simples e conseguir prever qual avaliação eles receberiam nessa plataforma, a qual muitos restaurantes utilizam como ferramenta para se adaptar ao ambiente competitivo.

5 Conclusão

O objetivo do presente trabalho foi apresentar o desenvolvimento de um projeto de Machine Learning para prever a classificação do cliente em relação ao restaurante, possibilitando dessa forma a utilização da Hospitalidade como um diferencial competitivo. Para tanto, no caso da presente pesquisa o fenômeno estudado foi o desenvolvimento de um modelo de classificação de avaliações de restaurantes japoneses na cidade de São Paulo, de forma a permitir que os proprietários desses estabelecimentos consigam prever de que maneira serão avaliados por seus clientes na Plataforma TripAdvisor, auxiliando aos

gestores na tomada de decisão para melhorar a gestão dos respectivos restaurantes.

Assim, como outros segmentos de negócios, os restaurantes estão em risco de falência desde a sua projeção, seja no plano de negócio, na implantação física, na contratação de mão de obra, nos recursos financeiros, principalmente e, sobretudo, na gestão dessas questões levantadas. Esse estabelecimento pode ser gerido por um profissional da área de alimentos e bebidas ou não, porém questões particulares do setor são essenciais para ter sucesso neste segmento.

Diante do exposto, pode-se afirmar que a hospitalidade ainda está em transformação no setor de alimentos e bebidas, visto que a cada dia surgem novas maneiras de medir o que entende-se por hospitalidade na visão do cliente, o atendimento bom ou ruim, os preços, o serviço com um todo, principalmente para os frequentadores de restaurantes japoneses, segmento ainda em crescimento e com muito potencial de explorar as múltiplas formas de receber, acolher e alimentar. Para o sucesso de uma empresa, é de se considerar que não só as estratégias executadas pelos empreendedores, como também a interação e envolvimento dos demais atores nessa cadeia, seja fornecedores, colaboradores e clientes, compõem o cenário exposto na hospitalidade do domínio comercial.

Considera-se, finalmente, que os resultados do estudo sugerem que, embora a hospitalidade seja um tema em amplas vertentes, o diferencial competitivo na gestão de restaurantes japoneses é um assunto ainda a

ser explorado. Fica claro que os resultados do estudo são especificamente para uma tipologia de restaurantes dentro de uma única plataforma de avaliação, para tanto, os resultados não podem ser generalizados.

Referências

- ABRASEL. (2019). *Home*. Abrasel.
<https://abrase.com.br/>
- Arora, N., & Kaur, P. D. (2020). A Bolasso based consistent feature selection enabled random forest classification algorithm: An application to credit risk assessment. *Applied Soft Computing*, 86, 105936.
<https://doi.org/10.1016/j.asoc.2019.105936>
- Asafu-Adjei, J. K., & Betensky, R. A. (2015). A Pairwise Naïve Bayes Approach to Bayesian Classification. *International Journal of Pattern Recognition and Artificial Intelligence*, 29(07), 1550023.
<https://doi.org/10.1142/S0218001415500238>
- Aulia, A., Jeong, D., Saaïd, I. M., Kania, D., Shuker, M. T., & El-Khatib, N. A. (2019). A Random Forests-based sensitivity analysis framework for assisted history matching. *Journal of Petroleum Science and Engineering*, 181, 106237.
<https://doi.org/10.1016/j.petrol.2019.106237>
- Babbie, E. R. (1999). *Metodos De Pesquisas De Survey*.
- Brotherton, B., & Wood, R. C. (2001). Hospitality and hospitality management. In *In search of hospitality* (pp. 134-156). Butterworth-Heinemann.
- Brotherton, B., & Wood, R. C. (2008). The nature and meanings of “hospitality.”. The SAGE handbook of hospitality management, 37-61.
- Chen, C., Zhang, G., Yang, J., Milton, J. C., & Alcántara, A. “Dely”. (2016). An explanatory analysis of driver injury severity in rear-end crashes using a decision table/Naïve Bayes (DTNB) hybrid classifier. *Accident Analysis & Prevention*, 90, 95–107.
<https://doi.org/10.1016/j.aap.2016.02.002>
- Chen, W., Zhang, S., Li, R., & Shahabi, H. (2018). Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. *Science of The Total Environment*, 644, 1006–1018.
<https://doi.org/10.1016/j.scitotenv.2018.06.389>
- Demo, P. (2000). *Metodologia do Conhecimento Científico* (Edição: 1ª). Atlas.
- Drucker, P. F. (2000). *Desafios Gerenciais Para o Século XXI* (1ª edição). Cengage.
- Fonseca, M. T. (2018). *Tecnologias gerenciais de restaurantes*. Senac.
- Gil, A. C. (2008). *Métodos e Técnicas de Pesquisa Social* (Edição: 6ª). Atlas.
- Godbout, J. T. (2019). To receive is to give. *Revue du MAUSS*, (1), 159-174.
- Jiang, L., Cai, Z., Zhang, H., & Wang, D. (2013). Naive Bayes text classifiers: A locally weighted learning approach. *Journal of Experimental & Theoretical Artificial Intelligence*, 25(2), 273–286.
<https://doi.org/10.1080/0952813X.2012.721010>
- Jones, P. (Ed.). (2002). *Introduction to hospitality operations: An indispensable guide to the industry*. Cengage Learning EMEA.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260.
<https://doi.org/10.1126/science.aaa8415>
- Karandikar, J., McLeay, T., Turner, S., & Schmitz, T. (2015). Tool wear monitoring using naïve Bayes classifiers. *The International Journal of Advanced Manufacturing Technology*, 77(9–12), 1613–1626. <https://doi.org/10.1007/s00170-014-6560-6>
- Kerlinger, F. N. (1988). *Metodologia da Pesquisa em Ciências Sociais. Um Tratamento Conceitual* (Edição: 8ª). EPU.
- King, C. A. (1995). What is hospitality? *International Journal of Hospitality Management*, 14(3-4), 219-234.
- Kou, G., Chao, X., Peng, Y., Alsaadi, F. E., & Herrera-Viedma, E. (2019). Machine learning methods for systemic risk analysis in financial sectors. *Technological and Economic Development of Economy*, 25(5), 716–742.
<https://doi.org/10.3846/tede.2019.8740>
- Kubat, M. (2017). *An Introduction to Machine Learning*. Springer International Publishing.
<https://doi.org/10.1007/978-3-319-63913-0>

- Lashley, C. (2004). Em busca da hospitalidade: Perspectivas para um mundo globalizado (1a edição). Editora Manole.
- Lashley, C. (2015). Hospitality and hospitableness. *Research in Hospitality Management*, 5(1), 1-7.
- Lashley, C., & Morrison, A. (2010). In search of hospitality. Routledge.
- Lashley, C., Lynch, P., & Morrison, A. (2007). Hospitality: an introduction. In *Hospitality: A social lens* (pp. 15-30). Routledge.
- Li, C., Jiang, L., & Li, H. (2014). Naive Bayes for value difference metric. *Frontiers of Computer Science*, 8(2), 255–264. <https://doi.org/10.1007/s11704-014-3038-5>
- Link, C. P., Silva, G., Barichello, R., & Magro, C. B. D. (2020). Fatores críticos no gerenciamento de projetos públicos sustentáveis. *Revista de Gestão e Projetos*, 11(2), 87–109. <https://doi.org/10.5585/gep.v11i2.17490>
- López-Cruz, P. L., Bielza, C., & Larrañaga, P. (2015). Directional naive Bayes classifiers. *Pattern Analysis and Applications*, 18(2), 225–246. <https://doi.org/10.1007/s10044-013-0340-z>
- Marconi, M. de A., & Lakatos, E. M. (2017). *Metodologia científica* (Edição: 7). Atlas.
- Mullins, L. J. (2001). *Hospitality management and organisational behaviour*. Pearson Education.
- Pitt-Rivers, J. (2017). *From Hospitality to Grace: A Julian Pitt-Rivers Omnibus*. Hau Books.
- Powers, T. F., & Barrows, C. W. (2003). *Introduction to the hospitality industry*. John Wiley & Sons.
- Prado, E. P. V., Mancini, M., & Grotta, A. (2020). Estilo de liderança e sucesso em projetos de TI em organizações brasileiras. *Revista de Gestão e Projetos*, 11(2), 64–86. <https://doi.org/10.5585/gep.v11i2.17387>
- R Foundation. (2019). *R: The R Foundation*. <https://www.r-project.org/foundation/>
- Recknagel, F. (2001). Applications of machine learning to ecological modelling. *Ecological Modelling*, 146(1–3), 303–310. [https://doi.org/10.1016/S0304-3800\(01\)00316-7](https://doi.org/10.1016/S0304-3800(01)00316-7)
- Russo, R. de F. S. M., & Silva, L. F. da. (2019). Critérios de sucesso e fatores de sucesso: É crítico distinguir o significado de ambos. *Revista de Gestão e Projetos*, 10(2), Article 2. <https://doi.org/10.5585/gep.v10i2.14614>
- Sánchez-Franco, M. J., Navarro-García, A., & Rondán-Cataluña, F. J. (2019). A naive Bayes strategy for classifying customer satisfaction: A study based on online reviews of hospitality services. *Journal of Business Research*, 101, 499–506. <https://doi.org/10.1016/j.jbusres.2018.12.051>
- Santos, A., Mendes, S., & Corte-Real, J. (2014). siC notícias online. (sicnoticias. sapo.pt/pais/2014/01/07/varios-restaurantes-e-baresnas-praias-destruidos-na-marginal-de-cascais). Last access on January, 7.
- SEBRAE. (2019). Sebrae SÃO PAULO SP - Sebrae. <https://www.sebrae.com.br/sites/PortalSebrae/ufs/sp?codUf=26>
- Sharma, D., Kumar, B., & Chand, S. (2019). A Trend Analysis of Machine Learning Research with Topic Models and Mann-Kendall Test. *International Journal of Intelligent Systems and Applications*, 11(2), 70–82. <https://doi.org/10.5815/ijisa.2019.02.08>
- Shiri Harzevili, N., & Alizadeh, S. H. (2018). Mixture of latent multinomial naive Bayes classifier. *Applied Soft Computing*, 69, 516–527. <https://doi.org/10.1016/j.asoc.2018.04.020>
- Silver, D. L., Yang, Q., & Li, L. (2013, março 15). Lifelong Machine Learning Systems: Beyond Learning Algorithms. *2013 AAAI Spring Symposium Series*. 2013 AAAI Spring Symposium Series. <https://www.aaai.org/ocs/index.php/SSS/SSS13/paper/view/5802>
- Skiena, S. S. (2017). *The Data Science Design Manual*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-55444-0>
- Spolon, A. P. G. (2015). Um exercício de hospitalidade. *Revista Hospitalidade*, 3-8.
- Tideman, M. C. (1983). External influences on the hospitality industry. *The management of hospitality.*, 1-23.
- Villani, J., Schully, S. D., Meyer, P., Myles, R. L., Lee, J. A., Murray, D. M., & Vargas, A. J. (2018). A Machine Learning Approach to Identify NIH-Funded Applied Prevention Research. *American Journal of Preventive Medicine*, 55(6), 926–931. <https://doi.org/10.1016/j.amepre.2018.07.024>

- Vural, M. S., & Gök, M. (2017). Criminal prediction using Naive Bayes theory. *Neural Computing and Applications*, 28(9), 2581–2592. <https://doi.org/10.1007/s00521-016-2205-z>
- Walker, J. R. (2002). Introduction to hospitality. Prentice Hall.
- Wang, S., Jiang, L., & Li, C. (2015). Adapting naive Bayes tree for text classification. *Knowledge and Information Systems*, 44(1), 77–89. <https://doi.org/10.1007/s10115-014-0746-y>
- Wen, S., Wang, C., Li, H., & Zheng, G. (2019). Parallel naïve Bayes regression model-based collaborative filtering recommendation algorithm and its realisation on Hadoop for big data. *International Journal of Information Technology and Management*, 18(2–3), 129–142. <https://doi.org/10.1504/IJITM.2019.099818>
- Wood, R. C. (1992). Hospitality industry labour trends: British and international experience. *Tourism Management*, 13(3), 297-304.
- Xu, S. (2018). Bayesian Naïve Bayes classifiers to text classification. *Journal of Information Science*, 44(1), 48–59. <https://doi.org/10.1177/0165551516677946>