

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

GUILHERME MORETE FELIX

**APLICAÇÃO DE TÉCNICAS DE MACHINE LEARNING PARA ESTUDAR A
SATISFAÇÃO DE PASSAGEIROS EM AEROPORTOS BRASILEIROS**

**LONDRINA
2023**

GUILHERME MORETE FELIX

**APLICAÇÃO DE TÉCNICAS DE MACHINE LEARNING PARA ESTUDAR A
SATISFAÇÃO DE PASSAGEIROS EM AEROPORTOS BRASILEIROS**

**Application of machine learning techniques to study passengers satisfaction in
brazilian airports**

Trabalho de conclusão de curso de graduação
apresentado como requisito para obtenção do título de
bacharel em Engenharia de Produção da Universidade
Tecnológica Federal do Paraná (UTFPR).

Orientador: Prof. Dr. Bruno Samways dos Santos

**LONDRINA
2023**



[4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

GUILHERME MORETE FELIX

**APLICAÇÃO DE TÉCNICAS DE MACHINE LEARNING PARA ESTUDAR A
SATISFAÇÃO DE PASSAGEIROS EM AEROPORTOS BRASILEIROS**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do título de
Bacharel em Engenharia de Produção da Universidade
Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 24/Novembro/2023

Bruno Samways dos Santos

Doutor

Universidade Tecnológica Federal do Paraná

Rafael Henrique Palma Lima

Doutor

Universidade Tecnológica Federal do Paraná

Rogério Tondato

Doutor

Universidade Tecnológica Federal do Paraná

LONDRINA

2023

Dedico esse trabalho à minha família, pelas provas
de amor e compreensão.

AGRADECIMENTOS

Primeiramente, agradeço a Deus pelo dom da vida e por me permitir levantar todos os dias e batalhar pelos meus sonhos. Graças a ele, pude conhecer pessoas incríveis durante todo o processo e encontrar a área que pretendo trabalhar e contribuir durante o restante da minha carreira.

Agradeço a minha família, principalmente aos meus pais, Dácio e Anagela, por toda a compreensão, companheirismo, auxílio e provas de amor nesse período, sou eternamente grato e privilegiado por poder chamar vocês de pais.

Agradeço também ao professor Dr. Bruno Samways dos Santos, por toda a orientação durante o trabalho e pelos ensinamentos ao longo da graduação, com toda certeza cresci muito como profissional devido a nossas conversas.

Enfim, gostaria de agradecer a todas as pessoas que tive contato durante o processo, a qual, de uma forma ou de outra, me auxiliaram a atingir o objetivo e de ser um ser humano melhor a cada dia.

RESUMO

A satisfação percebida do cliente em serviços e produtos é um fator chave para o sucesso de qualquer empresa. Com relação aos aeroportos, a satisfação do usuário compõe um dos quatorze indicadores de qualidade de serviço (IQS) elaborados pela Agência Nacional de Aviação Civil (ANAC), os quais descrevem o nível de qualidade do aeroporto e estão diretamente atrelados aos reajustes de tarifas aéreas. Dessa forma, para que aeroportos consigam melhorar a qualidade e aumentar o valor agregado de seu serviço, é necessário compreender as características e os aspectos que levam os usuários a terem a sensação de satisfação ou insatisfação do serviço prestado. Nesse contexto, o presente trabalho propôs de modelos de *machine learning* para a classificação da satisfação de usuários e identificar as variáveis mais importantes para a satisfação do usuário, a partir de dados fornecidos pela pesquisa de satisfação realizada pela Secretaria Nacional de Aviação Civil (SAC). Foram utilizados os algoritmos Árvore de Decisão, Floresta Aleatória, Regressão Logística, *MultiLayer Perceptron*, KNN, *Naive-Bayes* e *XGBoost*, sendo que a avaliação de cada modelo se baseou na acurácia, *recall*, precisão e *f1-score*. Hiperparametrização e estratégias de balanceamento de classe como *oversampling*, *undersampling* e *oversampling* por *resample* foram empregadas. Como resultado, os modelos com melhores performances foram a Regressão Logística e Floresta Aleatória, com acurácia acima de 85% e *recall* para usuários insatisfeitos acima de 75%, mas precisões abaixo de 50% para a maioria dos modelos. Com relação às variáveis mais importantes, os modelos Regressão Logística e Floresta Aleatória consideraram como melhores preditoras as variáveis ligadas a conforto e limpeza, bem como localização e deslocamento no aeroporto.

Palavras-chave: *Machine Learning*, Mineração de Dados, Aeroportos, Satisfação, Classificação.

ABSTRACT

Perceived customer satisfaction in services and products is a key factor in the success of any company. In relation to airports, user satisfaction is one of the fourteen service quality indicators (IQS) prepared by the National Civil Aviation Agency (ANAC), which describe the airport's quality level and are directly linked to air fare adjustments. Therefore, for airports to be able to improve the quality and increase the added value of their service, it is necessary to understand the characteristics and aspects that lead users to feel satisfied or dissatisfied with the service provided. In this context, the present work proposed machine learning models to classify user satisfaction and identify the most important variables for user satisfaction, based on data provided by the satisfaction survey carried out by the National Civil Aviation Secretariat (SAC). The Decision Tree, Random Forest, Linear Regression, MultiLayer Perceptron, KNN, Naive-Bayes and XGBoost algorithms were used, with the evaluation of each model being based on accuracy, recall, precision and f1-score. Hyperparameterization and class balancing strategies such as oversampling, undersampling and oversampling by resample were employed. As a result, the best performing models were Logistic Regression and Random Forest, with accuracy above 85% and recall for dissatisfied users above 75%, but precision under 50% for most models. Regarding the most important variables, the Logistic Regression and Random Forest models considered variables linked to comfort and cleanliness, as well as location and movement at the airport, the best predictors.

Keywords: Machine Learning, Data Mining, Airports, Satisfaction, Classification.

LISTA DE FIGURAS

Figura 1 -	Etapas do processo de KDD.....	5
Figura 2 -	Exemplo de conjunto de treino para tarefa de classificação.....	8
Figura 3 -	Fluxograma de um modelo supervisionado de <i>Machine Learning</i>	9
Figura 4 -	Modelo de Árvore de Decisão para jogar tênis.....	11
Figura 5 -	Modelo Floresta Aleatória para classificação.....	11
Figura 6 -	Representação da distância Euclidiana entre dois pontos.....	12
Figura 7 -	Jornada do passageiro nos aeroportos.....	18
Figura 8 -	Fluxograma das atividades realizadas no estudo.....	25
Figura 9 -	Análise de valores nulos de uma fração do conjunto de dados original através do heatmap.....	27
Figura 10 -	Histograma das variáveis numéricas.....	29
Figura 11 -	Boxplots das variáveis numéricas com relação ao nível de satisfação geral dos usuários.....	46

LISTA DE TABELAS

Tabela 1 -	Variáveis utilizadas na avaliação dos modelos.....	28
Tabela 2 -	Métricas sem balanceamento de classe.....	30
Tabela 3 -	Métricas com balanceamento por undersampling.....	31
Tabela 4 -	Métricas com balanceamento por oversampling.....	33
Tabela 5 -	Métricas com oversampling por resample.....	34
Tabela 6 -	Métricas no conjunto de undersampling com os melhores parâmetros obtidos no Grid-Search.....	35
Tabela 7 -	Variáveis mais importantes com os melhores parâmetros obtidos no grid-search e de trabalhos correlatos	37
Tabela 8 -	Métricas sem balanceamento de classe.....	39
Tabela 9 -	Métricas com balanceamento por undersampling.....	40
Tabela 10 -	Métricas com balanceamento por oversampling.....	41
Tabela 11 -	Métricas com oversampling por resample.....	42
Tabela 12 -	Métricas no conjunto de undersampling com os melhores parâmetros obtidos no grid-search.....	43
Tabela 13 -	Usuários com Satisfação Geral 3 - Acurácia no conjunto de undersampling com os melhores parâmetros obtidos no grid-search.....	48
Tabela 14 -	Natureza das variáveis do conjunto de dados.....	57
Tabela 15 -	Parâmetros utilizados no grid-search para cada algoritmo.....	61

LISTA DE QUADROS

Quadro 1 -	Representação da matriz de confusão.....	15
Quadro 2 -	Resumo de cada trabalho correlato.....	21

LISTA DE ABREVIATURAS E SIGLAS

ANAC	Agência Nacional de Aviação Civil
IQS	Índice de Qualidade de Serviço
KDD	<i>Knowledge Discovery in Databases</i> (Descoberta de conhecimento baseada em dados)
PND	Programa Nacional de Desestatização
SAC	Secretaria Nacional de Aviação Civil
MLP	<i>Multilayer Perceptron</i>

SUMÁRIO

1 INTRODUÇÃO	1
1.1 Objetivo Geral	3
1.2 Objetivos Específicos	3
2 REFERENCIAL TEÓRICO	4
2.1 Descoberta de Conhecimento e a Mineração de Dados	4
2.1.1 Seleção.....	5
2.1.2 Pré-Processamento.....	6
2.1.3 Transformação.....	6
2.1.4 Mineração de Dados.....	7
2.1.5 Interpretação e Avaliação.....	7
2.2 Machine Learning	7
2.2.1 Modelo Supervisionado.....	8
2.2.2 Modelo Não Supervisionado.....	9
2.3 Técnicas de Classificação	10
2.3.1 Árvore de Decisão.....	10
2.3.2 Floresta Aleatória.....	11
2.3.3 KNN.....	12
2.3.4 Naive Bayes.....	13
2.3.5 Regressão Logística.....	14
2.4 Métricas de Avaliação	14
2.4.1 Matriz de Confusão.....	15
2.5 Satisfação em Serviços	17
2.6 Trabalhos Correlatos	18
3 METODOLOGIA	23
3.1 Descrição do Conjunto de Dados	23

3.2 Atividades Desenvolvidas	23
3.3 Ferramentas Utilizadas	25
4. RESULTADOS E DISCUSSÕES.....	27
4.1 Análise Descritiva	27
4.2 Avaliação dos Modelos	29
4.2.1 Sem Balanceamento de Classe	30
4.2.2 Balanceamento por <i>Undersampling</i>	31
4.2.3 Balanceamento por <i>Oversampling</i>	32
4.2.4 <i>Oversampling</i> por <i>Resample</i>	34
4.2.5 Melhores parâmetros obtidos pelo <i>Grid-Search</i>	35
4.2.6 Avaliação do desempenho dos modelos	36
4.3 Variáveis Mais Importantes para Classificação da Satisfação.....	36
4.4 Avaliação do Modelos com Determinação Alternativa de Satisfeitos e Insatisfeitos	38
4.4.1 Sem Balanceamento de Classe	38
4.4.2 Balanceamento por <i>Undersampling</i>	39
4.4.3 Balanceamento por <i>Oversampling</i>	41
4.4.4 <i>Oversampling</i> por <i>Resample</i>	42
4.4.5 Melhores Parâmetros Obtidos pelo <i>Grid-Search</i>	43
4.4.6 Avaliação de Desempenho dos Modelos	44
4.5 Análise de Usuários com Satisfação Geral Intermediária	45
4.5.1 Análise Exploratória dos Usuários com Satisfação Intermediária	45
4.5.2 Avaliação dos Modelos dos Usuários com Satisfação Intermediária	48
5 CONCLUSÃO	50
REFERÊNCIAS	52
APÊNCIDE A.....	57
APÊNDICE B.....	61

1 INTRODUÇÃO

De acordo com Goldschmidt (2015), o avanço da tecnologia possibilitou a criação e armazenamento de grandes volumes de dados, que podem ser transformados em informação e conhecimento para a tomada de decisões mais assertivas. O valor agregado de tais dados está diretamente ligado à capacidade de extrair conhecimento a partir destes.

Nesse contexto, o processo de descoberta e conhecimento baseado em dados (*Knowledge Discovery in Databases*, KDD), que obteve muita popularidade a partir da década de 90, por meio da adoção da comunidade científica e de trabalhos como de FAYYAD *et al.* (1996), fornece um *framework* para, a partir de uma base de dados e um objetivo de negócios, etapas são estabelecidas para que informações valiosas possam ser extraídas.

Uma etapa fundamental do processo KDD é a de Mineração de Dados, onde a base de dados é passada por algoritmos de aprendizado de máquina (*machine learning*) capazes de identificar padrões que podem ser do tipo preditivos, que tem como objetivo prever um valor ou classificar uma categoria, ou descritivos, que busca apresentar padrões que seriam de difícil reconhecimento pela análise puramente humana. De acordo com Goldschmidt (2015), nessa etapa ocorre a busca por conhecimentos úteis a partir dos dados.

Com o avanço tecnológico, a Mineração de Dados pode ser aplicada em processos cada vez mais distintos, aos quais não era possível ter uma aplicação prática devido aos dados não serem armazenados digitalmente. Além disso, a Mineração de Dados pode trazer informações sobre o negócio em que está sendo aplicada que nenhuma outra técnica poderia obter (AMARAL, 2016).

Com a privatização de aeroportos brasileiros e o número crescente de passageiros ao longo dos anos, se torna necessário adequar as estruturas e serviços disponíveis nos mesmos para elevar a satisfação do cliente. Visando incentivar a melhora da qualidade do serviço nos aeroportos, a ANAC possui uma série de indicadores de qualidade de serviço que levam em conta a satisfação dos usuários obtidas por meio de pesquisas de satisfação (COSTA *et al.*, 2021).

Nesse âmbito, o processo de KDD pode auxiliar na classificação da satisfação dos passageiros com relação a diversos atributos oriundos de pesquisas de

satisfação, possibilitando que as empresas identifiquem o perfil do usuário de forma mais assertiva e foquem na otimização dos parâmetros mais importantes para a satisfação do usuário.

Um dos fatores que auxiliam qualquer empresa a ter sucesso no mercado é a oferta de serviços que atendam ou superem o nível de satisfação de seus clientes. Quando os clientes se sentem surpreendidos positivamente ou satisfeito com uma empresa, o grau de confiabilidade destes com a empresa aumenta (ALVES,2015).

Por meio do conhecimento das necessidades dos clientes, é possível elaborar estratégias que visam a implantação de ações para melhorar os aspectos que os clientes consideram como principais para a qualidade total do produto ou serviço utilizado. Uma das maneiras de compreender as necessidades dos clientes é por meio de pesquisas de satisfação, onde, através de questionários, é possível avaliar o quanto cada cliente está satisfeito com a empresa em diversos âmbitos, tais como atendimento, imagem, estrutura e valor agregado (SANTOS *et al.*, 2017).

No contexto dos aeroportos brasileiros, de acordo com ANAC (2017), a pesquisa de satisfação de passageiros é um dos aspectos que compõem os indicadores de qualidade de serviço (IQS). Tais indicadores são considerados para o cálculo anual do fator Q, que incide nos cálculos do reajuste das tarifas aéreas, já atrelado em contrato, ao qual o valor do reajuste é diretamente proporcional ao valor do fator, caracterizando como um incentivo para a melhora da qualidade dos aeroportos.

Dessa forma, modelos de *machine learning* podem auxiliar empresas que gerenciam aeroportos brasileiros a classificar a satisfação de seus usuários com relação as diversas questões apresentadas na pesquisa de satisfação, possibilitando identificar os pontos que mais necessitam de melhorias e aperfeiçoá-los, levando a uma possível consequência uma maior satisfação do usuário e melhores indicadores de qualidade de serviço.

1.1 Objetivo Geral

Desenvolver diferentes modelos de *machine learning* visando estudar a satisfação de usuários em aeroportos brasileiros participantes da pesquisa realizada pela Secretaria Nacional de Aviação Civil, incluindo uma análise dos fatores mais importantes para os modelos classificadores relacionados a pesquisa de satisfação.

1.2 Objetivos Específicos

- Implementar modelos de *machine learning* do tipo classificação para estudar a satisfação de usuários em aeroportos.
- Verificar e analisar as métricas de desempenho entre diversos algoritmos propostos.
- Discutir as variáveis mais relevantes e fornecer direcionamento para a melhoria da qualidade de serviços em aeroportos.

2 REFERENCIAL TEÓRICO

Esta seção apresenta a definição e as etapas do KDD, bem como os tipos de *machine learning*, o que é a tarefa de classificação e a apresentação de alguns de seus algoritmos.

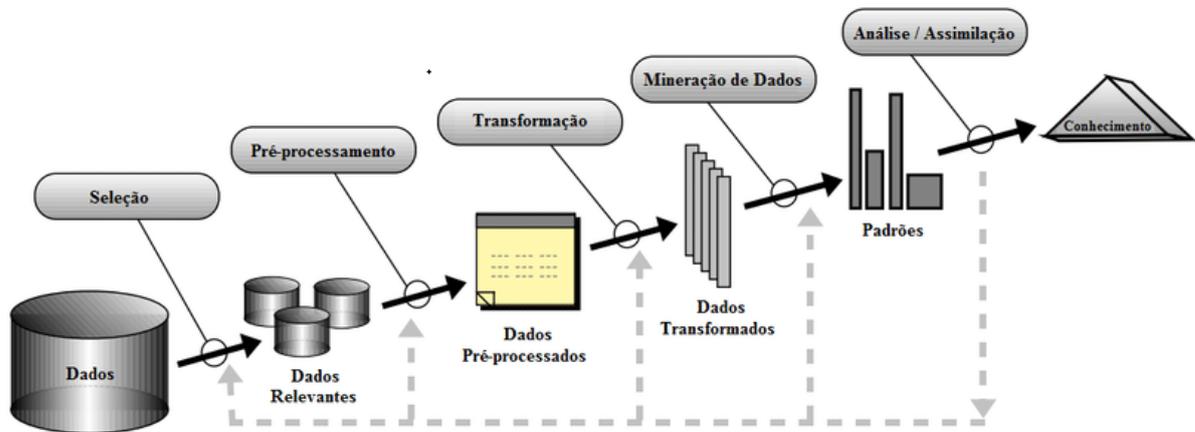
2.1 Descoberta de Conhecimento e a Mineração de Dados

Em um mundo globalizado e com dados sendo compartilhados em tempo real, é comum que se encontre base de dados de grande volume, necessitando que técnicas específicas sejam aplicadas para que informações e padrões possam ser extraídos, o que antes era feito, em grande maioria, por processos indutivos.

Com base nesse cenário, foi definido a Mineração de Dados como o processo de exploração analítica de dados para a obtenção de conhecimentos implícito e útil que podem ser utilizadas para auxiliar na tomada de decisões. (GOLDSCHMIDT, 2015).

O processo de descoberta de conhecimento baseada em dados, denominada KDD, é um processo não trivial em que é seguido uma série de etapas para a extração de conhecimento dos dados, sendo compreendido, de acordo com Fayyad *et al.* (1996), desde a seleção das fontes de dados utilizadas, pré-processamento, transformação e mineração de dados até a interpretação e avaliação das descobertas, tal como apresentado na Figura 1.

Figura 1 – Etapas do processo de KDD



Fonte: Adaptado de Fayyad *et al.* (1996, tradução livre).

2.1.1 Seleção

Nessa etapa, busca-se entender o problema de negócios ou o objetivo que se deseja alcançar no final da análise, bem como selecionar os tipos de dados, as variáveis e as amostras que irão ser utilizadas para as próximas etapas.

O conjunto de dados a ser analisado pode ser construído, por meio da junção direta ou orientada (GOLDSCHMIDT, 2015). Tais junções são definidas a seguir.

- **Junção Direta:** Todas as variáveis e registros de várias bases de dados são unidas em um novo conjunto de dados, sem uma análise prévia sobre quais variáveis podem contribuir para o processo de KDD.
- **Junção Orientada:** As bases de dados são previamente escolhidas por meio de um especialista no problema de negócio em questão e o profissional de KDD. Dessa forma, busca-se trazer para o conjunto de dados apenas as variáveis e registros que contribuam para o objetivo de análise.

É importante ressaltar que, independentemente do tipo de junção considerada para a criação do conjunto de dados, o mesmo deve passar por uma etapa de pré-processamento.

2.1.2 Pré-Processamento

A etapa de pré-processamento visa analisar a qualidade dos dados obtidos da etapa de seleção. Para isso, uma análise exploratória é realizada, visando encontrar e definir estratégias para tratar valores ausentes, mesmas variáveis com unidades de medidas diferentes, valores não condizentes com a realidade.

De acordo com Fayyad *et al.* (1996), também é visado a redução de variáveis no conjunto de dados por meio da seleção de variáveis que mostram dependência em relação ao objetivo de negócio.

A redução de variáveis também é importante para evitar a chamada 'Maldição da Dimensionalidade'. De acordo com Van der Maaten *et al.* (2007), tal fenômeno ocorre quando, à medida que as variáveis de um conjunto de dados aumentam, a quantidade de instâncias necessárias para que seja possível encontrar padrões significativos aumenta exponencialmente. Em muitas situações, não é possível obter o número de instâncias necessárias, trazendo perda de eficiências de modelos de *Machine Learning* e também podendo levar ao *overfitting*, quando o modelo tem uma boa performance apenas para as instâncias conhecidas, porém não é capaz de ter um desempenho aceitável para novas instâncias.

Dentro das correlações entre variáveis, podemos citar o teste do Qui Quadrado para a análise de variáveis qualitativas (não são variáveis numéricas) e o coeficiente de correlação de Pearson para variáveis quantitativas (variáveis numéricas, podendo ser de valores contínuos ou discretos). Além dessas técnicas, Mundstock (2006) relata que gráficos de dispersão e *boxplot* também podem ser utilizados para a detecção de *outliers*.

2.1.3 Transformação

A etapa de transformação consiste em escolher os algoritmos adequados de mineração e converter os dados para um padrão aceito pelos mesmos (FERREIRA *et al.*, 2018). É nessa fase em que variáveis quantitativas contínuas e discretas podem ser normalizadas, processo que consiste em padronizar os valores numéricos, como valores de uma variável que representa gastos mensais ou idade para pequenos

intervalos, tal como 0 a 1, bem como variáveis qualitativas podem ser binarizadas, transformando valores do tipo 'Aceito' e 'Rejeito' em 1 ou 0.

Como é levado em consideração o algoritmo em que os dados serão submetidos, alguns conjuntos de dados podem passar por maiores transformações do que outros.

2.1.4 Mineração de Dados

De acordo com Goldschmidt (2015), a etapa de mineração de dados é a parte principal do processo de KDD, sendo nessa etapa que padrões e novos conhecimentos sobre os dados são descobertos.

Nessa etapa, de acordo com Silva (2016), são aplicados algoritmos capazes de receber como entrada o conjunto de dados em análise e retornar um padrão de comportamento, podendo ser representado de maneira preditiva, ao qual tem o objetivo de prever valores futuros, ou descritiva, onde se busca encontrar padrões interpretáveis.

2.1.5 Interpretação e Avaliação

Nessa etapa, se busca apresentar o resultado de todo o processo de mineração de dados, podendo ser utilizadas ferramentas de visualização próprias para esse fim, tal como Power BI, Tableau ou Qlik Sense.

A partir da apresentação dos resultados, é possível definir novos objetivos de negócio, fazendo com que o processo retorne a etapas anteriores, aplicar os padrões encontrados, bem como documentar o trabalho realizado.

2.2 Machine Learning

De acordo com Mahesh (2020), *Machine Learning* é o campo de estudo em que dá aos computadores a habilidade de aprender sem serem explicitamente programados, sendo utilizados para lidar com conjunto de dados no processo KDD.

Diferentes modelos de aprendizado de máquina propõem algoritmos capazes de encontrar padrões nos dados e gerar hipóteses a partir desses. Ludermir (2018)

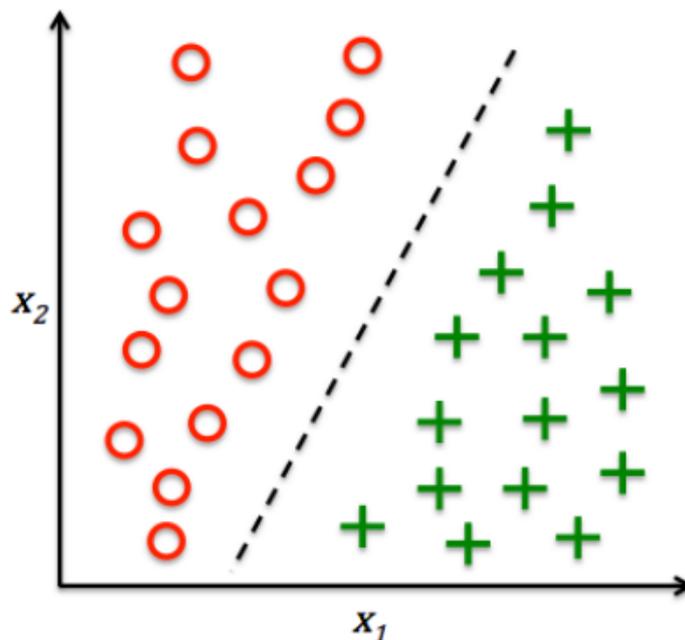
cita os principais modelos de aprendizado de máquina, sendo: Supervisionado, Não Supervisionado e por Reforço.

2.2.1 Modelo Supervisionado

Aprendizado supervisionado é quando o algoritmo se baseia em uma amostra de variáveis independentes e dependente (variável alvo) para determinar o valor da variável alvo de novas instâncias (LUDERMIR,2018).

O conjunto de dados é dividido em treino e teste, onde o algoritmo aprende padrões do conjunto de treino e o aplica para a predição ou classificação do conjunto de teste (MAHESH, 2020). Para tarefas de classificação, onde o objetivo é determinar a classe da variável alvo, Raschka (2015) mostra como exemplo um conjunto de treino com duas variáveis, 30 instâncias, sendo 15 dessas amostras de treino classificadas como negativas (círculos) e outras 15 como positivas (sinal de soma), mostrado na Figura 2. Utilizando esse conjunto de treino em tarefas de classificação, os modelos supervisionados de *Machine Learning* irão encontrar padrões nesses dados, criar uma fronteira de decisão, como sinalizada na linha preta pontilhada da Figura 2 e poderá classificar a variável alvo de novas instâncias do conjunto de teste.

Figura 2 – Exemplo de conjunto de treino para tarefa de classificação

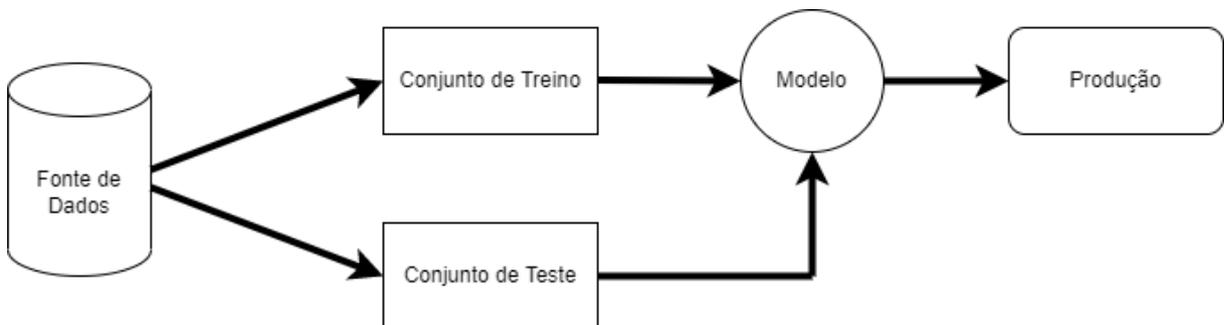


Fonte: Adaptado de Raschka (2015).

Dentre os algoritmos de *Machine Learning* para a tarefa de classificação, pode-se citar a Floresta Aleatória, Árvore de Decisão, KNN, *Navie Bayes*, Regressão Logística e *XGBoost*.

Com critérios de desempenho aceitos pela análise das métricas de cada modelo, é possível colocá-lo em produção (em uso), seguindo o fluxograma da Figura 3.

Figura 3 - Fluxograma de um modelo supervisionado de *Machine Learning*



Fonte: Adaptado de Mahesh (2020, tradução livre).

2.2.2 Modelo Não Supervisionado

Ao contrário do modelo supervisionado, onde existe um conjunto de treino, uma variável alvo e um conjunto de teste para avaliar o desempenho do modelo, no modelo não supervisionado, não existe nenhum tipo de separação dos dados, restando ao modelo encontrar padrões nos dados (MAHESH,2020).

Tais modelos possibilitam encontrar novas informações a partir dos dados que seriam muito custosos para serem observadas pela análise humana, podendo obter padrões de clusterização (relação entre as instâncias) ou associação (relação entre as variáveis) de um conjunto de dados.

Dentre os algoritmos não supervisionados, podemos citar o *K-Means Clustering*, *DBSCAN*, *Hierarchical* e *Apriori*.

2.3 Técnicas de Classificação

Atualmente existem vários algoritmos de classificação, indo desde técnicas já consolidadas, tal como Regressão Logística, até algoritmos mais recentes, como a Floresta Aleatória. Nesta seção, iremos apresentar as seguintes técnicas de classificação: Árvore de Decisão, Floresta Aleatória, KNN, *Naive Bayes* e Regressão Logística.

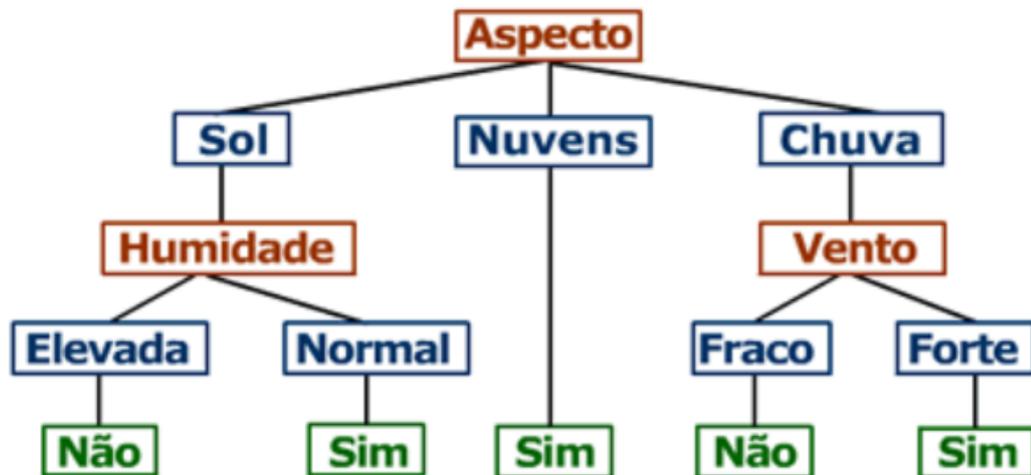
2.3.1 Árvore de Decisão

Árvore de Decisão é um dos algoritmos de classificação mais conhecidos na mineração de dados. Consiste de nós internos e nós folhas que seguem um modelo hierárquico e cada caminho percorrido segue uma regra do tipo 'Se, então'. Os nós internos são os atributos que o conjunto de dados possui e cada folha é um caminho sugerido pelo classificador.

De acordo com Schroff (2008), uma das desvantagens da árvore de decisão é a possibilidade da mesma apresentar problemas relacionados ao *overfitting*, ou seja, quando o modelo está muito alinhado com o conjunto de teste e não consegue fazer previsões assertivas para instâncias novas.

Como exemplo, na Figura 4, é possível analisar um modelo de decisão para jogar tênis, através de parâmetros como o aspecto do céu, umidade e vento. Cada sub-árvore é formada pelo valor que cada nó interno pode assumir (valores dos parâmetros), até que se chegue a decisão sugerida pelo classificador (nó folha, no caso, jogar ou não golfe).

Figura 4 – Modelo de Árvore de Decisão para jogar tênis

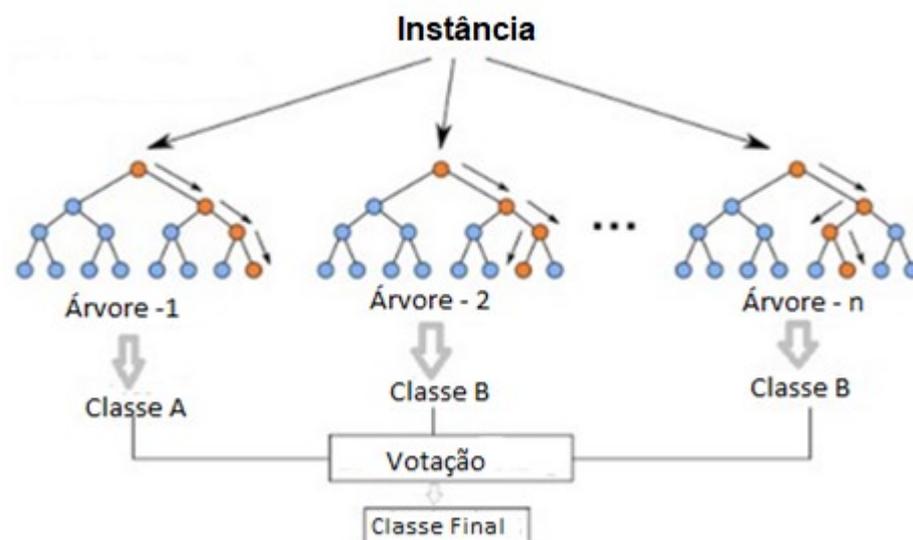


Fonte: Adaptado de GFBIOINFO (2023).

2.3.2 Floresta Aleatória

Com um funcionamento similar a Árvore de Decisão, de acordo com Schroff (2008) o algoritmo Floresta Aleatória busca mitigar o problema de *overfitting* por meio da criação de múltiplas árvores de decisão com instâncias aleatórias de seu conjunto de treino, sendo que a classificação da variável alvo se dá pelo valor resultante da maioria das árvores, como apresentado na Figura 5.

Figura 5 – Modelo Floresta Aleatória para classificação



Fonte: Adaptado de Lior Sinai (2020, tradução livre).

Apresentado pela primeira vez por Breiman (2001), possui bom comportamento tanto para classificação quanto para regressão, tendo como principais desvantagens o custo computacional para sua execução a depender do tamanho do conjunto de dados utilizado.

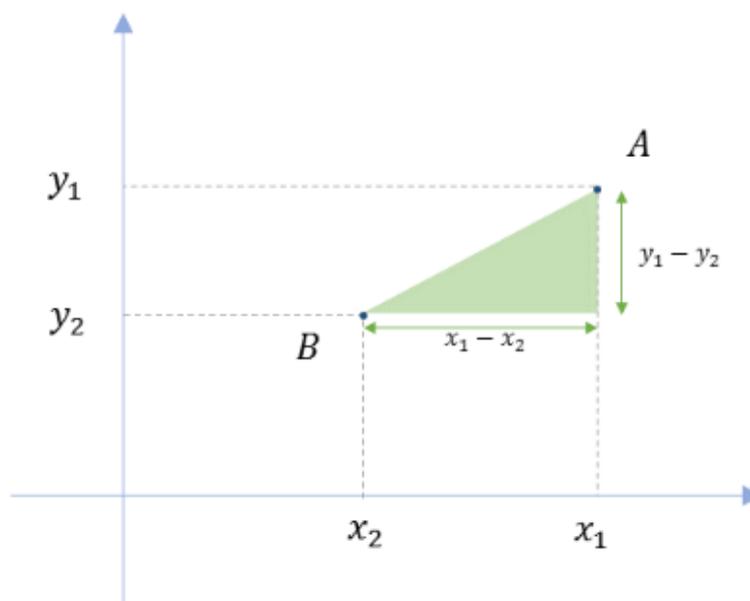
2.3.3 KNN

O algoritmo KNN utiliza as instâncias do conjunto de dados para basear seu aprendizado. Dessa forma, a variável alvo é determinada a partir da comparação das variáveis independentes com outras instâncias similares e que já possuem a variável alvo conhecida.

O processo de classificação é realizado por meio do valor mais frequente da variável alvo em instâncias similares com a variável já conhecida (instâncias do conjunto de teste). Tal relação de distância é definida por uma métrica de distância, onde diferentes métricas podem resultar em diferentes resultados (SILVA,2016).

Dentre as métricas de distância, a distância Euclidiana é a mais utilizada, pois representa a distância física em um espaço d-dimensional. Tal distância é representada pela linha contínua na Figura 6.

Figura 6 – Representação da distância Euclidiana entre dois pontos



Fonte: Adaptado de Oliveira (2018).

A distância Euclidiana para dois pontos A e B quaisquer pode ser representada pela equação 1.

$$d(A, B) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

Onde:

$d(A, B)$ = Distância Euclidiana entre os pontos A e B.

x_1 = Posição no eixo x do ponto A.

x_2 = Posição no eixo x do ponto B.

y_1 = Posição no eixo y do ponto A.

y_2 = Posição no eixo y do ponto B.

2.3.4 Naive Bayes

De acordo com Silva (2018), o algoritmo *Naive Bayes*, junto com a Floresta Aleatória, é um dos mais utilizados para tarefas de classificação devido ao seu desempenho tanto para dados categóricos quanto numéricos.

Ao ponto em que o algoritmo recebe uma instância para determinar a variável alvo, o algoritmo toma decisão de acordo com a probabilidade condicional da variável alvo daquela instância pertencer a cada uma de seus possíveis valores, tomando como base o conjunto de treino pelo Teorema de Bayes, indicado na equação 2.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (2)$$

Onde:

$P(A|B)$ = Probabilidade de pertencer a classe A, dado que as características B.

$P(B|A)$ = Probabilidade observar as características B, dado que pertence à classe A.

$P(A)$ = Probabilidade de pertencer a classe A.

$P(B)$ = Probabilidade de observar as características B.

2.3.5 Regressão Logística

Comumente utilizada para problemas de classificação em duas classes, é um algoritmo utilizado nos mais diversos ramos da ciência a partir do início do século XX e trata a variável alvo (variável dependente) com dois valores possíveis, 0 ou 1, podendo ser tais valores representados para 'Satisfeito' e 'Insatisfeito', 'Comprou' e 'Não Comprou', 'Verdadeiro' e 'Positivo', dentre outros.

De acordo com Netto (2021), o algoritmo trabalha estabelecendo um ponto de corte, que pode ser definido como a probabilidade no qual a variável alvo de uma determinada instância será considerada 1, ou, caso contrário, 0.

Para a Regressão Logística, a probabilidade de valores de uma amostra pertença a determinada classe é obtida por meio da função sigmoide, também chamada de função logística, obtida a partir da equação 3 (NETTO, 2021).

$$\text{sigmoide}(y) = \frac{1}{1 + e^y} \quad (3)$$

Onde y pode ser representado pela equação 4.

$$y = \beta_0 + \beta_1 * X_1 + \dots + \beta_n * X_n \quad (4)$$

Onde:

y = Probabilidade de uma amostra pertencer a determinada classe.

X = Variáveis independentes do modelo.

β = Coeficientes do modelo.

2.4 Métricas de Avaliação

Nesta seção, serão apresentados o conceito da matriz de confusão e as métricas utilizadas para a avaliação de desempenho da tarefa de classificação.

2.4.1 Matriz de Confusão

A matriz de confusão é uma ferramenta muito útil para avaliar modelos de classificação, pois resume os resultados entre o que foi previsto para o conjunto de teste com a valor da variável alvo real de tal conjunto, como pode ser observado no quadro 1.

Quadro 1 – Representação da matriz de confusão

		Valor Predito	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Fonte: Autoria Própria (2023).

Cada resultado da matriz de confusão é encaixado dentro de quatro categorias:

- Verdadeiro Positivo (VP): O modelo classificou a instância alvo como positivo e a mesma é de fato positiva.
- Falso Positivo (FP): O modelo classificou a variável alvo como positivo, porém a mesma é negativa.
- Falso Negativo (FN): O modelo classificou a variável como negativa, porém a mesma é positiva.
- Verdadeiro Negativo (VN): O modelo classificou a variável alvo como negativa e a mesma é de fato negativa.

A partir dos resultados da matriz de confusão, pode-se obter mais três importantes métricas de avaliação de modelos de classificação, sendo elas a Acurácia, Precisão, *recall*, Especificidade e f1-score.

- Acurácia: Mede quantas previsões do modelo foram de fato corretas, sendo obtida pela equação 5.

$$Acurácia = \frac{VP+VN}{VP+FP+VN+FN} \quad (5)$$

- Precisão: Mede, das instâncias em que o modelo classifica como positivo, qual é a razão entre as que foram classificadas positivas corretamente e incorretamente (Verdadeiro Positivo e Falso Positivo), sendo obtida pela equação 6.

$$Precisão = \frac{VP}{VP+FP} \quad (6)$$

- Recall: Também chamada de sensibilidade, mede, das instâncias em que o modelo classificou corretamente como positivo (Verdadeiro Positivo), qual é a razão entre as que de fato são positivas no conjunto de teste (Verdadeiro Positivo e Falso Negativo), sendo obtida pela equação 7.

$$Recall = \frac{VP}{VP+FN} \quad (7)$$

- Especificidade: Mede, das instâncias em que o modelo classificou corretamente como negativo (Verdadeiro Negativo), qual é a razão entre as que de fato são negativas no conjunto de teste (Verdadeiro Negativo e Falso Positivo), sendo obtida pela equação 9.

$$Especificidade = \frac{VN}{VN+FP} \quad (8)$$

- f1- score: É a média entre o recall e a especificidade, sendo obtida pela equação 9.

$$f1 - score = 2 \times \frac{(Precisão \times Sensibilidade)}{(Precisão+Sensibilidade)} \quad (9)$$

2.5 Satisfação em Serviços

Com o aumento da concorrência em todos os segmentos de mercado, várias empresas têm adotado a satisfação do cliente como um de seus objetivos centrais, buscando que o mesmo retorne e utilize novamente o serviço ou faça novamente uma compra, constituindo um ciclo virtuoso, pois, de acordo com Boulter(2013, *apud ALVES et al.,2015*) manter um cliente é mais rentável do que adquirir novos.

Com o cenário atual, as empresas se veem obrigadas a melhorar seus serviços, pois, de acordo com Lages e Piercy (2012 *apud Alves et al., 2015*), a satisfação do cliente possui uma relação estreita ligada com a qualidade do serviço. Como o ambiente de avaliação é dinâmico, é comum observar, tanto no meio acadêmico quanto gerencial, que novas variáveis sejam incluídas na análise de satisfação, sendo um desafio a quantificação da percepção do cliente acerca do serviço prestado.

Devido ao aumento da procura pelo serviço aéreo, sobretudo na Europa, a partir da década de 80, foi necessário a adoção de políticas de privatização de aeroportos para que os mesmos fornecessem aos usuários uma estrutura de qualidade. No Brasil, esse movimento teve maior incidência a partir do ano de 2008, em que foi incluído o aeroporto de São Gonçalo do Amarante no Programa Nacional de Desestatização (PND) (COSTA *et al.,2021*).

É importante salientar que, junto com a privatização, os aeroportos brasileiros privatizados possuem cláusulas contratuais onde a satisfação do cliente é um item chave. Uma das pesquisas de satisfação do cliente aplicadas atualmente em aeroportos é realizada pela Secretaria Nacional de Aviação Civil (SAC), em que avalia a percepção dos passageiros referente a infraestrutura, atendimento, serviços e processos aeronáuticos, ao qual busca mensurar a satisfação de toda a jornada do passageiro, desde o processo de aquisição de passagem até a restituição da bagagem, conforme mostrado na Figura 7.

Figura 7 – Jornada do passageiro nos aeroportos



Fonte: Ministério dos Transportes (2020).

Tal pesquisa é realizada desde janeiro de 2013, sendo que, de acordo com SAC (2023), existe uma evolução expressiva da satisfação dos passageiros em aeroportos brasileiros. Entre o período de 2013 a 2020, aproximadamente 85% dos entrevistados atribuíram notas 4 ou 5 (de no máximo 5) para os indicadores da entrevista.

A partir de 2020, através de discussões do Comitê Técnico de Desempenho Operacional (CTDO) da SAC, uma nova metodologia começou a ser utilizada para buscar um maior aprofundamento sobre os passageiros com opinião neutra e negativa sobre os processos e serviços prestados pelos aeroportos.

2.6 Trabalhos Correlatos

Homaid e Moulitsas (2022) avaliam a qualidade de serviços em aeroportos utilizando métodos de *Machine Learning* baseado em análise de sentimentos, a qual busca retirar o sentimento que determinado texto busca expressar. Para tal estudo, foi utilizado o conjunto de dados SKYTRAX, que conta com *reviews* de aproximadamente 600 aeroportos, a etapa de pré-processamento desconsiderou as

stopwords, que são palavras que não trazem sentido ao texto (que, ou, para, etc), foram utilizados os algoritmos de Regressão Logística, XGBoost, *Support Vector Machine*, Floresta Aleatória e Naiva Bayes e as métricas de avaliação usadas foram a precisão, *recall*, acurácia e *f1-score*. Como resultado do estudo, o algoritmo XGBoost apresentou melhor performance, obtendo valores em todas as métricas acima de 80%.

Bhaskar Sankaranarayanan, BV e Rathod (2016) utilizam dados de vários aeroportos ao redor do mundo para prever a satisfação dos usuários. No estudo, as variáveis utilizadas se baseiam na performance do aeroporto, no número e tempo de voos atrasados, bem como o número de estações que o aeroporto possui. Foi utilizado o modelo Árvore de Decisão, com o nível de satisfação sendo dividido em três classes: Alto, baixo e médio. Como resultado, obteve-se valores de acurácia superiores a 80% e de precisão acima de 70%, porém observou-se uma região confusa entre as classes de satisfação Alta e Baixo.

Singh e Dalpatadu (2018) utilizam como base a pesquisa de satisfação dos usuários do aeroporto internacional de São Francisco (EUA) para o nível de satisfação dos mesmos com relação às variáveis disponíveis na pesquisa, bem como determinar as *features* mais importantes para a determinação do nível de satisfação do usuário.

Para tal estudo, foi utilizado o algoritmo Floresta Aleatória e os valores da variável de satisfação foram agrupados em 'Inaceitável' e 'Abaixo da Média' como 'Insatisfeito' bem como na 'Média', 'Bom' e 'Sensacional' como 'Satisfeito'. Como resultado, o modelo apresentou uma acurácia de 98,5%, bem como a variável de limpeza geral, sinalização, exposições de arte e restaurantes foram as mais determinantes para predizer o nível de satisfação dos usuários.

Barakat, Yeniterzi e Martín-Domingo (2021) fazem uma análise de sentimentos de tweets na linguagem inglesa e árabe de usuários de aeroporto para medir a qualidade do serviço. Para o estudo, foram utilizados dois modelos de *Deep Learning*, sendo a rede neural convolucional (CNN) e a rede neural recorrente (LSTM). A análise considera as classificações de satisfação dos *tweets* como positiva, negativa e neutra. Ambos os modelos foram testados levando em consideração o conjunto de dados apenas com textos em inglês e outro apenas com textos em árabe. Como resultado, ambos os modelos tiveram valores de acurácia e *f1-score* acima de 70%.

Lucini et. al (2020) realizaram um estudo de mineração de texto para identificar a satisfação dos usuários de aeroportos com base em revisões online dos passageiros. Para tal trabalho, foram utilizadas mais de 55000 revisões de mais de 400 aeroportos diferentes, considerando como avaliação positiva as que possuem um score de 10, enquanto negativa as que possuem um score de 1. A mineração de texto foi feita utilizando o modelo Naive-Bayes, ao qual resultou em uma acurácia de 79,95%, enquanto que a variável mais determinante para a determinação da satisfação do usuário foi o tipo de cabine utilizada no voo.

Khedkar e Shinde (2018) utilizaram conjunto de dados de revisões de usuários de vários aeroportos para realizarem uma análise de sentimentos dos textos, ao qual podem ser classificadas em textos positivos ou negativos. Para o estudo, houve a *remoção das chamadas stopwords* na fase de pré-processamento e o modelo utilizado é resultando de um *ensemble* (combinação) de vários outros modelos, tal como Floresta Aleatória, *XGBoost*, Árvore de Decisão e *Naive Bayes*. Como resultado, houve um valor de acurácia de 79,83% e 78,97% de *f1-score*.

Koonsanit e Nishiuchi (2020) trazem uma proposta de classificação da satisfação de usuários de diversos serviços de acordo com a expressão facial que esses apresentam durante a utilização, levando também em consideração a idade e o gênero. Para tal, as expressões faciais dos usuários poderiam assumir, de acordo com a resposta que os mesmos forneceram através de um formulário, nas categorias de 'Raiva', 'Nojo', 'Medo', 'Felicidade', 'Tristeza', 'Surpresa' e 'Neutro'. Para o estudo, o método mais performático foi o *Support Vector Machine* com uma acurácia no *cross-validation* de 86%.

Mattozo et al.(2012) realizaram um estudo para determinar a satisfação global dos usuários do aeroporto de Parnamirim - RN e identificar as variáveis mais determinantes para a satisfação do usuário. O conjunto de dados utilizado foi a pesquisa de satisfação do aeroporto no mês de fevereiro de 2011 e o modelo de classificação utilizado foi a Regressão Logística. Como resultado, o modelo obteve acurácia maior que 88%, bem como as principais variáveis para a classificação de satisfação dos usuários foram o ambiente físico, a qualidade do relacionamento e a qualidade dos serviços prestados.

Kumas e Zymbler (2019) realizaram uma análise de sentimentos com base em tweets de usuários de diversas linhas aéreas com o objetivo de determinar o nível de

satisfação desses clientes. Para o estudo, os modelos utilizados foram o *Support Vector Machine* (SVM), rede neural convolucional (CNN) e rede neural artificial (ANN), com diversas alterações nos hiperparâmetros no *grid-search* para buscar a melhor acurácia. Como resultado, o modelo CNN obteve o melhor desempenho, com maior valor de acurácia de 92,3% no conjunto de validação, seguida da ANN e SVM, com 79,4% e 76,5% valor de acurácia no conjunto de validação, respectivamente.

O quadro 2 a seguir apresenta um resumo do contexto e dos resultados obtidos em cada trabalho correlato.

Quadro 2 - Resumo de cada trabalho correlato

Autores	Contexto	Modelos Utilizados	Resultados
Homaid e Moulitsas	Análise de sentimentos de avaliações de usuários de aeroporto.	Regressão Logística, XGBoost, <i>Support Vector Machine</i> , Floresta Aleatória e Naiva Bayes	Modelo XGBoost com acurácia e <i>f1-score</i> acima de 80%.
Bhaskar Sankaranarayanan, BV e Rathod	Satisfação com base em indicadores de performance de aeroportos.	Árvore de Decisão	Acurácia acima de 80% e precisão acima de 70%.
Singh e Dalpatadu	Satisfação dos usuários do aeroporto internacional de São Francisco a partir de pesquisas dos usuários.	Floresta Aleatória	Acurácia de 98,5%.
Barakat, Yeniterzi e Martín-Domingo	Análise de sentimentos de tweets de usuários de aeroportos.	Rede neural convolucional (CNN) e rede neural recorrente (LSTM)	Ambos modelos apresentaram valores de acurácia e <i>f1-score</i> acima de 70%.
Lucini et al.	Mineração de texto a partir de avaliações online de usuários de aeroportos.	Naive-Bayes	Acurácia de 79,95%.
Khedkar e Shinde	Análise de sentimentos de avaliações de usuários de aeroportos.	Ensemble dos modelos Floresta Aleatória, XGBoost, Árvore de Decisão e Naive Bayes	Acurácia de 79,83% e <i>f1-score</i> de 78,97%.
Koonsanit e Nishiuchi	Classificação da satisfação dos usuários de diversos produtos de acordo com a expressão facial.	<i>Support Vector Machine</i>	Acurácia no <i>cross-validation</i> de 86%.

Quadro 2 - Resumo de cada trabalho correlato

(Continuação)

Mattozo et al.	Satisfação dos usuários do aeroporto de Parnamirim - RN a partir de pesquisas dos usuários.	Regressão Logística	Acurácia de 88%.
Kumas e Zymbler	Análise de sentimentos de tweets de usuários de aeroportos.	<i>Support Vector Machine</i> (SVM), rede neural convolucional (CNN) e rede neural artificial (ANN).	Rede neural convolucional obteve o melhor desempenho com acurácia de 92,3% no conjunto de validação.

Fonte: Autoria Própria (2023).

3 METODOLOGIA

Esta seção apresenta a descrição do conjunto de dados, a sequência das atividades realizadas, bem como a descrição da linguagem, ambiente de trabalho e bibliotecas utilizadas.

3.1 Descrição do Conjunto de Dados

Para o presente estudo, foi utilizada a base de dados da pesquisa de satisfação em aeroportos a partir do ano de 2020 até o primeiro trimestre de 2023. Tal pesquisa, disponível em <https://dados.gov.br/dados/conjuntos-dados/pesquisa-de-satisfacao-do-passageiro-em-aeroportos>, é realizada de maneira presencial em uma série de aeroportos do país, com variáveis podendo ser de natureza tanto qualitativa, quanto quantitativa, essa última estando em uma escala de cinco pontos, ou em formato de data e hora. Todas as variáveis estão descritas de acordo com seu tipo de natureza no Apêndice A.

3.2 Atividades Desenvolvidas

Para o estudo, foram definidos como “Insatisfeitos” os usuários que deram nota 1,2 ou 3 na variável alvo “Satisfação Geral”, enquanto “Satisfeitos” se referem aos usuários que atribuíram valores 4 ou 5. Visando obter um número mais balanceado de instâncias entre usuários considerados insatisfeitos e satisfeitos, foram utilizados apenas os usuários que atribuíram notas de “4” ou “5” no primeiro semestre dos anos de 2020, 2021, 2022 e 2023, resultando em um conjunto de dados com 72.847 instâncias.

Com o conjunto de dados consolidado, foi realizada uma análise exploratória e retiradas todas as variáveis com mais de 35% de valores nulos, variáveis que não refletem a visão dos entrevistados (tal como aeroporto, data, mês, início e fim da coleta, terminal, portão e voo), e as instâncias que possuem valor nulo da variável “Satisfação Geral” e “Limpeza Geral do Aeroporto”, identificada como a variável com maior número de valores nulos.

Como resultado da retirada de instâncias e variáveis por valores nulos, o conjunto de dados possui 18 variáveis com 37.829 instâncias e nenhum valor nulo. As variáveis qualitativas, em virtude de cada época em que a pesquisa foi realizada, traz a mesma informação com grafias diferentes, como por exemplo, a escrita da linha aérea “Azul”, onde encontraram-se as variações do nome da empresa como “Azul”, “azul” e “AZUL”. Dessa forma, foi realizado um trabalho de agrupamento de/para dos valores das variáveis qualitativas.

Em seguida, com relação às variáveis quantitativas, foram retiradas variáveis que tinham um coeficiente de correlação de Pearson acima de 0,7 ou abaixo de -0,7 e, visando selecionar as variáveis qualitativas, foram retiradas variáveis que não possuíam associação significativa com a variável alvo em estudo (aceita a hipótese nula) através do método do Qui-quadrado, considerando um nível de significância de 0,05.

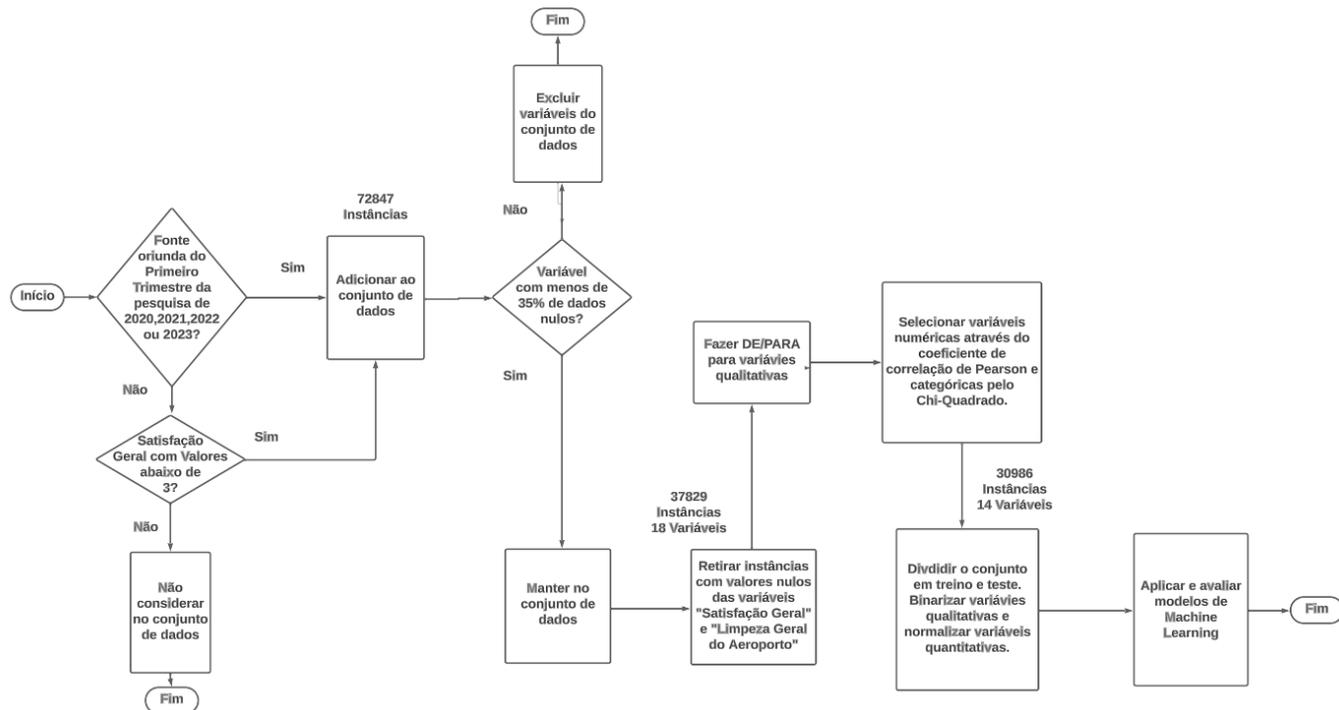
Com o conjunto de dados consolidado, o mesmo possui 13 variáveis preditoras e uma variável alvo, com 30986 instâncias, sendo 25587 instâncias de usuários classificados como satisfeitos e 5399 instâncias de usuários classificados como insatisfeitos. Em seguida, o conjunto foi dividido em conjunto de treino e teste, com 70% das instâncias destinadas para treino (21.690) e o restante para teste (9.296). As variáveis quantitativas foram normalizadas através do método *minmaxscaler*, que transforma os dados de uma variável para o intervalo fechado de [0,1], enquanto as variáveis qualitativas categóricas foram binarizadas.

Na sequência, a tarefa de classificação foi realizada no conjunto de dados resultante. Devido ao desbalanceamento na contagem de instâncias da variável classe, foram aplicadas técnicas de balanceamento por *undersampling*, *oversampling* e *resample*. Também houve a análise dos hiperparâmetros, obtendo-se os melhores parâmetros a partir do método *Grid-Search* de cada modelo. Os modelos utilizados no estudo foram: KNN, Floresta Aleatória, *Multilayer Perceptron* (MLP), Regressão Logística, *Naive-Bayes* e *XGBoost*. Além disso, também foram analisadas as variáveis mais importantes para a classificação da satisfação dos usuários dos modelos a partir do ranqueamento de atributos pela sua importância (*feature importances*).

Uma análise alternativa considerando como usuários insatisfeitos os passageiros que atribuíram nota menor ou igual a 2 e satisfeitos os que determinaram nota maior ou igual a 4 para a variável ‘Satisfação Geral’.

Por fim, realizou-se uma investigação de cada modelo treinado no intuito de classificar usuários que determinaram a nota 3 para a variável “Satisfação Geral”. Essa análise ocorreu com o objetivo de averiguar se os usuários com avaliação intermediária poderiam ser de fato classificados como ‘Insatisfeitos’. A sequência do estudo pode ser observada no fluxograma da figura 8.

Figura 8 – Fluxograma das atividades realizadas no estudo



Fonte: Autoria Própria (2023).

3.3 Ferramentas Utilizadas

O estudo foi realizado utilizando a linguagem *Python 3* através do ambiente de desenvolvimento integrado (IDE) *Jupyter Notebook* e *distribuição Anaconda* (disponível em <https://www.anaconda.com/>). Foram utilizadas bibliotecas para manipulação e visualização de dados, bem como bibliotecas específicas para *machine learning*. Com relação à manipulação de dados, foram utilizadas as bibliotecas *pandas* e *numpy* (disponíveis em <https://pandas.pydata.org/> e <https://numpy.org/>, respectivamente), enquanto que para a visualização de dados, foram empregadas as bibliotecas *seaborn* (disponível em <https://seaborn.pydata.org/>) e *matplotlib*

(disponível em: <https://matplotlib.org/>). Por fim, as bibliotecas *sklearn* (<https://scikit-learn.org/stable/>), *imblearn* (imbalanced-learn documentation — Version 0.11.0), *mlxtend* (<https://rasbt.github.io/mlxtend/>) e *xgboost* (<https://xgboost.readthedocs.io/en/stable/python/index.html>) foram utilizadas para utilizar e avaliar os modelos de *machine learning*.

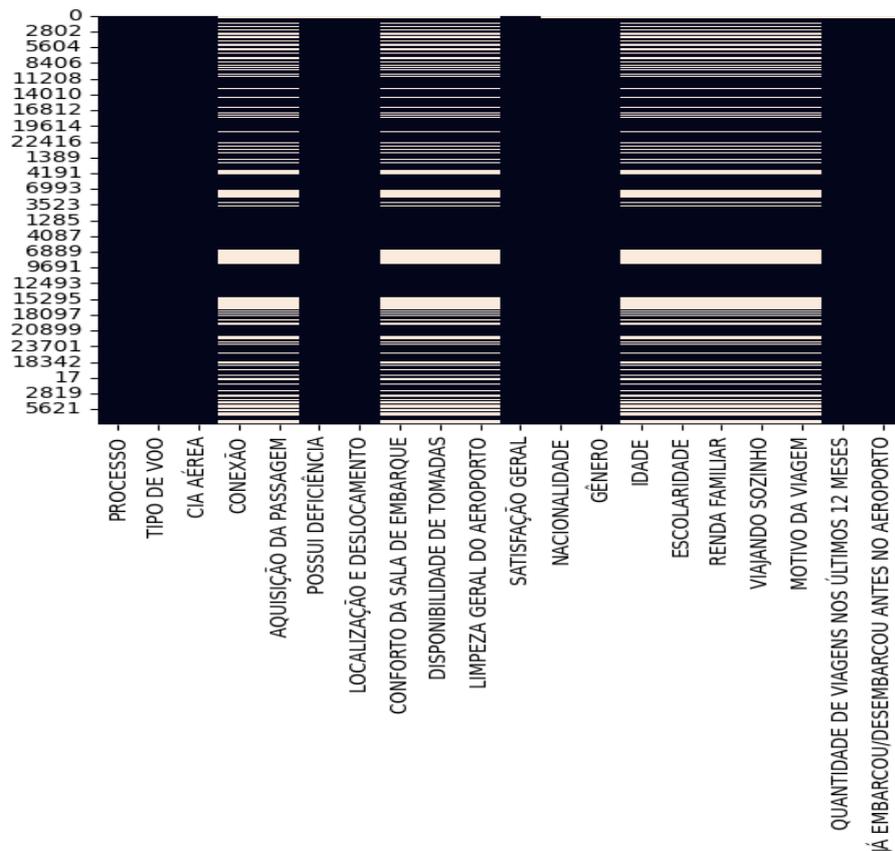
4 RESULTADOS E DISCUSSÕES

Esta seção apresenta os resultados do trabalho, incluindo a tarefa de pré-processamento do conjunto de dados, desempenho de cada modelo para classificação da satisfação dos usuários, bem como as variáveis mais importantes para utilizadas pelos modelos com os melhores desempenhos.

4.1 Análise Descritiva

O conjunto de dados original possuía instâncias com muitos valores nulos, ao ponto que determinadas variáveis tivessem mais de 60% de valores que não são válidos, como pode ser observado no *heatmap* da figura 9.

Figura 9 – Análise de valores nulos de uma fração conjunto de dados original através do *heatmap*



Fonte: Autoria Própria (2023).

Com o tratamento dos valores faltantes realizado conforme descrito na seção 3.2, o conjunto de dados aparece sem nenhuma instância com valores nulos. Para a seleção de variáveis do conjunto de dados, foi utilizado o teste do Chi-Quadrado para variáveis categóricas e a análise do coeficiente de correlação e Pearson para as variáveis numéricas. Dessa forma, as variáveis a serem utilizadas na avaliação dos modelos estão representadas na tabela 1.

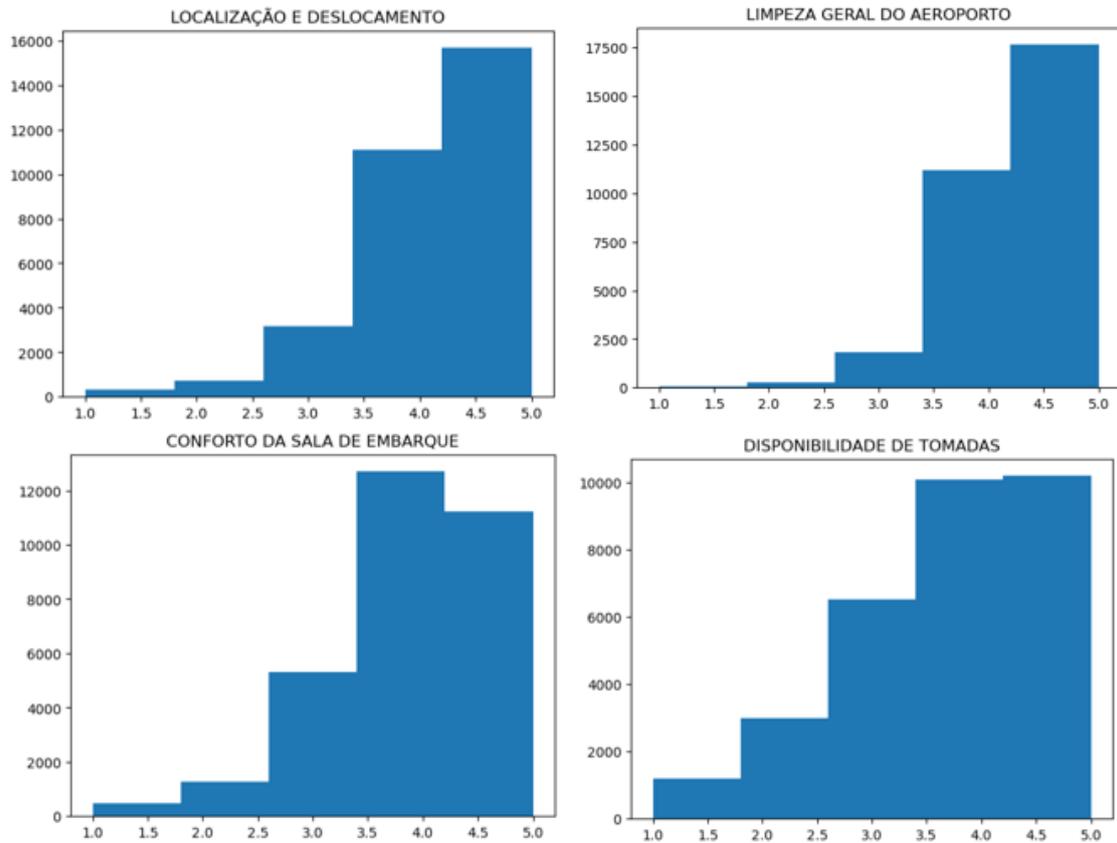
Tabela 1 – Variáveis utilizadas na avaliação dos modelos

Variável	Tipo	Natureza
Conexão	Categórica	Nominal
Localização e Deslocamento	Numérica	Discreta
Conforto da Sala de Embarque	Numérica	Discreta
Disponibilidade de Tomadas	Numérica	Discreta
Limpeza Geral do Aeroporto	Numérica	Discreta
Gênero	Categórica	Nominal
Idade	Categórica	Ordinal
Escolaridade	Categórica	Ordinal
Renda Familiar	Categórica	Ordinal
Viajando Sozinho	Categórica	Nominal
Motivo da Viagem	Categórica	Nominal
Quantidade de Viagens nos Últimos 12 Meses	Categórica	Ordinal
Já Embarcou/Desembarcou Antes no Aeroporto	Categórica	Nominal
Satisfação Geral	Numérica	Binária

Fonte: Autoria Própria (2023).

Com relação às variáveis numéricas, é observado uma grande concentração de valores na faixa de 4 e 5, como é indicado na distribuição por histograma da figura 10. Isso implica que grande parte dos usuários, independente da nota atribuída para a satisfação geral, atribuíram valores de 4 e 5 para as outras variáveis de avaliação dos aeroportos.

Figura 10 – Histograma das variáveis numéricas



Fonte: Autoria Própria (2023).

4.2 Avaliação dos Modelos

No presente estudo, os modelos de classificação testados foram KNN, Floresta Aleatória, *Multilayer Perceptron* (MLP), Regressão Logística, *Naive Bayes - Bernoulli*, *XGBoost* e *Árvore de Decisão*. Para todos os modelos, o conjunto de dados foi separado destinando 70% das instâncias para treino e 30% para teste.

Para avaliação, foi utilizada quatro estratégias de análise com relação ao balanceamento de classes, sendo a avaliação sem nenhum tipo de balanceamento, balanceamento por *undersampling*, *oversampling*, *resample* e pelos melhores parâmetros de cada modelo no conjunto *undersampling* obtidos através do *grid-search*. Com relação ao conjunto de teste, 7635 instâncias são de usuários considerados satisfeitos e 1661 de usuários insatisfeitos.

4.2.1 Sem Balanceamento de Classe

A tabela 2 apresenta as métricas de avaliação de cada modelo sem nenhum tipo de balanceamento de classe da variável alvo 'Satisfação Geral'. Para o conjunto de treino, 17952 instâncias são de usuários considerados satisfeitos e 3738 instâncias são de usuários considerados insatisfeitos.

Tabela 2 – Métricas sem balanceamento de classe

Métricas	KNN		Floresta Aleatória		MLP	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,85	0,63	0,89	0,66	0,90	0,67
Recall	0,98	0,19	0,95	0,46	0,95	0,49
F1-score	0,91	0,29	0,92	0,54	0,92	0,57
Acurácia	0,84		0,86		0,87	
Métricas	Naive Bayes		Regressão Logística		Árvore de Decisão	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,85	0,55	0,94	0,47	0,88	0,46
Recall	0,96	0,24	0,81	0,78	0,88	0,47
F1-score	0,90	0,33	0,87	0,59	0,88	0,47
Acurácia	0,83		0,81		0,81	
Métricas	XGBoost					
	Satisfeito	Insatisfeito				
Precisão	0,88	0,64				
Recall	0,95	0,42				
F1-score	0,91	0,51				
Acurácia	0,85					

Fonte: Autoria própria (2023).

Mesmo com todos os modelos apresentando valores de acurácia maiores que 80%, a métrica *recall* está muito baixa para os usuários insatisfeitos, dessa forma, os modelos apresentam grande dificuldade em identificar usuários insatisfeitos e tem grande tendência para classificar o usuário como satisfeito.

O único modelo que apresentou desempenho superior aos outros foi a Regressão Logística, com valor de acurácia de 81% e recall para usuários insatisfeitos de 78%.

4.2.2 Balanceamento por *Undersampling*

Nesse caso, foram retiradas amostras da instância mais favorecida (satisfeitos), de modo que a quantidade de instâncias de usuários insatisfeitos correspondesse a 90% da quantidade de amostras das instâncias de usuários satisfeitos. Para o conjunto de treino, 4154 instâncias são de usuários considerados satisfeitos e 3738 instâncias são de usuários considerados insatisfeitos. O desempenho de cada modelo é apresentado na tabela 3.

Tabela 3 – Métricas com balanceamento por undersampling

Métricas	KNN		Floresta Aleatória		MLP	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,90	0,34	0,94	0,46	0,94	0,44
Recall	0,73	0,64	0,81	0,76	0,79	0,76
F1-score	0,81	0,44	0,87	0,57	0,86	0,56
Acurácia	0,71		0,80		0,78	
Métricas	Naive Bayes		Regressão Logística		Árvore de Decisão	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,89	0,30	0,95	0,48	0,91	0,36
Recall	0,71	0,58	0,81	0,79	0,73	0,68
F1-score	0,79	0,40	0,87	0,59	0,81	0,47

Tabela 3 – Métricas com balanceamento por undersampling

(Continuação)

Acurácia	0,69	0,81	0,73
Métricas	XGBoost		
	Satisfeito	Insatisfeito	
Precisão	0,92	0,47	
Recall	0,83	0,68	
F1-score	0,87	0,55	
Acurácia	0,80		

Fonte: Autoria própria (2023).

Nessa estratégia, mesmo com os modelos tendo acurácia maior que 80%, a métrica de recall para os usuários insatisfeitos ainda não é satisfatória, salvo para os modelos Regressão Logística, Floresta Aleatória e *Multilayer Perceptron* (MLP), que apresentam valores de recall da amostra menos favorecida com valores maiores que 70%.

4.2.3 Balanceamento por *Oversampling*

Nesse caso, instâncias sintéticas de usuários insatisfeitos foram imputadas até o número de usuários dessa classe representar 80% da quantidade de instâncias dos usuários satisfeitos. Para o conjunto de treino, 17952 instâncias são de usuários considerados satisfeitos e 14362 instâncias são de usuários considerados insatisfeitos. O desempenho dos modelos está representado na tabela 4.

Tabela 4 – Métricas com balanceamento por oversampling

Métricas	KNN		Floresta Aleatória		MLP	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,89	0,35	0,90	0,59	0,92	0,51
Recall	0,78	0,56	0,92	0,55	0,87	0,65
F1-score	0,83	0,43	0,91	0,57	0,89	0,57
Acurácia	0,74		0,85		0,83	
Métricas	Naive Bayes		Regressão Logística		Árvore de Decisão	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,89	0,30	0,94	0,47	0,88	0,46
Recall	0,71	0,58	0,81	0,78	0,88	0,46
F1-score	0,79	0,40	0,87	0,59	0,88	0,46
Acurácia	0,69		0,80		0,81	
Métricas	XGBoost					
	Satisfeito		Insatisfeito			
Precisão	0,91		0,47			
Recall	0,83		0,68			
F1-score	0,87		0,55			
Acurácia	0,80					

Fonte: Autoria própria (2023).

É observado que nenhum modelo possui métricas satisfatórias, com exceção do modelo Regressão Logística, o qual apresenta valores de acurácia acima de 80% e *recall* para usuários insatisfeitos de 78%.

4.2.4 Oversampling por Resample

Nessa estratégia, foram selecionadas e replicadas instâncias de forma aleatória de usuários insatisfeitos de modo que o conjunto de dados ficasse com o mesmo número de instâncias de usuários satisfeitos e insatisfeitos. O resultado das métricas de avaliação de cada modelo é apresentado na tabela 5. Dessa forma, com relação ao conjunto de teste, o mesmo continha 17952 de instâncias tanto para os usuários classificados como insatisfeitos e satisfeitos.

Tabela 5 – Métricas com oversampling por resample

Métricas	KNN		Floresta Aleatória		MLP	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,88	0,37	0,90	0,59	0,93	0,47
Recall	0,82	0,50	0,91	0,55	0,83	0,70
F1-score	0,85	0,42	0,91	0,57	0,88	0,57
Acurácia	0,76		0,85		0,81	
Métricas	Naive Bayes		Regressão Logística		Árvore de Decisão	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,89	0,29	0,95	0,48	0,88	0,46
Recall	0,69	0,60	0,81	0,78	0,88	0,46
F1-score	0,77	0,39	0,87	0,59	0,88	0,46
Acurácia	0,67		0,81		0,81	
Métricas	XGBoost					
	Satisfeito		Insatisfeito			
Precisão	0,88		0,64			
Recall	0,95		0,42			
F1-score	0,91		0,51			
Acurácia	0,85					

Fonte: Autoria própria (2023).

Nesse caso, os modelos apresentam valores de recall muito baixo para usuários insatisfeitos, com exceção do modelo Regressão Logística, o qual apresenta acurácia de 81% e *recall* de 78%.

4.2.5 Melhores parâmetros obtidos pelo *Grid-Search*

Nessa estratégia, utilizando o conjunto de dados obtido pelo método *Undersampling*, foi realizado uma validação cruzada com o objetivo de obter os melhores parâmetros que cada modelo apresenta para maximizar o valor da métrica *f1-score*. Dessa forma, cada modelo foi avaliado com tais valores. O resultado da performance pode ser observado na tabela 6. Os parâmetros e valores testados através do *grid-search* podem ser observados no apêndice B. Os modelos XGBoost e *Multilayer Perceptron* (MLP) não foram utilizados devido ao alto custo computacional.

Tabela 6 – Métricas no conjunto de undersampling com os melhores parâmetros obtidos no grid-search

Métricas	KNN		Floresta Aleatória		Naive Bayes	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,91	0,42	0,94	0,47	0,89	0,30
Recall	0,81	0,63	0,81	0,76	0,71	0,57
F1-score	0,86	0,45	0,87	0,58	0,79	0,40
Acurácia	0,78		0,80		0,69	
Métricas	Regressão Logística		Árvore de Decisão			
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito		
Precisão	0,94	0,48	0,91	0,35		
Recall	0,82	0,78	0,73	0,69		
F1-score	0,88	0,59	0,81	0,47		
Acurácia	0,81		0,72			

Fonte: Autoria própria (2023).

Nesse caso, os modelos de Regressão Logística e Floresta Aleatória apresentam valores de acurácia acima de 80% e de recall para usuários insatisfeitos de 75%.

4.2.6 Avaliação do desempenho dos modelos

No geral, os modelos apresentam altos valores de acurácia, porém valores baixos de *recall* para usuários insatisfeitos (para a maioria dos modelos, a métrica variou de 40% a 65%), a qual são os usuários que representam a classe menos favorecida. Dessa forma, é possível observar que os modelos possuem um viés para classificar os usuários em satisfeitos, o que demonstra que as avaliações de modelos de classificação não devem levar em conta apenas a acurácia, mas sim também o desempenho da classe de interesse.

Por outro lado, o modelo Regressão Logística apresentou desempenho maior que os outros modelos em todas as estratégias realizadas, com acurácia maior que 80% e *recall* para usuários insatisfeitos maior que 75%. O modelo Floresta Aleatória também teve uma melhor performance para o conjunto de dados tratado com *undersampling* e pelos melhores parâmetros através do *Grid-Search*, com desempenho semelhante ao Regressão Logística.

O resultado do modelo Regressão Logística se assemelha com o obtido no trabalho de Mattozo et al (2012), ao qual a métrica de acurácia também foi superior a 80%.

4.3 Variáveis Mais Importantes para Classificação da Satisfação

Levando em consideração que os modelos de melhor desempenho foram a Regressão Logística e Floresta Aleatória, foi realizada uma análise das variáveis mais importantes, para a estratégia em que se utilizou os melhores parâmetros através do *Grid-Search*, para a classificação de satisfação dos usuários.

Para a Regressão Logística, a determinação de importância é feita através do valor dos coeficientes de cada variável na equação da regressão logística, enquanto para a Floresta Aleatória, o método *Feature Importance* busca determinar as variáveis com menor grau de impureza. As cinco principais variáveis obtidas nos modelos e dos

trabalhos de Singh e Dalpatadu (2018) e de Mattozo et al. (2012) estão apresentadas na tabela 7.

Tabela 7– Variáveis mais importantes com os melhores parâmetros obtidos no grid-search e de trabalhos correlatos

Modelo	Regressão Logística	Floresta Aleatória	Singh Dalpatadu (2018)	e Mattozo et al (2012)
Variáveis	Conforto da Sala de Embarque	Conforto da Sala de Embarque	Sinalização Interna	Qualidade Externa
	Limpeza Geral do Aeroporto	Disponibilidade de Tomadas	Limpeza Geral do Aeroporto	Disponibilidade de Assentos
	Disponibilidade de Tomadas	Limpeza Geral do Aeroporto	Exibições de Arte	Preço dos Restaurantes
	Localização e Deslocamento	Localização e Deslocamento	Sinalização Externa	Conforto do Terminal
	Conexão	Gênero	Lojas Internas	

Fonte: Autoria própria (2023).

Observa-se que para ambos os modelos, a variável mais importante foi o conforto da sala de embarque. Além disso, das cinco principais variáveis, quatro delas são semelhantes em ambos os modelos, sendo que a única divergência é que o modelo Regressão Logística possui como quinta variável mais importante a conexão, enquanto a Floresta Aleatória apresenta a variável gênero.

O resultado é condizente com as cinco principais variáveis determinantes para a satisfação do usuário de acordo com o trabalho de Singh e Dalpatadu (2018), a qual a 'limpeza geral do aeroporto' é considerada como a segunda variável mais importante, bem como as variáveis de 'sinalização interna e externa', que são relacionadas a sinalização dentro do aeroporto e facilidade para deslocamento até o mesmo, também são consideradas como fatores importantes. No presente trabalho, essas duas variáveis podem ser comparadas com a variável sinalização e deslocamento.

O estudo de Mattozo et al (2012), a qual relata que a qualidade externa, relacionada a disponibilidade de táxis e tempo de espera para transporte no aeroporto é uma variável determinante para a satisfação dos usuários. Além disso, o mesmo estudo dos autores também traz a variável 'qualidade de assentos e conforto do terminal' determinantes para a insatisfação dos usuários. Essas duas variáveis podem

ser relacionadas as variáveis de 'Conforto da Sala de Embarque' e 'Localização e Deslocamento' do presente trabalho.

4.4 Avaliação do Modelos com Determinação Alternativa de Satisfeitos e Insatisfeitos

Nessa seção, é apresentada uma estratégia alternativa para a classificação de satisfação dos usuários. Nela, foi considerada na análise como usuários insatisfeitos os que determinaram nota 2 ou inferior na variável 'Satisfação Geral', enquanto os usuários considerados satisfeitos foram os que deram nota 4 ou superior.

Os modelos de classificação testados foram KNN, Floresta Aleatória, *Multilayer Perceptron* (MLP), Regressão Logística, *Naive Bayes - Bernoulli*, *XGBoost* e Árvore de Decisão. Para todos os modelos, o conjunto de dados foi separado destinando 70% das instâncias para treino e 30% para teste. Nessa estratégia, o conjunto de teste continha 3672 instâncias de usuários satisfeitos e 192 instâncias de usuários insatisfeitos.

Para avaliação, foram utilizadas quatro estratégias de análise com relação ao balanceamento de classes, sendo a avaliação sem nenhum tipo de balanceamento, balanceamento por *undersampling*, *oversampling*, *resample* e pelos melhores parâmetros de cada modelo no conjunto *undersampling* obtidos através do *grid-search*.

4.4.1 Sem Balanceamento de Classe

A tabela 8 apresenta as métricas de avaliação de cada modelo sem nenhum tipo de balanceamento de classe da variável alvo 'Satisfação Geral'. Com relação ao conjunto de treino, 8572 instâncias são de usuários considerados satisfeitos e 442 de usuários insatisfeitos.

Tabela 8 – Métricas sem balanceamento de classe

Métricas	KNN		Floresta Aleatória		MLP	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,95	0,92	0,97	0,80	0,97	0,62
Recall	1,00	0,06	1,00	0,34	0,99	0,42
F1-score	0,98	0,12	0,98	0,48	0,98	0,50
Acurácia	0,95		0,96		0,96	
Métricas	Naive Bayes		Regressão Logística		Árvore de Decisão	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,97	0,63	0,99	0,23	0,97	0,45
Recall	0,99	0,34	0,86	0,77	0,97	0,39
F1-score	0,98	0,44	0,92	0,35	0,97	0,41
Acurácia	0,96		0,86		0,95	
Métricas	XGBoost					
	Satisfeito		Insatisfeito			
Precisão	0,97		0,69			
Recall	0,99		0,39			
F1-score	0,98		0,49			
Acurácia	0,96					

Fonte: Autoria própria (2023).

Os valores de acurácia de todos os modelos são elevados, porém observa-se que o recall para os usuários insatisfeitos não é satisfatório, tendo seu melhor resultado no modelo Regressão Logística, sendo de aproximadamente 77%, ao ponto que o valor da acurácia no mesmo modelo é de 86%.

4.4.2 Balanceamento por *Undersampling*

Nesse caso, foi retirado amostras da instância mais favorecida (satisfeitos), de modo que a quantidade de instâncias de usuários insatisfeitos correspondesse a 90% da quantidade de amostras das instâncias de usuários satisfeitos. Para o conjunto de treino, 492 instâncias são de usuários considerados satisfeitos e 442 instâncias são

de usuários considerados insatisfeitos. O desempenho de cada modelo é apresentado na tabela 9.

Tabela 9 – Métricas com balanceamento por undersampling

Métricas	KNN		Floresta Aleatória		MLP	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,97	0,14	0,98	0,24	0,98	0,16
Recall	0,82	0,58	0,88	0,74	0,81	0,70
F1-score	0,89	0,23	0,93	0,36	0,89	0,26
Acurácia	0,81		0,87		0,80	
Métricas	Naive Bayes		Regressão Logística		Árvore de Decisão	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,98	0,17	0,99	0,24	0,98	0,15
Recall	0,83	0,64	0,88	0,74	0,78	0,77
F1-score	0,90	0,20	0,93	0,37	0,87	0,26
Acurácia	0,82		0,87		0,78	
Métricas	XGBoost					
	Satisfeito			Insatisfeito		
Precisão	0,98			0,24		
Recall	0,88			0,70		
F1-score	0,93			0,36		
Acurácia	0,88					

Fonte: Autoria própria (2023).

Nessa estratégia, o modelo mais performático foi a Regressão Logística e Floresta Aleatória, com valores de acurácia acima de 85% e de recall acima de 70%. Também foi observado uma melhora na métrica de recall de todos os outros modelos tendo em vista a avaliação sem balanceamento de classe, porém ainda com resultados não satisfatórios.

4.4.3 Balanceamento por *Oversampling*

Nesse caso, instâncias genéricas de usuários insatisfeitos foram imputadas até o número de usuários dessa classe representar 80% da quantidade de instâncias dos usuários satisfeitos. Com relação ao conjunto de treino, 8572 instâncias são de usuários considerados satisfeitos e 6858 de usuários insatisfeitos. O resultado da performance de cada modelo está representado na tabela 10.

Tabela 10 – Métricas com balanceamento por oversampling

Métricas	KNN		Floresta Aleatória		MLP	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,97	0,16	0,97	0,67	0,97	0,42
Recall	0,88	0,41	0,99	0,41	0,97	0,42
F1-score	0,92	0,23	0,98	0,51	0,97	0,42
Acurácia	0,86		0,96		0,94	
Métricas	Naive Bayes		Regressão Logística		Árvore de Decisão	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,98	0,17	0,99	0,23	0,97	0,42
Recall	0,83	0,64	0,87	0,76	0,97	0,41
F1-score	0,90	0,26	0,92	0,35	0,97	0,41
Acurácia	0,82		0,86		0,94	
Métricas	XGBoost					
	Satisfeito		Insatisfeito			
Precisão	0,98		0,24			
Recall	0,88		0,70			
F1-score	0,93		0,36			
Acurácia	0,88					

Fonte: Autoria própria (2023).

Os modelos com melhor performance foram o *XGBoost* e Regressão Logística, ambos com mais de 85% de acurácia e mais de 70% de recall para usuários

insatisfeitos. Observa-se também uma queda de desempenho do modelo Floresta Aleatória em comparação com estratégias anteriores.

4.4.4 *Oversampling* por Resample

Nessa estratégia, foram selecionadas e replicadas instâncias de forma aleatória de usuários insatisfeitos de modo que o conjunto de dados ficasse com o mesmo número de instâncias de usuários satisfeitos e insatisfeitos.

Dessa forma, com relação ao conjunto de teste, o mesmo continha 8572 de instâncias tanto para os usuários classificados como insatisfeitos e satisfeitos. O resultado das métricas de avaliação de cada modelo é apresentado na tabela 11.

Tabela 11 – Métricas com oversampling por resample

Métricas	KNN		Floresta Aleatória		MLP	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,96	0,22	0,97	0,66	0,98	0,36
Recall	0,95	0,29	0,99	0,40	0,95	0,54
F1-score	0,95	0,25	0,98	0,50	0,96	0,43
Acurácia	0,91		0,96		0,93	
Métricas	Naive Bayes		Regressão Logística		Árvore de Decisão	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,98	0,15	0,99	0,23	0,97	0,40
Recall	0,81	0,66	0,87	0,76	0,97	0,39
F1-score	0,88	0,24	0,92	0,35	0,97	0,39
Acurácia	0,80		0,86		0,94	
Métricas	XGBoost					
	Satisfeito			Insatisfeito		
Precisão	0,97			0,69		
Recall	0,99			0,39		
F1-score	0,98			0,49		
Acurácia	0,96					

Fonte: Autoria própria (2023).

Nessa estratégia, ainda que o modelo XGboost apresenta valor de acurácia de 96%, a métrica de recall para a classe de insatisfeitos apresentou um decréscimo com relação a estratégias anteriores. O único modelo que apresenta resultados de acurácia acima de 85% e de recall para a classe de insatisfeitos acima de 75% continua sendo a Regressão Logística.

4.4.5 Melhores Parâmetros Obtidos pelo *Grid-Search*

Nessa estratégia, utilizando o conjunto de dados obtido pelo método *undersampling*, foi realizada uma validação cruzada com o objetivo de obter os melhores parâmetros que cada modelo apresenta para maximizar o valor da métrica *f1-score*. Dessa forma, cada modelo foi avaliado com tais valores. O resultado da performance pode ser observado na tabela 12. Os parâmetros e valores testados através do *grid-search* podem ser observados no apêndice B. Os modelos XGBoost e *Multilayer Perceptron* (MLP) não foram utilizados devido ao alto custo computacional.

Tabela 12 – Métricas no conjunto de undersampling com os melhores parâmetros obtidos no grid-search

Métricas	KNN		Floresta Aleatória		Naive Bayes	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Precisão	0,98	0,23	0,98	0,25	0,98	0,17
Recall	0,90	0,60	0,88	0,73	0,84	0,62
F1-score	0,94	0,34	0,93	0,37	0,91	0,27
Acurácia	0,88		0,88		0,83	
Métricas	Regressão Logística		Árvore de Decisão			
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito		
Precisão	0,99	0,22	0,98	0,15		
Recall	0,86	0,78	0,77	0,76		
F1-score	0,92	0,34	0,86	0,25		
Acurácia	0,85		0,77			

Fonte: Autoria própria (2023).

Ao ponto que o valor de acurácia foi o menor entre todas as estratégias utilizadas em todos os modelos, o *recall* da classe de indivíduos insatisfeitos apresenta uma melhora em todos os modelos, muito em função da escolha dos melhores parâmetros possíveis.

Nessa estratégia, os modelos que apresentam melhor performance foram a Regressão Logística, com valores de acurácia e *recall* para insatisfeitos de 85% e 78%, respectivamente, bem como a Floresta Aleatória, com valores de acurácia e *recall* para insatisfeitos de 88% e 73%.

É importante frisar que, mesmo nos modelos com melhores performances, a precisão para os usuários insatisfeitos não é satisfatória, ou seja, os modelos classificam usuários satisfeitos como insatisfeitos. Para a Regressão Logística e Floresta Aleatória, a precisão para usuários insatisfeitos foi de aproximadamente 23%, o que indica que os modelos não obtiveram sucesso em distinguir tais usuários de grande parte dos usuários satisfeitos.

4.4.6 Avaliação de Desempenho dos Modelos

Embora os valores de acurácia e *recall* para indivíduos insatisfeitos apresentaram uma melhora com relação a estratégia adotada no item 4.2, os modelos ainda apresentam uma dificuldade em determinar os indivíduos insatisfeitos, salvo os modelos Regressão Logística, a qual apresenta métricas com bom desempenho em todas as estratégias, com acurácia acima de 85%, *recall* para usuários insatisfeitos de aproximadamente 75%, e o modelo Floresta Aleatória, que apresenta boa performance com os melhores parâmetros obtidos pelo *Grid-Search* e pelo balanceamento por *undersampling*, com acurácia de em torno de 88% e *recall* para usuários insatisfeitos de aproximadamente 74%.

Os desempenhos dos modelos com melhor performance não trouxeram métricas satisfatórias de precisão para usuários insatisfeitos, sendo que tanto a Regressão Logística quanto a Floresta Aleatória trazem valores em torno de 23%. É importante salientar que, apesar do *recall* desses usuários indicar que esses dois modelos classificaram corretamente 3 a cada 4 usuário insatisfeito, uma grande parte dos usuários satisfeitos também estão sendo classificados como insatisfeitos.

Os resultados do modelo Floresta Aleatória, quando balanceamento por meio do *oversampling* e *oversampling* por *resample*, com acurácia acima de 95% são semelhantes ao apresentado no trabalho de Singh e Dalpatadu (2018), a qual, através da pesquisa de satisfação dos usuários do aeroporto de São Francisco utilizando o modelo Floresta Aleatória para determinar o nível de satisfação, obteve acurácia de 98,5%. É necessário salientar que tal resultado foi possível devido a classificação correta dos usuários satisfeitos, ao ponto que houve uma dificuldade em classificar corretamente os usuários insatisfeitos, vide o *recall* para tais usuários de aproximadamente 40%, indicando dificuldade do modelo em identificar os padrões dos usuários insatisfeitos.

4.5 Análise de Usuários com Satisfação Geral Intermediária

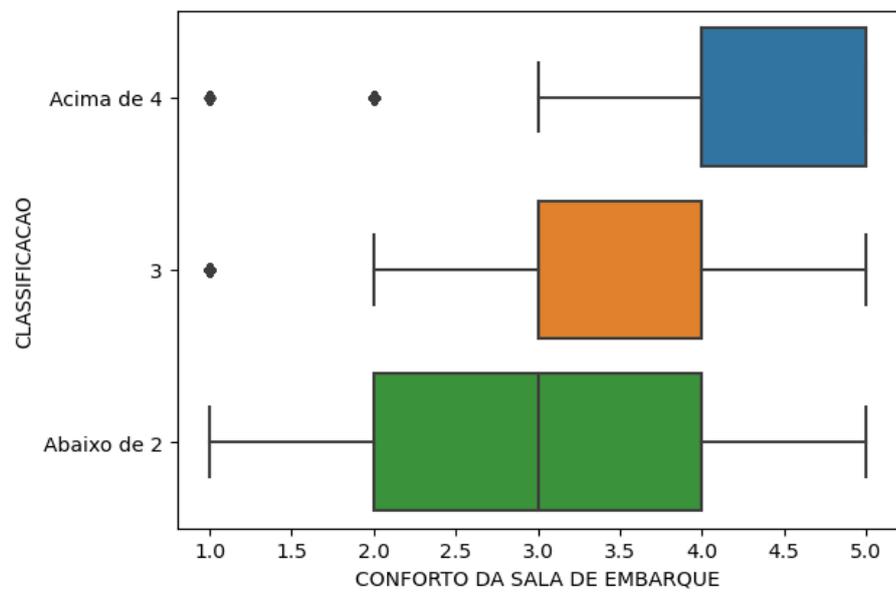
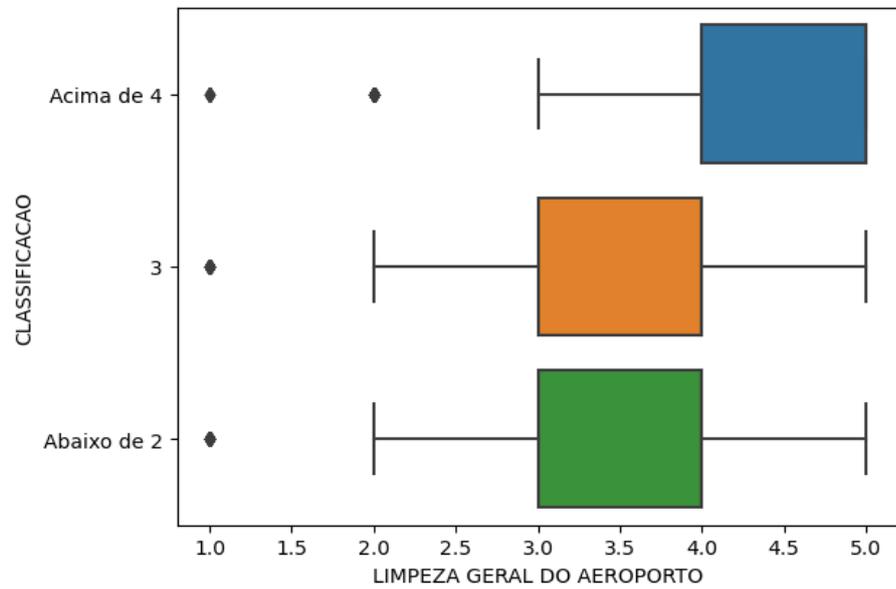
Como os valores da variável 'Satisfação Geral' vão de 1 até 5 e que o estudo considera como insatisfeitos usuários com avaliação de 1 até 3 e satisfeitos os usuários que determinaram notas de 4 e 5, foi investigado como os usuários com avaliação 3 (em teoria, avaliação neutra) se comportam. Para isso, foi realizada uma análise descritiva e avaliado como cada modelo determina o nível de satisfação desses usuários.

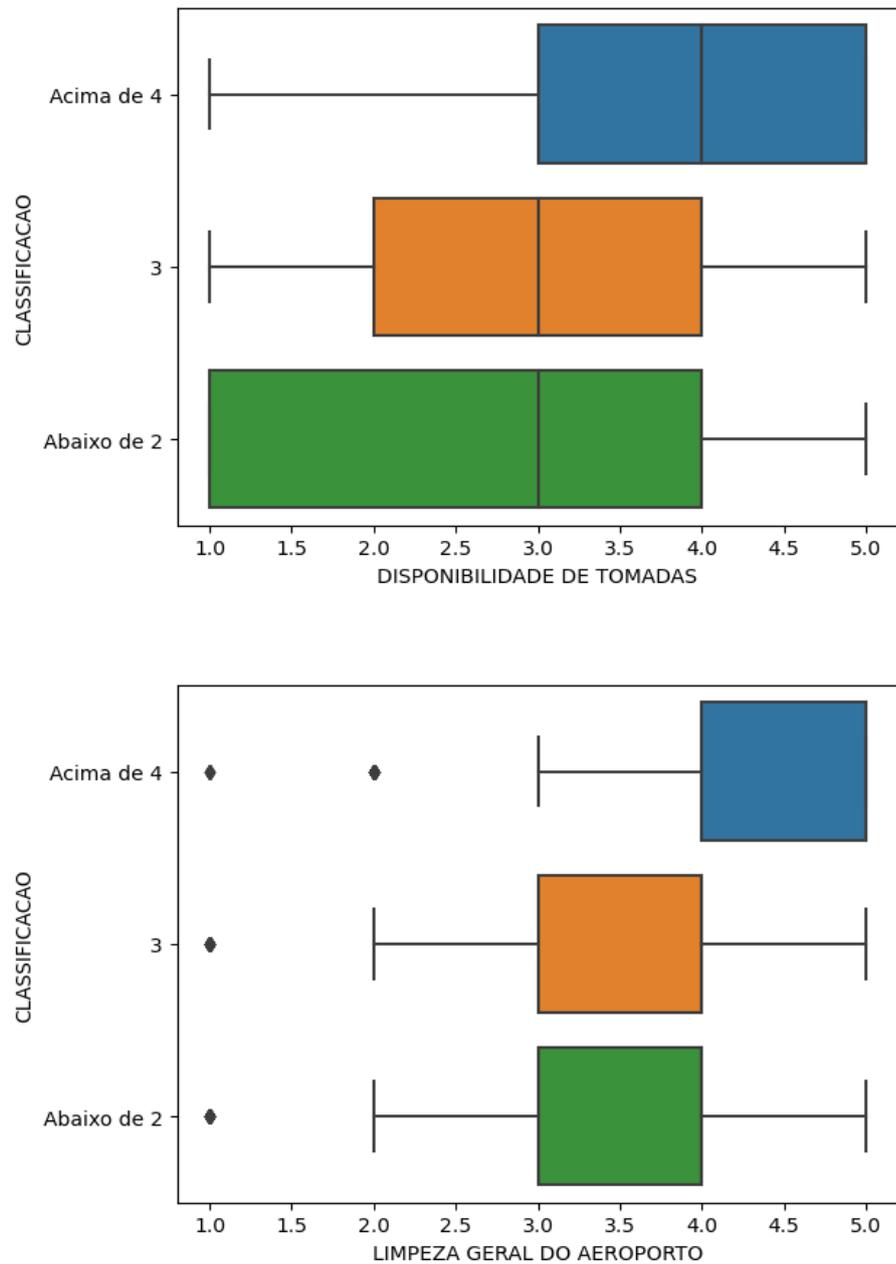
4.5.1 Análise Exploratória dos Usuários com Satisfação Intermediária

Para avaliar como os usuários que determinaram a nota para a variável de 'Satisfação Geral' como 3 se comportam, tal grupo foi comparado com os usuários que deram notas maiores ou iguais a 4 e com usuários com avaliação menores ou iguais a 2. Para isso, foram analisadas as variáveis numéricas de acordo com cada faixa de nota determinada pelos usuários.

Os *boxplots* das variáveis numéricas utilizadas para a avaliação dos modelos estão representados na figura 11.

Figura 11 – Boxplots das variáveis numéricas com relação ao nível de satisfação geral dos usuários





Fonte: Autoria própria (2023).

Através da análise dos *boxplots*, observa-se que os usuários com valor de satisfação geral igual a 3 se comportam majoritariamente como os usuários com avaliação igual ou menor a 2. Porém, uma parte de usuários com nota superior ou igual a 4 também compartilham as mesmas notas que os usuários dos outros grupos.

Nas variáveis 'Limpeza Geral do Aeroporto' e 'Conforto da Sala de Embarque', o primeiro quartil dos usuários com nota superior ou igual a 4 compartilham a mesma avaliação dos usuários dos outros grupos de satisfação geral.

Por outro lado, na variável ‘Disponibilidade de Tomadas’, os usuários do grupo de avaliação geral acima de 4 possuem uma parcela significativa de usuários com notas equivalentes aos usuários das outras categorias.

Com tal similaridade entre os grupos para os valores das variáveis numéricas, os modelos de classificação podem encontrar dificuldade em definir a qual nível de satisfação o usuário pertence.

4.5.2 Avaliação dos Modelos dos Usuários com Satisfação Intermediária

Para a avaliação dos usuários com satisfação intermediária, assim como na avaliação do conjunto de dados completo, os modelos de classificação testados foram KNN, Floresta Aleatória, *Multilayer Perceptron*, Regressão Logística, *Naive Bayes - Bernoulli*, *XGBoost* e Árvore de Decisão. Para tal procedimento, foi considerado que o conjunto de teste será os usuários com valores de satisfação geral igual a 3 e nesse caso são considerados como insatisfeitos. Além disso, a estratégia utilizada será utilizar os melhores parâmetros de cada modelo obtidos anteriormente pelo *grid-search* na seção 4.4.4, ou seja, na estratégia onde os usuários que determinaram nota 3 para a ‘Satisfação Geral’ não foram levados em consideração no conjunto de treino. Os modelos *XGBoost* e *Multilayer Perceptron* (MLP) não foram utilizados devido ao alto custo computacional.

A tabela 13 apresenta a acurácia de cada modelo para a classificação dos usuários com avaliação de ‘Satisfação Geral’ igual a 3.

Tabela 13 – Usuários com Satisfação Geral 3 - Acurácia no conjunto de undersampling com os melhores parâmetros obtidos no grid-search

Métricas	KNN		Floresta Aleatória		Naive Bayes	
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito
Acurácia	0,44		0,66		0,40	
Métricas	Regressão Logística		Árvore de Decisão			
	Satisfeito	Insatisfeito	Satisfeito	Insatisfeito		
Acurácia	0,69		0,62			

Fonte: Autoria própria (2023).

Avaliando as métricas obtidas em todas as estratégias, todos os modelos apresentaram uma dificuldade em classificar os usuários com avaliação de 3 na variável 'Satisfação Geral' como insatisfeitos. Tal fato pode indicar uma região nebulosa, onde não necessariamente os usuários dessa categoria são todos satisfeitos ou insatisfeitos.

5 CONCLUSÃO

O presente estudo buscou avaliar modelos de *machine learning* para a tarefa de classificação com o objetivo de classificar a satisfação geral dos usuários dos aeroportos brasileiros. Para isso, foi utilizada a pesquisa realizada pela Secretaria Nacional de Aviação Civil com dados a partir do ano de 2020 e determinado como a variável alvo do estudo a variável ‘Satisfação Geral’

Para que o trabalho pudesse ser realizado, foi feito um pré-processamento no conjunto de dados, ao ponto que foram excluídas da análise variáveis com mais de 35% de valores nulos, mapeamento de valores similares e seleção de variáveis pelo teste do Qui-Quadrado para variáveis categóricas e pelo teste de correlação de Pearson para variáveis numéricas.

Durante todo o estudo, os modelos testados incluíram os algoritmos KNN, Floresta Aleatória, Naive Bayes, Regressão Logística, *Decision Tree*, MLP e XGBoost, sendo que a estratégia de balanceamento dos dados passou por avaliar o conjunto sem nenhum tipo de balanceamento de classe, balanceamento por *undersampling* e *oversampling*, *resample* e pelos melhores parâmetros de cada modelo obtidos a partir do conjunto de dados da estratégia por *undersampling* visando maximizar a métrica *f1-score*.

Em um primeiro momento, os usuários que determinaram notas de 1,2 e 3 para a variável ‘Satisfação Geral’ foram considerados ‘Insatisfeitos’ e os que determinaram notas de 4 e 5 como ‘Satisfeitos’. Os modelos mais performáticos nessa situação foram os que utilizaram os algoritmos Regressão Logística e Floresta Aleatória, sendo que a Regressão Logística apresentou valores de acurácia acima de 80% e de recall para os usuários insatisfeitos acima de 70% para todas as estratégias de balanceamento adotadas. Com relação ao Floresta Aleatória, o mesmo obteve desempenho semelhante ao Regressão Logística quando utilizado no balanceamento por *undersampling* e com os melhores parâmetros obtidos no *grid-search*.

Das cinco variáveis mais importantes para a classificação da satisfação dos usuários em cada modelo, quatro foram idênticas, sendo elas o ‘Conforto da Sala de Embarque’, ‘Limpeza Geral do Aeroporto’, ‘Disponibilidade de Tomadas’ e ‘Localização e Deslocamento’. Com relação a única variável divergente em ordem de importância entre os dois modelos, a Regressão Logística considerou como

importante a variável 'Conexão', enquanto a Floresta Aleatória considerou a variável 'Gênero'.

Em seguida, foi utilizada uma abordagem em que os usuários 'Insatisfeitos' foram considerados os que determinaram as notas de 1 e 2 para a variável 'Satisfação Geral', enquanto os usuários 'Satisfeitos' foram os que deram notas de 4 e 5. Nessa nova abordagem, os usuários com nota 3 foram excluídos da análise. Como resultado, os modelos que obtiveram as melhores performances utilizaram os algoritmos Regressão Logística, com valores de acurácia acima de 80% e de *recall* para os usuários insatisfeitos acima de 70% em todas as estratégias de balanceamento utilizadas, e a Floresta Aleatória, que obteve desempenho semelhante ao Regressão Logística quando adotou-se a estratégia de balanceamento por *undersampling*.

Tendo em vista que o trabalho considerou os usuários com avaliação intermediária (3) na variável 'Satisfação Geral' como 'Insatisfeitos', essas instâncias foram levadas ao conjunto de teste e avaliadas pelos modelos de classificação com os melhores parâmetros obtidos anteriormente. Como resultado, os modelos classificaram como insatisfeitos por volta de 40% (modelo utilizando o algoritmo *Naïve Bayes*) até 69% (modelo utilizando o algoritmo Regressão Logística) dessas instâncias, o que mostra que não é totalmente confiável determinar que todos os usuários dessa avaliação se comportam como insatisfeitos.

É importante salientar que, com o objetivo de implementar os modelos elaborados no trabalho, é necessário verificar os parâmetros mais importantes para os tomadores de decisão. Caso o objetivo seja identificar usuários insatisfeitos, os recalls obtidos pelos modelos da Regressão Logística podem ser satisfatórios, porém, se o objetivo for classificar todos os usuários corretamente, é necessário um aperfeiçoamento para que a métrica de precisão seja aperfeiçoada.

Como sugestão para futuros trabalhos, é aconselhável a hiperparametrização de mais parâmetros, valores e estratégias, tal como a *Random Search*. Além disso, é possível utilizar variáveis similares da pesquisa realizada pela SAC desde seu início, ao ponto que o estudo utilizou a metodologia aplicada a partir de 2020. A busca por mais instâncias de usuários que possuem como satisfação geral notas de 1 e 2 também é interessante para melhorar a performance dos modelos. Por fim, a análise da satisfação dos usuários considerando as instâncias separadas em um conjunto de dados para cada aeroporto também é válida.

REFERÊNCIAS

ALVES, Allan Carlos *et al.* A QUALIDADE DO SERVIÇO E A SATISFAÇÃO DO CLIENTE. **Revista Pensamento e Realidade**, [s. l.], v. 30, ed. 4, 2015. Disponível em:

<https://revistas.pucsp.br/index.php/pensamentorealidade/article/view/25533/19168>.

Acesso em: 23 maio 2023.

AMARAL, Fernando. *Aprenda Mineração de Dados: Teoria e Prática*. 1. ed. Rio de Janeiro: Alta Books, 2016.

ANAC. Indicadores de Qualidade de Serviço: Anexo 7. **Nota Explicativa**, [s. l.], 2017.

Disponível em: [https://antt-hml.antt.gov.br/documents/359170/1022362/Fator+Q+-+Anexo+7.pdf/87380afa-0940-2094-5aa6-](https://antt-hml.antt.gov.br/documents/359170/1022362/Fator+Q+-+Anexo+7.pdf/87380afa-0940-2094-5aa6-3e5213f54721?t=1593197327627#:~:text=1.3%20O%20Fator%20Q%20%C3%A9,Acidentes%20com%20v%C3%ADtimas%20na%20Rodovia)

[3e5213f54721?t=1593197327627#:~:text=1.3%20O%20Fator%20Q%20%C3%A9,Acidentes%20com%20v%C3%ADtimas%20na%20Rodovia](https://antt-hml.antt.gov.br/documents/359170/1022362/Fator+Q+-+Anexo+7.pdf/87380afa-0940-2094-5aa6-3e5213f54721?t=1593197327627#:~:text=1.3%20O%20Fator%20Q%20%C3%A9,Acidentes%20com%20v%C3%ADtimas%20na%20Rodovia). Acesso em: 8 jun. 2023.

AYUSH, Kalla. K Means Clustering. **DataDrivenInvestor**, [s. l.], 2019. Disponível em:

<https://medium.datadriveninvestor.com/k-means-clustering-4a700d4a4720>. Acesso

em: 22 maio 2023.

BARAKAT, H.; YENITERZI, R.; MARTÍN-DOMINGO, L. Applying deep learning models to twitter data to detect airport service quality. *Journal of Air Transport Management*, [s. l.], v. 91, 2021. DOI <https://doi.org/10.1016/j.jairtraman.2020.102003>.

Disponível em:

<https://www.sciencedirect.com/science/article/pii/S0969699720305846>. Acesso em:

25 set. 2023.

BHASKAR SANKARANARAYANAN, Hari; BV, Vishwanath; RATHOD, Viral. An Exploratory Analysis for Predicting Passenger Satisfaction at Global Hub Airports using Logistic Model Trees. 2016 Second International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), [s. l.], 2016.

Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7813672>.

Acesso em: 24 set. 2023.

BREIMAN, Leo. *Random Forest*. **Statistics Department, University of California, Berkeley**, CA, 2001. Disponível em:

<https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf>. Acesso em: 28 maio 2023.

COSTA, Luís Henrique Gonçalves *et al.* Avaliação da qualidade de serviço dos aeroportos brasileiros concedidos: uma abordagem multicritério aplicando o MACBETH. **Revista Principia**, [s. l.], v. 60, ed. 1, 30 mar. 2023. DOI <http://dx.doi.org/10.18265/1517-0306a2021id6129>. Disponível em: <https://periodicos.ifpb.edu.br/index.php/principia/article/viewFile/6129/1893>. Acesso em: 30 maio 2023.

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P. **Advances in Knowledge Discovery & Data Mining**. California: American Association for Artificial Intelligence, 1996.

FERREIRA, Julio César *et al.* KNOWLEDGE DISCOVERY IN DATABASE E DATA MINING: UMA CONTRIBUIÇÃO BIBLIOMÉTRICA. **XXXVIII ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO**, [s. l.], 2018. Disponível em: https://abepro.org.br/biblioteca/TN_STO_263_509_36492.pdf. Acesso em: 22 maio 2023.

GFBIINFO. Árvores de Decisão. **E-learning**, [s. l.], 2023. Disponível em: <http://web.tecnico.ulisboa.pt/ana.freitas/bioinformatics.ath.cx/bioinformatics.ath.cx/indexf23d.html?id=199>. Acesso em: 9 jun. 2023.

Goldschmidt, Ronaldo. **Data mining : conceitos, técnicas, algoritmos, orientações e aplicações** - 2. ed. - Rio de Janeiro : Elsevier, 2015.

HOMAID, Mohammed Salih; MOULITSAS, Irene. Measuring Airport Service Quality Using Machine Learning Algorithms. **Association for Computing Machinery**, [s. l.], 2022. Disponível em: <https://dl.acm.org/doi/fullHtml/10.1145/3571560.3571562>. Acesso em: 24 set. 2023.

KHEDKAR, Sujata Atul; SHINDE, Subash K. Customer Review Analytics for Business Intelligence. IEEE International Conference on Computational Intelligence and Computing Research, [s. l.], 2018. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8782305>. Acesso em: 25 set. 2023.

KUMAR, Sachin; ZYMBLER, Mikhail. A machine learning approach to analyze customer satisfaction from airline tweets. *Journal of Big Data*, [s. l.], 2019. DOI <https://doi.org/10.1186/s40537-019-0224-1>. Disponível em: <https://link.springer.com/article/10.1186/s40537-019-0224-1>. Acesso em: 25 set. 2023.

- KOONSANIT, Kitti; NISHIUCHI, Nobuyuki. Classification of User Satisfaction Using Facial Expression Recognition and Machine Learning. 2020 IEEE REGION 10 CONFERENCE (TENCON), Osaka, Japão, 2020. DOI 10.1109/TENCON50793.2020.9293912. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9293912>. Acesso em: 25 set. 2023.
- LIOR SINAI. Random Forest classifier from scratch in Python. **Lior Sinai**, [s. l.], 2020. Disponível em: <https://liorsinai.github.io/coding/2020/09/29/random-forests.html>. Acesso em: 28 maio 2023.
- LUCINI, Filipe R. Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. *Journal of Air Transport Management*, [s. l.], v. 83, 2020. DOI <https://doi.org/10.1016/j.jairtraman.2019.101760>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0969699719302959>. Acesso em: 25 set. 2023.
- LUDERMIR, Teresa Bernarda. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. **Estudos Avançados**, [s. l.], 2021. DOI <https://doi.org/10.1590/s0103-4014.2021.35101.007>. Disponível em: <https://www.scielo.br/j/ea/a/wXBdv8yHBV9xHz8qG5RCgZd/>. Acesso em: 21 maio 2023.
- MATTOZO, T.C. *et al.* *Logistic Regression Applied to Airport Customer Satisfaction Using Hierarchical Quality Model*. *Intelligent Data Engineering and Automated Learning - IDEAL 2012*, Berlin, v. 7435, 2012. DOI https://doi.org/10.1007/978-3-642-32639-4_68. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-32639-4_68. Acesso em: 25 set. 2023.
- MAHESH, Batta. Machine Learning Algorithms - A Review. **International Journal of Science and Research (IJSR)**, [s. l.], v. 9, 2020. Disponível em: https://www.researchgate.net/publication/344717762_Machine_Learning_Algorithms_-_A_Review. Acesso em: 22 maio 2023.
- MINISTÉRIO DOS TRANSPORTES. Brasil. Metodologia 2020. Pesquisa de Satisfação e Desempenho Operacional, [s. l.], 2020. Disponível em: https://dados.transportes.gov.br/dataset/e2fdb29b-7b7d-42f9-b1a6-ff3b881ef1ef/resource/00a0e2cc-356c-4b2a-8648-3ea1d11671d8/download/nova-metodologia-da-pesquisa-de-satisfacao_a-partir-de-2020.pdf. Acesso em: 30 maio 2023.

MUNDSTOCK, Elsa *et al.* INTRODUÇÃO À ANÁLISE ESTATÍSTICA: UTILIZANDO O SPSS 13.0. **INSTITUTO DE MATEMÁTICA** : Cadernos de Matemática e Estatística, UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, 2006. Disponível em: http://euler.mat.ufrgs.br/~camey/SPSS/Introdu%20%E3o%20%E0%20An%20E1lise%20Estat%20stica%20utilizando%20o%20SPSS%2013_0.pdf. Acesso em: 8 jun. 2023.

NETTO, Amilcar. **Python para Data Science e Machine Learning:**

Descomplicado. 1. ed. Rio de Janeiro: Alta Books, 2021. 416 p. ISBN 978-65-5520-217-2.

NOGARE, Diego. Performance de Machine Learning – Matriz de Confusão. **INTELIGÊNCIA ARTIFICIAL & MACHINE LEARNING**, [s. l.], 2020. Disponível em: <https://diegonogare.net/2020/04/performance-de-machine-learning-matriz-de-confusao/>. Acesso em: 8 jun. 2023.

OLIVEIRA, Raul Rodrigues de. Distância entre dois pontos. **Mundo Educação**, [s. l.], 2018.

RASCHKA, Sebastian. **Python Machine Learning**: Unlock deeper insights into machine learning with this vital guide to cutting-edge predictive analytics. 1. ed. Reino Unido: Packt Publishing Ltd., 2015. ISBN 978-1-78355-513-0.

SANTOS, Mariana *et al.* SATISFAÇÃO DE CLIENTES: ANÁLISE DO ATENDIMENTO PRESTADO PELAS EMPRESAS NO COMÉRCIO DE SÃO JOÃO DEL REI-MG. **Simpósio de Excelência em Gestão e Tecnologia**, [s. l.], 207. Disponível em: <https://www.aedb.br/seget/arquivos/artigos17/542539.pdf>. Acesso em: 8 jun. 2023.

SCHROFF, Florian. Object Class Segmentation using Random Forest. **Dept. of Engineering Science, University of Oxford**, [s. l.], 2008. Disponível em: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/Criminisi_bmvc2008.pdf. Acesso em: 28 maio 2023.

SINGH, A.K; DALPATADU, Rohan J. Determinants of Customer Satisfaction at the San Francisco International Airport. *Journal of Tourism & Hospitality*, [s. l.], 2019. DOI 10.35248/2167-0269.19.8.398. Disponível em:

https://www.researchgate.net/profile/Rohan-Dalpatadu/publication/335446770_Determinants_of_Customer_Satisfaction_at_the_San_Francisco_International_Airport/links/5dd3297b92851c382f49acaa/Determinants-of-Customer-Satisfaction-at-the-San-Francisco-International-Airport.pdf. Acesso em: 24 set. 2023.

SILVA, Leandro Augusto da. **Introdução à Mineração de Dados - Com Aplicações em R**. 1. ed. Rio de Janeiro: Elsevier, 2016. 259 p. ISBN 978-85-352-8446-1.

VAN DER MAATEN, Laurens *et al.* Dimensionality Reduction: A Comparative Review. **Journal of Machine Learning Research**, [s. l.], 2007. Disponível em: https://www.researchgate.net/publication/228657549_Dimensionality_Reduction_A_Comparative_Review. Acesso em: 11 jun. 2023.

APÊNCIDE A

Tabela 14 – Natureza das variáveis do conjunto de dados

Variável	Tipo	Natureza
Processo	Categórica	Nominal
Aeroporto	Categórica	Nominal
Data	Numérica	Discreta
Mês	Categórica	Nominal
Início Coleta	Numérica	Discreta
Fim Coleta	Numérica	Discreta
Terminal	Numérica	Discreta
Portão	Numérica	Discreta
Tipo De Voo	Categórica	Nominal
Cia Aérea	Categórica	Nominal
Voo	Categórica	Nominal
Conexão	Categórica	Nominal
Aquisição Da Passagem	Categórica	Nominal
Meio De Aquisição Da Passagem	Categórica	Nominal
Meio De Transporte Para O Aeroporto	Categórica	Nominal
Possui Deficiência	Categórica	Nominal
Utiliza Recurso Assistivo	Categórica	Nominal
Solicitou Assistência Especial	Categórica	Nominal
Forma De Desembarque Utilizada	Categórica	Nominal
Avaliação Do Método De Desembarque	Numérica	Discreta
Utilizou O Estacionamento	Categórica	Nominal
Facilidade De Desembarque No Meio-Fio	Numérica	Discreta
Opções De Transporte Até O Aeroporto	Numérica	Discreta
Forma De Check-In	Categórica	Nominal

Processo De Check In	Numérica	Discreta
Tempo De Espera Na Fila	Numérica	Discreta
Organização Das Filas	Numérica	Discreta
Quantidade De Totens	Numérica	Discreta
Quantidade De Balcões	Numérica	Discreta
Cordialidade Dos Funcionários	Numérica	Discreta
Tempo De Atendimento	Numérica	Discreta
Processo De Aquisição Da Passagem	Numérica	Discreta
Atendimento Da Cia. Aérea	Numérica	Discreta
Processo De Inspeção De Segurança	Numérica	Discreta
Tempo De Espera Em Fila	Numérica	Discreta
Atendimento Dos Funcionários	Numérica	Discreta
Controle Migratório	Numérica	Discreta
Quantidade De Guichês	Numérica	Discreta
Controle Aduaneiro	Numérica	Discreta
Estabelecimentos De Alimentação	Numérica	Discreta
Quantidade De Estabelecimentos De Alimentação	Numérica	Discreta
Qualidade E Variedade De Opções De Estabelecimentos De Alimentação	Numérica	Discreta
Relação Preço X Qualidade Dos Estabelecimentos De Alimentação	Numérica	Discreta
Estabelecimentos Comerciais	Numérica	Discreta
Quantidade De Estabelecimentos Comerciais	Numérica	Discreta
Qualidade E Variedade De Opções De Estabelecimentos Comerciais	Numérica	Discreta
Relação Preço X Qualidade Dos Estabelecimentos Comerciais	Numérica	Discreta
Estacionamento	Numérica	Discreta
Qualidade Das Instalações De Estacionamento	Numérica	Discreta
Facilidade Para Encontrar Vagas	Numérica	Discreta

Facilidade De Acesso Ao Terminal	Numérica	Discreta
Relação Custo X Benefício	Numérica	Discreta
Localização E Deslocamento	Numérica	Discreta
Sinalização	Numérica	Discreta
Disponibilidade De Painéis De Informações De Voo	Numérica	Discreta
Acessibilidade Do Terminal	Numérica	Discreta
Conforto Da Sala De Embarque	Numérica	Discreta
Conforto Térmico	Numérica	Discreta
Conforto Acústico	Numérica	Discreta
Disponibilidade De Assentos	Numérica	Discreta
Disponibilidade De Assentos Reservados	Numérica	Discreta
Disponibilidade De Tomadas	Numérica	Discreta
Internet Disponibilizada Pelo Aeroporto	Numérica	Discreta
Velocidade De Conexão	Numérica	Discreta
Facilidade De Acesso À Rede	Numérica	Discreta
Sanitários	Numérica	Discreta
Quantidade De Banheiros	Numérica	Discreta
Limpeza Dos Banheiros	Numérica	Discreta
Manutenção Geral Dos Sanitários	Numérica	Discreta
Limpeza Geral Do Aeroporto	Numérica	Discreta
Processo De Restituição De Bagagens	Numérica	Discreta
Facilidade De Identificação Da Esteira De Restituição	Numérica	Discreta
Tempo De Restituição	Numérica	Discreta
Integridade Da Bagagem	Numérica	Discreta
Satisfação Geral	Numérica	Discreta
Motivo	Categórica	Nominal
Nacionalidade	Categórica	Nominal

Gênero	Categórica	Nominal
Idade	Categórica	Ordinal
Escolaridade	Categórica	Ordinal
Renda Familiar	Categórica	Ordinal
Viajando Sozinho	Categórica	Nominal
Número De Acompanhantes	Categórica	Ordinal
Motivo Da Viagem	Categórica	Nominal
Quantidade De Viagens Nos Últimos 12 Meses	Categórica	Ordinal
Já Embarcou/Desembarcou Antes No Aeroporto	Categórica	Nominal
Antecedência	Categórica	Ordinal
Tempo De Espera	Categórica	Ordinal
Comentários Adicionais"	Categórica	Nominal

Fonte: Autoria Própria (2023).

APÊNDICE B

Tabela 15 – Parâmetros utilizados no Grid-Search para cada algoritmo

Algoritmo	Parâmetro	Valores
Regressão Logística	C	0.1,1,10
	Penalty	l1,l2
	Solver	liblinear,saga
	Max_iter	50,100,250,500
Floresta Aleatória	N_estimators	10,300
	Max_features	sqrt,log2
	Max_depth	50,100,250,500
	Class_weight	balanced,balanced_subsample
Árvore de Decisão	Criterion	gini,entropy,log_loss
	Splitter	best,random
	Max_depth	100,200,300
	Max_features	Auto,sqrt,log2
Naive-Bayes	alpha	10,20,100
	fit_prior	True,False
KNN	N_neighbors	10,50,100
	Weights	uniform,distance
	Algorithm	auto,ball_tree,kd_tree,brute