

Article

Exploring Net Promoter Score with Machine Learning and Explainable Artificial Intelligence: Evidence from Brazilian Broadband Services

Matheus Raphael Elero ¹, Rafael Henrique Palma Lima ², Bruno Samways dos Santos ^{2,*}
and Gislaïne Camila Lapasini Leal ¹

¹ Graduate Program in Production Engineering, State University of Maringá (UEM), Maringá 87020-900, Brazil; matheuselero1@gmail.com (M.R.E.); gclleal@uem.br (G.C.L.L.)

² Department of Production Engineering, Federal University of Technology Paraná (UTFPR), Londrina 86036-700, Brazil; rafaelhlma@utfpr.edu.br

* Correspondence: brunosantos@utfpr.edu.br; Tel.: +55-42-99900-3605

Abstract

Despite the growing use of machine learning (ML) for analyzing service quality and customer satisfaction, empirical studies based on Brazilian broadband telecommunications data remain scarce. This is especially true for those who leverage publicly available nationwide datasets. To address this gap, this study investigates customer satisfaction with broadband internet services in Brazil using supervised ML and explainable artificial intelligence (XAI) techniques applied to survey data collected by ANATEL between 2017 and 2020. Customer satisfaction was operationalized using the Net Promoter Score (NPS) reference scale, and three modifications in the scale were evaluated: (i) a binary model grouping ratings ≥ 8 as satisfied and ≤ 7 as dissatisfied (portion of the neutrals as satisfied and another as dissatisfied); (ii) a binary model excluding neutral responses (ratings 7–8) and retaining only detractors (≤ 6) and promoters (≥ 9); and (iii) a multiclass model following the original NPS categories (detractors, neutrals, and promoters). Nine ML classifiers were trained and validated on tabular data for each formulation. Model interpretability was addressed through SHAP and feature importance analysis using tree-based models. The results indicate that Histogram Gradient Boosting and Random Forest achieve the most robust and stable performance, particularly in binary classification scenarios. The analysis of neutral customers reveals classification ambiguity, showing scores of “7” tend toward dissatisfaction, while scores of “8” tend toward satisfaction. XAI analyses consistently identify browsing speed, billing accuracy, fulfillment of advertised service conditions, and connection stability as the most influential predictors of satisfaction. By combining predictive performance with model transparency, this study provides computational evidence for explainable satisfaction modeling and highlights the value of public regulatory datasets for reproducible ML research.

Keywords: machine learning; customer satisfaction; service quality; NPS; SHAP



Academic Editors: Paolo Bellavista
and Fernando Bobillo

Received: 2 October 2025

Revised: 13 January 2026

Accepted: 14 January 2026

Published: 2 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The scientific literature describes how Machine Learning (ML) assists service quality management. Research identifies several ways these algorithms enhance service provider management, including the development of predictive customer satisfaction systems and the analysis of key satisfaction features [1–3].

This diversity of applications arises from the characteristics of the service industry itself, which can be subdivided into transportation, communications, hospitality, and healthcare, among others [4]. Each of these sectors has its own specific management methods, operations, and end products. However, all service providers share a common objective: meeting consumer expectations through effective planning, failure analysis, and quality improvement [5].

Unlike a physical product, whose quality can be measured concretely, service quality is harder to assess because it is intangible [6]. To address this, Altuntas and Kansu [5] suggest managing service quality by aligning expectations with perceived value, planning effectively, and reducing failures. Further, the literature shows customer satisfaction is linked to higher retention and loyalty [7,8].

This has led companies and regulatory agencies to develop methods for obtaining customer satisfaction data. Commonly, questionnaires such as SERVQUAL measure quality via tangibility, reliability, responsiveness, assurance, and empathy [9,10]. The Net Promoter Score (NPS) measures satisfaction on a 0–10 scale and classifies customers as detractors, neutrals, or promoters [11,12]. Sentiment analysis of online reviews [13] and social media [14] offers alternative ways to assess satisfaction.

Recent literature demonstrates the wide range of applications of machine learning models focused on customer satisfaction and service quality. For example, studies can be found in the fields of tourism and hospitality [15–17], financial institutions [18–21], aviation [22–25], product design [26–28], e-commerce [29–32], and distribution logistics [33,34], with the objective of predicting churn, complaints, analyzing sentiment in reviews, and understanding factors related to customer retention and loyalty. In the telecommunications field, works such as [35–45] are the most recent found in the literature and will be described later.

The telecommunications sector encompasses internet service providers (ISPs) involved in fixed-line telephony, mobile telephony, broadband, and pay-TV services [46]. The academic literature on the telecommunications sector covers articles applied to streaming video transmission quality, customer retention, complaints predictions, analysis of the most relevant features, construction of automated systems for customer satisfaction classification and prediction, among others [47–50].

Despite abundant international literature, few Brazilian publications focus on using ML to study service quality. Rožman et al. [51] and Grebovic et al. [52] argue that ML models outperform traditional statistical models on nonlinear datasets due to higher accuracy. Some Brazilian applications exist, such as Sibert et al. [2] on predictive customer satisfaction in energy, and Lucini et al. [25] on ML-driven analyses of air transport customer satisfaction. These examples illustrate ML's potential in enhancing Brazilian service quality management.

Despite the limited number of national publications, Brazil has been developing databases that could be explored. The National Telecommunications Agency (ANATEL) conducts an annual nationwide survey to gather information on customer satisfaction indices for mobile telephony and broadband services [46]. The resulting database from this survey is publicly available, but few scientific works have used this dataset for research. In the literature, two studies explore these data, focusing on customer satisfaction with telephony services; however, neither addresses customer satisfaction with broadband services [53,54].

Considering the challenges and importance of service quality, the potential benefits that machine learning (ML) techniques can offer to this domain, the scarcity of publications in Brazil, and the availability of data from ANATEL, this research presents a novel approach that applies machine learning and explainable artificial intelligence techniques

to analyze and interpret models related to customer satisfaction with internet services in Brazil. Specifically, ML techniques were employed to classify customer satisfaction, and the most relevant features influencing the outcomes were subsequently investigated using the explainable artificial intelligence method SHAP (SHapley Additive eXplanations), in conjunction with algorithms such as Random Forest (RF), Decision Tree (DT), and Gradient Boosting (GB).

Accordingly, this article makes three main contributions:

1. The exploration of ANATEL data on the quality of broadband internet service providers, which are publicly available at the national level, is released periodically, and based on reliable measurements, yet remains underexplored in the literature.
2. Application of ML algorithms to predict (in this case, classify) customer satisfaction, assessing the contribution of variables to deepen the understanding of the most relevant predictors through strategies embedded in tree-based algorithms, as well as SHAP, which provides a global interpretation of the models. This enhances model transparency and interpretability [55], representing a strong alternative to traditional statistical approaches.
3. Using the Net Promoter Score (NPS) as a reference value to operationalize customer satisfaction monitoring, as noted by [56], enhances the applicability of classification algorithms and the interpretability of the models. Additionally, the NPS categories were adapted to examine model training behavior and the predictive power of the algorithms under different categorization schemes.

The remainder of this article is organized into six additional sections. Section 2 (Theoretical Background) discusses the concepts of customer satisfaction, NPS, ML algorithms, and explainable artificial intelligence. Section 3 (Literature Review) presents prior studies on the application of ML across various service sectors. Section 4 (Materials and Methods) details the entire research process, from data characteristics and extraction to the application and analysis of the techniques. Section 5 (Results) reports the performance of the classification techniques and identifies the most relevant features. Section 6 (Discussion) provides a synthesis of the findings. Finally, Section 7 (Conclusions) presents the conclusions derived from the data and research objectives, along with directions for future research.

2. Theoretical Background

2.1. Customer Satisfaction and the Net Promoter Score (NPS)

Customer satisfaction can be considered a determining factor in assessing whether a product or service meets customer needs, reflecting the status of a business [57]. It has been systematically studied since 1989, particularly through the Customer Satisfaction Barometer [58]. Since then, customer satisfaction has been measured across various fields of knowledge, such as total quality management, economic globalization, and strategic management, thereby permeating the entire organization [58,59]. At the same time, customer satisfaction can be measured as a variable that influences brand loyalty or trust [60], making it an essential tool for marketing strategies [61].

However, tracking customer satisfaction is not a trivial task, as it requires organizational resources to remain up to date with customer perceptions. Several methods are employed for this purpose, commonly in the form of on-site or electronic surveys that collect customer feedback. This process involves data collection, processing, and analysis to support managerial decision-making [62].

In such studies, the SERVQUAL and SERVPERF models are widely used. The former measures the gap between customer expectations and perceptions of service performance, while the latter provides a direct measure of perceived service quality [63]. Baquero [64] also discusses the Customer Satisfaction Score (CSAT), Net Promoter Score (NPS), and

Customer Effort Scale (CES), whereas Patti et al. [65] refer to the Service Experience Scale (EXQ) and the American Customer Satisfaction Index (ACSI). Apart from NPS and CES, all these methods rely on multiple evaluation items and require mathematical aggregation to produce a final score or index.

In the present study, the Net Promoter Score (NPS) was adopted, as it is based on a single question: “How likely are you to recommend our company to a friend or colleague?”. Respondents rate this question on a scale from 0 to 10, being classified as “detractors” if they assign scores from 0 to 6, “neutrals” if they select 7 or 8, and “promoters” if they choose 9 or 10. For a comprehensive analysis of customer responses, the NPS index is calculated as a percentage by subtracting the proportion of detractors from the proportion of promoters, dividing by the total number of respondents, and excluding “neutral” respondents [64,66].

It is noteworthy that NPS is not a direct measure of satisfaction, but rather of customer loyalty [49]. Nevertheless, the literature suggests that customer satisfaction, loyalty, and business profitability are associated, although a clear causal relationship has not yet been established [67]. Complementarily, some studies indicate that NPS primarily measures customer loyalty, with satisfaction emerging as a consequence of loyalty [64]. In this context, treating the NPS index as a proxy for customer satisfaction appears to be a practical approach, as it facilitates categorizing ANATEL survey respondents into “detractors,” “neutral,” and “promoters” based on the score directly attributed to overall satisfaction with internet services.

2.2. Machine Learning and Explainable Artificial Intelligence

Machine learning (ML) is a subfield of artificial intelligence in which systems learn specific tasks from training datasets and can make predictions on previously unseen test data [68]. In classification tasks, the algorithm is trained on a dataset composed of multiple input features (predictor variables) and a known label y (a categorical variable), with the objective of classifying new instances in a test set whose labels are unknown [69].

Several algorithms can be applied to classification tasks. In this study, nine classification algorithms were trained, and their hyperparameters were validated. These algorithms were selected due to their strong performance on tabular datasets. The algorithms used, as well as details of the training and validation process, are presented in Section 4.

However, many ML algorithms are designed solely for prediction and frequently lack interpretability regarding how input variables affect model outcomes; such models are termed “black boxes” [70]. In response, Explainable Artificial Intelligence (XAI) methods have been developed to enhance the transparency and interpretability of analytical processes, increasing both trust in and usability of AI-based systems [70,71]. This subfield of artificial intelligence centers on clarifying the reasoning behind system decisions [72].

Tree-induction-based methods, such as Decision Trees and Random Forests, inherently provide mechanisms to rank variable importance during model construction [69]. Alternatively, model-agnostic approaches offer interpretability regardless of the underlying ML model. Examples of these include Anchors, Prediction Difference Analysis, SHAP, Break-Down, and Local Interpretable Model-Agnostic Explanations (LIME), among others [73].

SHAP, the method used in this study, is based on cooperative game theory. It produces “SHAP values,” which quantify each feature’s contribution to a specific instance, and the predictive model thus yields an output for each observation [74]. The main advantage of SHAP is its clear, graphical interpretation of results and feature contributions; however, its high computational cost makes it unsuitable for real-time systems [70].

3. Literature Review

The earliest relevant studies applying machine learning (ML) to customer satisfaction in telecommunications date back to 1995 [75] and 1996 [76]. The former was conducted in partnership with the telecommunications company US WEST and the University of Colorado, using the KnowledgeSEEKER software (version 3.0) to analyze customer data, with a focus on classifying service quality through tree-based techniques. The second study is closely related to the first but places greater emphasis on decision-tree-based strategies aimed at cost reduction. In this case, the most relevant factors were modified due to issues related to the company's competitive environment.

As previously noted, numerous studies have been published on ML applications focusing on customers in the telecommunications sector. Pacheco et al. [35] proposed an approach to internet traffic classification using ML and deep learning models to improve Quality of Service (QoS). This application was more technical in nature, addressing internet protocols and satellite communications, and classified traffic using algorithms such as k-Nearest Neighbors (KNN), Decision Tree (DT), Extra Trees, Ensemble Voting, and AdaBoost, with automatic feature extraction via one- and two-dimensional convolutional neural networks (1-D and 2-D CNNs).

Some studies explicitly used customer satisfaction as the target variable. For example, Wassouf et al. [36] adopted a predictive big data approach to analyze customer loyalty in Syria, focusing on the Syriatel Telecom Company. Based on 127 million records, users were segmented into groups, and loyalty was measured for each segment. Classifiers such as Random Forest (RF), DT, Gradient-Boosted Machine (GBM), and Artificial Neural Network (ANN) were employed. A total of 223 features—covering customer segments, individual behavior, social and spatial aspects, and service and contract types—were used. The binary classification achieved accuracies of up to 87% and a precision of 76%, while multiclass models yielded lower accuracy.

Internet speed was the focus of Namli et al. [37], who proposed a framework using ANN classifiers (multilayer perceptron and radial basis function network types) for initial problem detection, followed by clustering with Kohonen networks and Adaptive Resonance Theory to identify issues, and concluding with a bagging ensemble for both classification and clustering. Principal Component Analysis was applied to reduce dimensionality in a dataset originally containing over 33,000 instances and 270 variables from a call center, and class-balancing strategies were used, achieving accuracies of up to 92%. Espadinha-Cruz et al. [38] proposed a methodology to improve lead management efficiency and increase customer acquisition in a telecommunications company. Using a dataset of more than 546,000 records and 18 variables, each instance was classified into probability ranges of potential customer conversion using ANN, DT, GBM, and ensemble-based algorithms. Ensemble models showed errors below 0.17.

YahYah et al. [39] classified customer trouble tickets in a telecommunications company to improve operational efficiency. The study was conducted across several company units in Malaysia and used data including creation date, speed, type, status, internet signal strength, configuration, response time, and customer characteristics. Natural Language Processing (NLP), SQA-based data selection, and data mining techniques were applied. Average accuracies of 75.35% were achieved using single-machine boosting ensemble models, while a Hadoop-based DT system reached an average accuracy of 84.86%. Rintyarna et al. [40] analyzed Twitter data to understand customer perspectives on internet service quality during the COVID-19 pandemic. Using sentiment analysis and Naïve Bayes (NB) classifiers on datasets from two Indonesian companies, the proposed framework achieved accuracy slightly above 80%, recall of approximately 90%, and precision around 75%, with positive and neutral sentiments predominating.

Haq et al. [41] investigated factors affecting 3G and 4G telecommunication services in Punjab, Pakistan, analyzing network coverage, customer service, video calls, and download speed using descriptive statistics, correlation, and regression analysis. With data from 300 respondents, results indicated that internet speed was the most important factor for increasing company revenues in the region. Saleh and Saha [42] examined customer retention through churn analysis using four datasets, including one collected in person at Aalborg University. Models such as RF, AdaBoost, Logistic Regression (LR), DT, and Extreme GBM were applied. RF, AdaBoost, and XGBoost achieved the best performance on this dataset, with accuracies above 90%. Feature importance analysis highlighted package upgrades as key for the Aalborg dataset, while contract type, monthly charges, and service stability were more relevant in other datasets.

Afzal et al. [43] applied classification algorithms to predict churn in a telecommunications service company using RF, KNN, DT, and XGBoost, achieving test accuracies of up to 98.25%. Variables included internet service types, device protection plans, online security and backup, contract types, and demographic factors. However, a limitation of this study is its reliance on a single, well-known dataset available on the Kaggle platform. Chajia and Nfaoui [44] also addressed churn prediction, leveraging large language model embeddings and classifiers such as LR, SVM, RF, KNN, ANN, NB, DT, and zero-shot classification. Using a Kaggle dataset with nearly 37,000 customers, including demographic, spatial, customer category, preferred offers, operations, internet options, discounts, and complaints, the authors found that embeddings from OpenAI combined with LR achieved 89% accuracy and an average F1-score of 88%. Finally, AbdelAziz et al. [45] analyzed churn data using ML and deep learning models applied to Customer Relationship Management (CRM) data from insurance, internet service provider, and telecommunications companies. Using XGBoost, CNN, and deep learning ensembles, accuracies of up to 95.96% for insurance and 98.42% for telecommunications were achieved.

There is also a substantial body of literature addressing service quality and customer satisfaction in telecommunications. Cox and Bell and Cox et al. [75,76] used Decision Trees to classify mobile telephony customer satisfaction based on perceived quality questionnaires and demonstrated how DTs can extract the most relevant features influencing satisfaction.

Specifically, regarding the use of the Net Promoter Score (NPS) in telecommunications, authors such as Tong et al. [49] and Markoulidakis et al. [77] used questionnaire data and treated NPS as the output variable in ML applications. Tong et al. [49] applied XGBoost and DT models to consumer information, pricing, usage profiles, and behavior, finding that average call duration over the previous three months, average 2G internet traffic, and Average Revenue Per User (ARPU) were the most important variables for identifying customer loyalty. They also applied k-means clustering to segment customers into six groups aligned with business needs. In contrast, Markoulidakis et al. [77] conducted extensive ML experiments to analyze the association between NPS and customer experience (CX). DT, KNN, SVM, RF, ANN, CNN, NB, and LR were applied to classify 10,800 surveys across two dataset versions: one with nine CX-related attributes and the class label, and another including an additional NPS bias variable. NB and CNN achieved the best F1-scores in the first version, while RF, ANN, and CNN performed best overall in the second version, reaching accuracies and precision of approximately 80%.

Suchanek and Bucicova [78] modeled customer satisfaction in the Slovak telecommunications industry using structural equation modeling and identified a positive relationship between corporate image and perceived service quality, as well as a negative relationship between customer expectations and perceived quality, underscoring the importance of expectation–perception gaps. Abdullah et al. [79] surveyed customers of telecommuni-

cations providers in Kurdistan and found that prices, network quality, and connectivity were the main drivers of satisfaction, concluding that value-added services did not affect quality perception.

Applications focused on telecommunications in Brazil remain limited. Study [80] examined customer satisfaction in call center management within a telecommunications company using stepwise multiple linear regression, identifying “First-Contact Resolution Rate” and “Average Handle Time After the Call” as the variables most strongly associated with satisfaction. Alternative indicators closely related to satisfaction were also proposed, including responsiveness, abandonment rate, company satisfaction, and problem-solving capability. Siebert et al. [2] investigated satisfaction prediction using DT, SVM, Gaussian process regression, ANN, and ensemble models applied to electricity distribution service data from COPEL, the utility responsible for this service in the state of Paraná. The dependent variable was an internal indicator used by the company, and regression models were evaluated using mean absolute percentage error (MAPE), yielding an average deviation of 1.36% from company-reported results.

Overall, this literature review highlights that Brazilian studies are scarce and generally focused on different scales or types of companies. While there is extensive research in the telecommunications sector, studies often rely on distinct independent variables and satisfaction indicators. Regarding NPS specifically, only two relevant articles were identified. Consequently, this research aims to address this gap by leveraging public, underutilized data to apply ML techniques, using the NPS to categorize the dependent variable (satisfaction) and incorporating variable importance analyses not presented in prior studies, particularly SHAP, a widely adopted and contemporary method compared to traditional statistical techniques and tree-based feature importance measures.

4. Materials and Methods

4.1. ML Models

This research analyzes customer satisfaction with broadband services in Brazil. We used data from ANATEL’s annual nationwide survey, which covers all Brazilian states [15]. The dataset’s overall customer satisfaction feature (“J1”) is the analysis’s target variable.

The NPS (Net Promoter Score) is an indicator used by service providers to assess customer loyalty [81]. Customers are asked, “On a scale of 0 to 10, how likely are you to recommend a particular company to someone else?” Based on their responses, customers are classified into different categories: those who respond with a score of 9 or 10 are referred to as “Promoters,” those who respond with a score of 7 or 8 are labeled as “Neutrals,” and those who respond with a score of 0 to 6 are categorized as “Detractors” [27]. Since the overall service satisfaction index is also on a 10-point Likert scale, the NPS was used as a reference to construct three classification models.

Although NPS is widely used, the literature contains articles that criticize its applications. Kristensen and Eskildsen [82] compared NPS and other metrics using simulated data, showing that grouping into promoters, neutrals, and detractors can vary and may make the final indicator imprecise. The same authors also reported that the indicator lacked precision and that its use could lead to poor management decisions [82].

Despite these issues with the NPS, Baquero [64] argues that many companies ultimately use NPS in conjunction with additional questions in multidimensional questionnaires, balancing detailed assessments of products and services with the item related to customer loyalty. In the present study, considering the assertion that there are potential associations between customer loyalty and satisfaction [69], and acknowledging that NPS alone is insufficient to analyze customer loyalty and business growth, as emphasized by [82], the NPS scale classification (detractors, neutrals, and promoters) was adopted to

categorize the score of interest representing overall customer satisfaction with the service analyzed in this work. This approach facilitates analysis within discrete groups by machine learning models.

Figure 1 presents a description of the three models used in the research. Based on the NPS scale, each model was obtained by applying distinct strategies to convert respondents' scores into satisfaction classes. Model 1 represents a binary classification dataset: customers who responded with a score of 8 or higher are classified as satisfied, and those with a score of 7 or lower are classified as dissatisfied. In Model 2, only the Detractor and Promoter classes of the NPS are considered, excluding the "Neutrals". Therefore, all customers who scored 7 or 8 were removed, leaving only the two extreme categories. Finally, Model 3 considers all three NPS classes simultaneously.

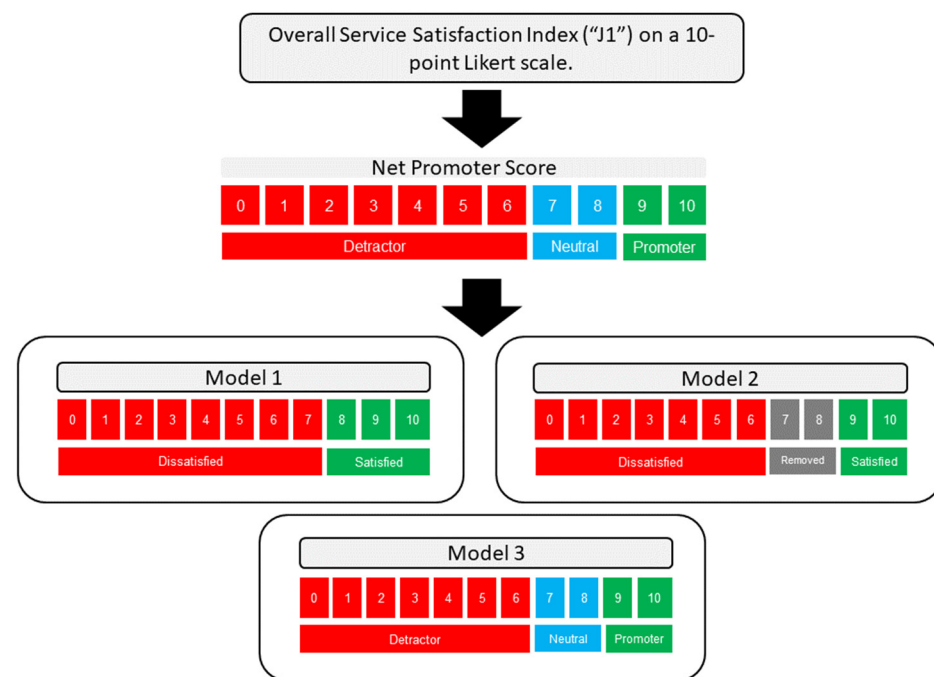


Figure 1. Description of the Models Applied in Research.

In Model 1, customers who rated 8 are classified as satisfied, while those who rated 7 are classified as dissatisfied. Since customers rated as neutral by the NPS represent an uncertain group, Model 1 aimed to analyze the ML application by treating a portion of the neutrals as satisfied and another as dissatisfied. The results were essential for identifying the impact on the techniques' performance and determining the most relevant features for classification.

Model 2, on the other hand, excludes neutral customers from the training dataset and therefore aims to analyze the impact on the techniques' performance without considering these instances. Therefore, this model is important for understanding and exploring the profile of customers classified as neutral, comparing the data resulting from Model 1, and for analyzing how ML techniques classify these uncertain data according to their training criteria.

The objective of Model 3 was to apply ML using the full set of NPS subdivision criteria. Model 3 aimed to assess the performance of the techniques in classifying data into promoters, neutrals, or detractors. While Models 1 and 2 explore the inclusion or exclusion of neutrals in the dataset, Model 3 is relevant because it uses all NPS classes without manipulation and provides a comprehensive analysis of classification results for each category.

Figure 2 illustrates the overall research workflow. In line with the general objective of analyzing customer satisfaction with broadband services in Brazil through the application of machine learning (ML) techniques to ANATEL data, the study was structured into three main stages: (1) Dataset preparation, where the data were explored and pre-processed for the next phase, (2) Application of ML algorithms, where the techniques were trained and tested in all models. Finally, (3) an analysis of neutral customers was performed using the algorithms trained in models 1 and 2.

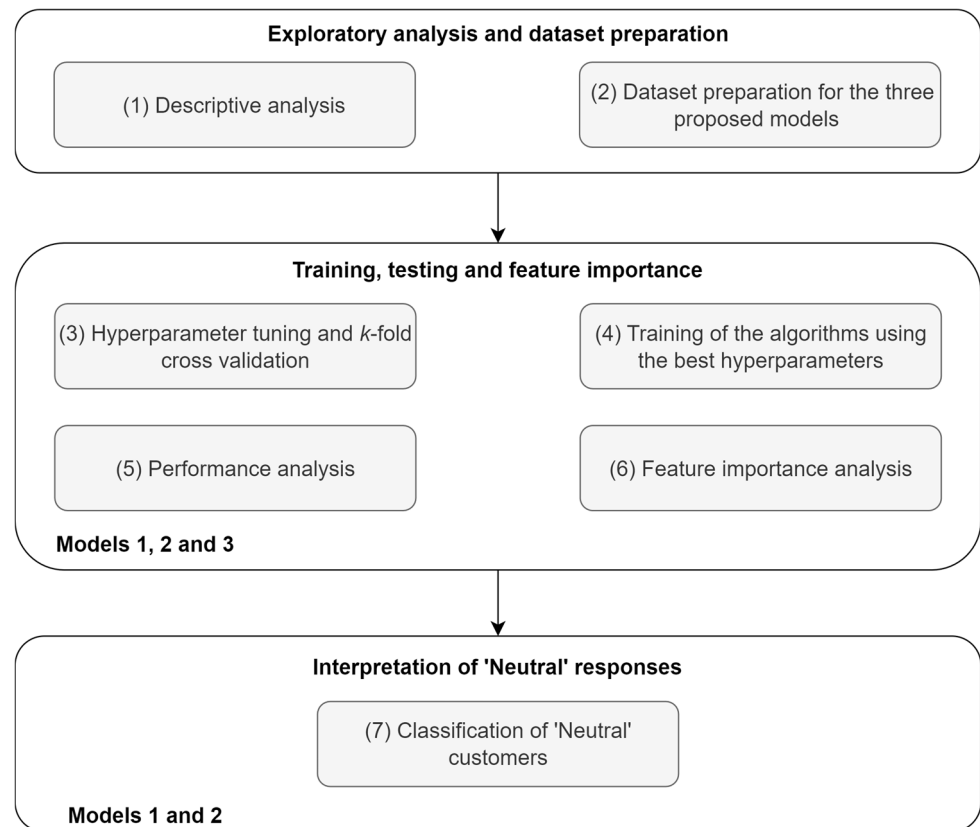


Figure 2. Research Process.

The steps and details of the proposed methodology are summarized in Table 1. All models use ML for classification and feature analysis, differing only in the characteristics and composition of their outputs. It is worth noting that Models 1 and 2 include an additional step: applying the trained algorithms to classify customers who are categorized as neutrals by the NPS.

Table 1. Description of the research process.

Steps	Description	Resources	Details
1	Descriptive analysis	Research glossary, Python 3.11 with Pandas, Matplotlib, Missigno, Numpy libraries.	Download the database, construction of a general research flowchart, and descriptive analysis of the data.

Table 1. Cont.

Steps	Description	Resources	Details
2	Dataset preparation for the three proposed models	Python 3.11 with Pandas, Numpy, and scikit-learn libraries.	Removal of empty cells, selection of a subset of the database, construction of the three models, and preparation of training and test datasets.
3	Hyperparameter tuning and k-fold cross validation	Python 3.11 with scikit-learn and pandas libraries.	Construction of the search space with hyperparameters and application of Grid Search.
4	Training of the algorithms using the best hyperparameters	Python 3.11 with scikit-learn and pandas libraries.	Training the algorithms using the parameters found in Grid Search, with a 50% holdout of the dataset for training.
5	Performance analysis	Python 3.11 with scikit-learn and pandas libraries.	Analysis of the performance of trained techniques through the metrics accuracy, precision, recall, and F1 score.
6	Feature importance analysis	Python 3.11 with scikit-learn, pandas, SHAP, and matplotlib libraries.	Analysis of the most relevant attributes for the classification of satisfaction, according to RF, DT, GB techniques, and the SHAP method.
7	Classification of ‘Neutral’ customers	Python 3.11 with scikit-learn, pandas, and matplotlib libraries.	Models 1 and 2 were employed to reclassify customers who had been initially categorized as Neutral based on the NPS reference scale.

4.2. Dataset

The database was extracted from a customer satisfaction survey conducted by ANATEL (<https://github.com/mrelero/dados-pesquisa-matheus-elero-UEM>, accessed on 23 February 2023), which was carried out annually between 2015 and 2020 and targeted broadband service users [82]. For detailed information on the database composition, ANATEL provides a glossary with descriptions, field types, and the survey flow [46].

The dataset, covering the period from 2015 to 2020, comprises 147,777 responses, with 63 variables corresponding to the survey questions. Table 2 presents the types of questions and responses in the ANATEL questionnaire. The 63 columns in the database consist of four data types, with the “open-ended” and “selection” variables containing both text and numbers. For binary responses, the value can take two values: “1” for “yes” and “2” for “no.” Lastly, the Likert scale ranges from 0 to 10, indicating the level of satisfaction with a specific feature.

Table 2. Questionnaire Question Formats.

Type	Description
Open	Open field for filling, such as the number of people in the family and the monthly family income

Table 2. *Cont.*

Type	Description
Selection	Selection from a dropdown list of options, such as state, service provider, interviewee’s occupation, gender, individual or legal entity, and age groups
“Yes” or “No”	Responses to objective questions about the execution of a specific action by the client. For example, if the consumer requested repairs on the internet installation.
10-point Likert scale	Level of satisfaction from 0 to 10 regarding a specific service attribute.

4.3. Data Preparation

The investigated database allows for various configurations and manipulations. As the dataset contains conditional questions, covers 5 years, has nationwide coverage, and comprises 147,777 rows, it can be used to generate other databases for ML applications. For this research, initial data preparation was conducted, and then the data were used to create three models based on the output column “J1” and the NPS indicator.

Table 3 presents the three steps carried out for data preparation. The first preparation consisted of identifying the subset used, removing columns and rows with empty cells, and responses that customers did not know how to answer. The period covered was 2017 to 2020, during which the questionnaire maintained the same structure each year. There were no restrictions based on age, state, income, gender, or the number of people using the broadband service. After this step, the database was reduced from 147,777 rows to 91,935 rows.

Table 3. Steps of data preparation.

Steps	The Number of Columns After the Stage	The Number of Rows After the Stage
1. Date limitation from 2017 to 2020.	63	91,935
2. Column removal	25	91,935
3. Removal of instances that had identifiers indicating that the question was not answered.	25	64,647

To expand the proportion used from the dataset, columns in which customers responded “yes” or “no” to specific subjects, such as E2, E4, E6, and E8, were considered, and the questions conditioned on these responses were consequently removed. Therefore, the following columns were removed: A2_1, A2_2, A2_3, A1_4, A3, A4, E2, E4, E6, E8, F2_1, F2_2, F2_3, F4_1, F4_2, and F4_3. Other columns with repeated, irrelevant, generalized, or sparse data were also removed: Q2, Q2_1, Q2_2, Q3, Q4, Q7a, IDTNS, G1, G2_1, G2_2, G2_3, PESO, Q1, Q6, H3, COD_IBGE, H2a, I2, and H0. As the objective of the project was to explore the most important features influencing service satisfaction, demographic data were excluded. After this processing, the database was reduced from 63 columns to 25.

Columns referring to optional response variables have been removed. Lines containing empty values or in which the respondent did not know how to answer the question were also removed. Finally, columns with string and categorical values were converted to numerical values, and columns with empty or filled “yes” or “no” responses were converted to binary values (0 or 1), where 0 represents “No” and 1 represents “Yes”. The

resulting dataset had 64,647 rows and 25 columns, with column “J1” representing the overall satisfaction level, which served as a reference for constructing the three models.

Table 4 presents the class distribution for each generated model. The results show that few customers are considered promoters, indicating that the balance in Model 3 is only 19%. The outcome indicates that models 1 and 2 have a predominance of dissatisfied customers, with 60% in the first and 69% in the second.

Table 4. Format of each model.

Model	Class	Proportion	Instances
Model 1	Satisfied	40%	38,890
	Dissatisfied	60%	26,057
Model 2	Satisfied	31%	12,463
	Dissatisfied	69%	27,108
Model 3	Promoter	19%	12,463
	Neutral	39%	25,057
	Detractor	42%	27,108

4.4. Application of ML Techniques

The application of ML in the three models was carried out using the scikit-learn and SciKeras libraries (<https://scikit-learn.org/stable/>, accessed on 13 February 2023). First, the models to be applied were selected, and then a hyperparameter tuning was conducted using the Grid Search method. With the selection of optimal parameters, nine algorithms were trained, tested, and analyzed using accuracy, precision, F1 score, recall, AUC (Area Under the Curve), and k-fold cross-validation. All 25 variables were used directly in the training set without normalization.

Table 5 presents all the techniques and the parameter values chosen for the Grid Search. This method searches for the best parameter combination using cross-validation [31]. All values were determined through individual tests of the techniques, with execution time and performance considered. The Grid Search method was applied using 5-fold cross-validation. The resulting parameters were selected to train and test all techniques. Finally, an additional test was performed using 10-fold cross-validation across all techniques, with accuracy as the evaluation metric. Results were visualized using box plots.

Table 5. Techniques and Hyperparameter tested.

ML	Parameters	Field
AdaBoost (AB)	n_estimators	1; 5; 10
Decision Tree (DT)	learning_rate	0.1; 1.2
	splitter	Best; Random
	criterion	Gini; entropy
	max_depth	6; 12; 16; 24
	min_samples_leaf	2; 4
Gradient Boosting (GB)	min_samples_split	2; 4
	loss	log_loss; exponential
	max_features	log2; sqrt
	learning_rate	0.01; 0.05; 0.1; 0.5; 1.0
	max_depth	3; 4; 5
	min_samples_leaf	4; 5; 6
	subsample	0.6; 0.7; 0.8
	n_estimators	5; 10; 15; 20

Table 5. Cont.

ML	Parameters	Field
Histogram Gradient Boosting (HGB)	max_iter	1.000; 1.200; 1.500
	learning_rate	0.1
	max_depth	25; 50; 75
	l2_regularization	1,5
k-Nearest Neighbors (KNN)	n_neighbors	1; 2; 4; 6; 7; 10
	weights	Uniform; distance
Logistic Regression (LR)	penalty	l2; none
	tol	1×10^{-6} ; 1×10^{-2}
	fit_intercept	True; False
	max_iter	10.000
Random Forest (RF)	n_estimators	200; 400
	max_depth	6; 12; 24
	min_samples_split	2; 4
	min_samples_leaf	2; 4
	criterion	Gini; entropy
Artificial Neural Network (ANN)	hidden_layer_dim	24; 48; 96
	loss	binary_crossentropy
	optimizer	adam; SGD
	optimizer_learning_rate	0.0001; 0.001; 0.1
	activation_hidden_layer	Sigmoid; relu
Support Vector Machine (SVM)	gamma	1; 0.1; 0.01; 0.001
	kernel	rbf; sigmoid
	C	1; 10; 100; 1.000

Table 6 presents the techniques and best resulting parameters from the Grid Search process. Considering the three models, it is possible to identify that in some techniques, such as SVM, RL, and AB, there was little variation in the chosen parameters, resulting in the same parameter composition across all three models. Since Grid Search can be time-consuming, the joblib library was used to save the trained models with the configured hyperparameters, preventing loss of development progress. A pseudocode is presented in Appendix A.

Table 6. Grid Search Results by Model and Technique.

ML	Parameters	Model 1	Model 2	Model 3
AB	n_estimators	10	10	10
	learning_rate	1	1	1
DT	splitter	"Best"	"Best"	"Best"
	criterion	"gini"	"entropy"	"entropy"
	max_depth	6	6	6
	min_samples_leaf	2	2	2
	min_samples_split	2	4	2
GB	loss	"exponential"	"log_loss"	"log_loss"
	max_features	"sqrt"	"sqrt"	"log2"
	learning_rate	0.1	0.1	0.1
	max_depth	5	5	5
	min_samples_leaf	5	6	5
	subsample	0.8	0.8	0.7
	n_estimators	20	20	20
HGB	max_iter	1200	1000	1000
	learning_rate	0.1	0.1	0.1
	max_depth	25	50	25
	l2_regularization	1.5	1.5	1.5

Table 6. Cont.

ML	Parameters	Model 1	Model 2	Model 3
KNN	n_neighbors	10	7	10
	weights	"distance"	"distance"	"distance"
LR	penalty	l2'	l2'	l2'
	tol	1.00×10^{-6}	1.00×10^{-6}	1.00×10^{-6}
	fit_intercept	"False"	"False"	"False"
	max_iter	10,000	10,000	10,000
RF	n_estimators	400	200	400
	max_depth	12	12	24
	min_samples_split	4	2	2
	min_samples_leaf	4	2	4
	criterion	"gini"	"entropy"	"gini"
ANN	"hidden_layer_dim"	96	48	96
	"loss"	"binary_crossentropy"	"binary_crossentropy"	"categorical_crossentropy"
	"optimizer"	"adam"	"adam"	"adam"
	"optimizer_learning_rate"	0.001	0.0001	0.001
	"activation_hidden_layer"	"relu"	"sigmoid"	"relu"
SVM	gamma	0.001	0.001	0.001
	kernel	rbf	rbf	rbf
	C	10	10	10

4.5. Feature Importance

Two approaches were used to find which factors matter most for broadband customer satisfaction in Brazil. The first method is to assess the importance of each feature using three computer models: RF, DT, and AB. These models measure how much each feature helps make better predictions during training. The process was done for all three models, and their results were compared.

In contrast to the first approach, the second one involved applying the SHAP (Shapley Additive exPlanations) analysis using the Python SHAP library, version 0.41 [83]. As presented by Lundberg and Lee [84], SHAP assigns an importance value to each feature for a given prediction. For this research, the RF technique was selected for SHAP application due to its strong performance relative to others, and the Tree Explainer method introduced by Lundberg et al. [85] was used in conjunction with the test set. The results were analyzed through bar charts and beeswarm plots for all three models.

4.6. Neutral Customers Analysis

Model 2 uses a dataset with customers who rated their satisfaction as "7" or "8" removed from the training set. Here, ML techniques are trained only on the extremes: customers who rate 9 or higher are satisfied, and those who rate 6 or lower are dissatisfied. This format mirrors NPS classification and excludes neutral customers.

Figure 3 presents a flowchart of the neutral analysis. For the Model 2 techniques, data from customers who rated their satisfaction as "7" and "8" were removed and used as input to the classifiers after training. This allowed classification according to the trained criteria. Model 1 classifiers also classified these neutral customers. The results were analyzed by calculating the percentages of satisfied and dissatisfied classifications, determined by ratings of "7" or "8." Heatmaps and cross-tabulation aided the analysis.

Neutral customers represent indecision regarding their satisfaction with the services provided. Therefore, it is necessary to determine whether such customers exhibit more characteristics of satisfaction or dissatisfaction. Model 1 and Model 2 were trained with different labeling strategies, so their application to classify neutral customers aims to better understand the characteristics of these instances by comparing the results.

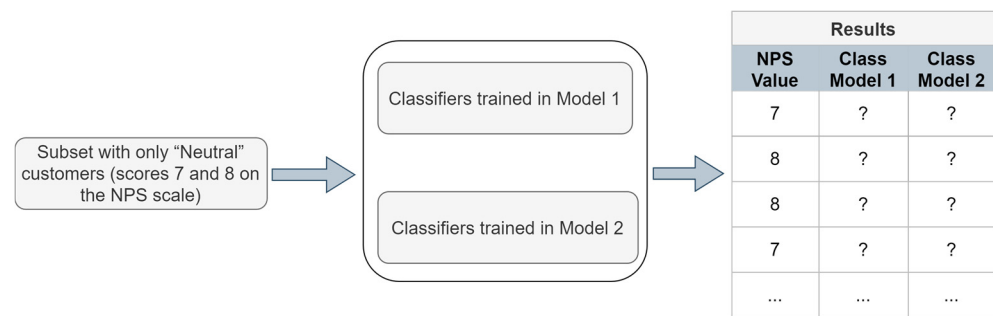


Figure 3. Neutral customers analysis diagram. The “?” corresponds to the class predicted by the algorithm.

5. Results

5.1. Exploratory Data Analysis

Figure 4 shows four histograms of overall satisfaction per year. There are no variations in ratings over time because the distribution of ratings is similar across years, with subtle differences in certain ratings. The average satisfaction score in 2017 was 6.40; in 2018, 6.48; in 2019, 6.54; and in 2020, 6.42. Therefore, the average satisfaction of broadband customers in the studied data was close to 6.5 between 2017 and 2020.

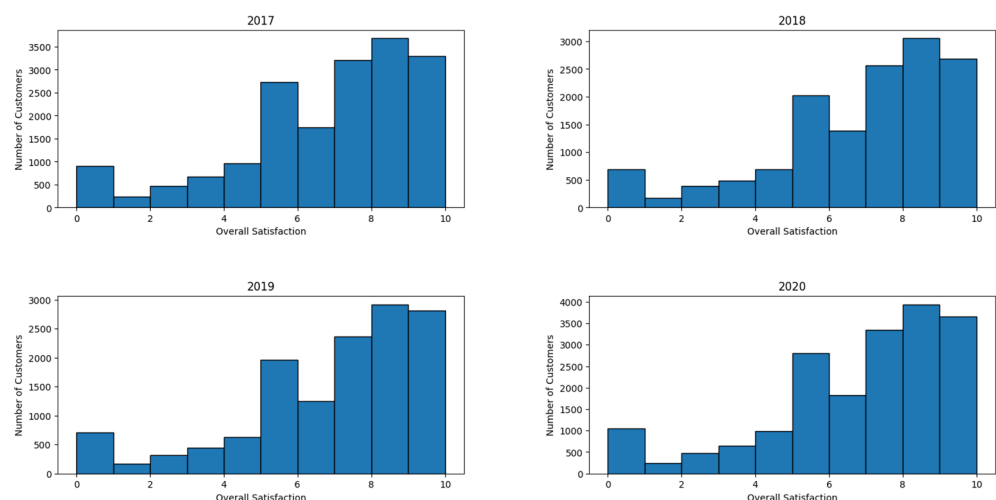


Figure 4. Histogram of satisfaction per year.

Figure 5 shows the number of instances per year, categorized by Model 3 classes. Between 2017 and 2020, the dataset exhibits minor fluctuations in the number of responses; starting with 17,912 in 2017, the volume decreased through 2018 and 2019 before rising to 18,987 in 2020. Furthermore, the proportion of customers classified according to NPS criteria remained consistent over time. Consequently, the NPS remained stable throughout the study period.

Figure 6 presents the number of responses per carrier. The survey results include data from customers of 16 different carriers, with OI (44,732) and NET (33,858) accounting for the highest volume of responses. These are followed by VIVO with 21,662 responses and SKY with 15,353 responses.

Figure 7 shows response counts by gender. Responses and class distribution are balanced between both genders in the survey.

Finally, Table 7 presents respondents' age and income. Per the survey manual, all participants were 18 or older, resulting in a minimum age of 18, a mean of 41, and a standard deviation of 13. The average income was R\$5631.67, ranging from R\$170.00 to R\$800,000.00.

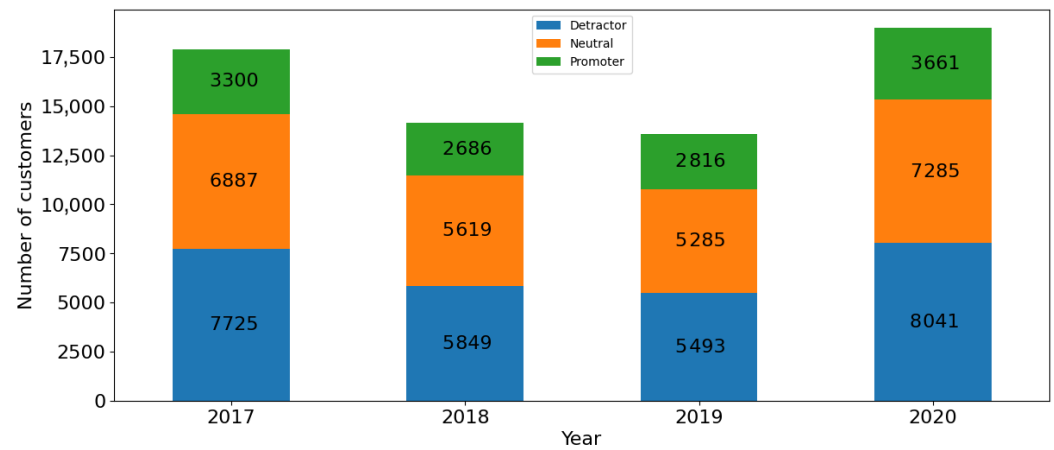


Figure 5. Number of responses per year.

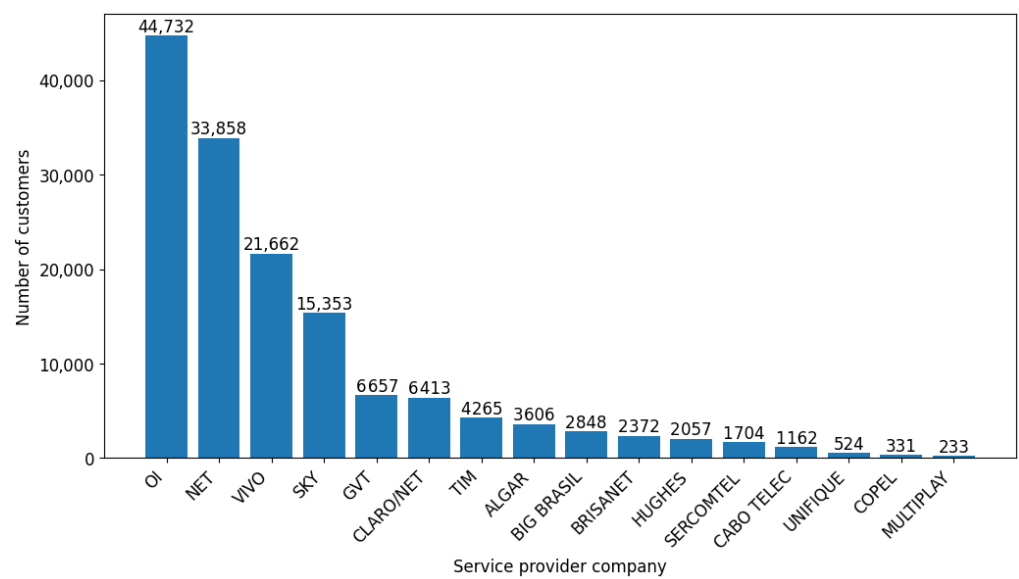


Figure 6. Number of responses by company.

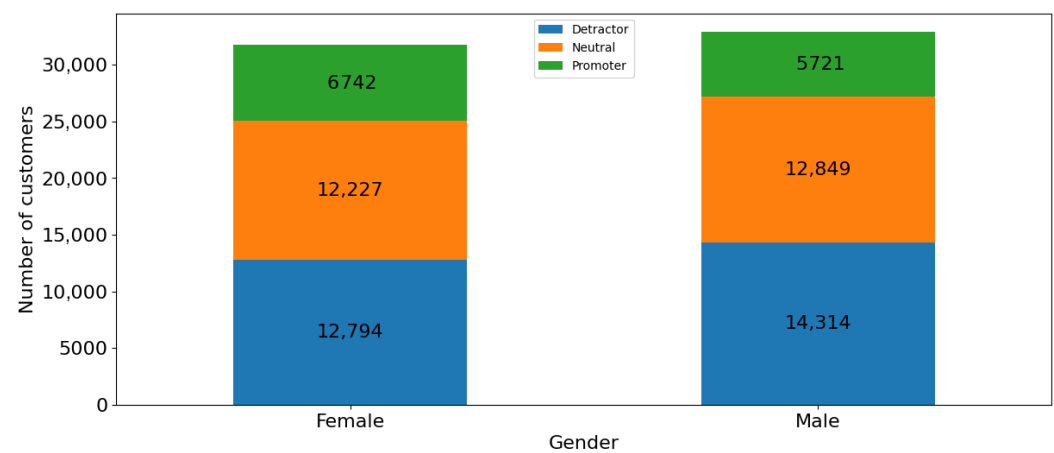


Figure 7. Number of responses by gender.

Table 7. Descriptive analysis of age and income data.

Statistic	Age	Income
Count	64,647	64,647
mean	41.77	5631.66
std	13.58	9229.15
min	18	170
25%	31	2000
50%	40	3500
75%	51	6000
max	97	800,000

5.2. Model 1

Table 8 presents the performance of all techniques trained with the parameters selected by Grid Search. The best results for some indicators are highlighted in green, and the worst results are shown in red. As observed, all techniques achieved accuracies ranging from 0.81 to 0.86, with RF, HGB, and GB achieving the highest accuracy. However, the KNN technique had the lowest accuracy, averaging 0.8132.

Table 8. Performance of Model 1 Techniques.

ML	Accuracy	Precision		Recall		F1		AUC-Score
		Satisfied	Dissatisfied	Satisfied	Dissatisfied	Satisfied	Dissatisfied	
AB	0.8502	0.8104	0.8776	0.8202	0.8705	0.8153	0.8741	0.9254
DT	0.8529	0.8205	0.8745	0.8128	0.8800	0.8167	0.8772	0.9270
GB	0.8576	0.8363	0.8711	0.8042	0.8937	0.8199	0.8823	0.9326
HGB	0.8596	0.8298	0.8794	0.8198	0.8865	0.8248	0.8829	0.9345
KNN	0.8132	0.7568	0.8541	0.7903	0.8286	0.7732	0.8412	0.8908
LR	0.8545	0.8143	0.8824	0.8277	0.8726	0.8209	0.8774	0.9277
RF	0.8596	0.8335	0.8765	0.8142	0.8902	0.8237	0.8833	0.9344
ANN	0.8394	0.8790	0.8208	0.6975	0.9352	0.7778	0.8743	0.9259
SVM	0.8427	0.8055	0.8678	0.8039	0.8690	0.8047	0.8684	0.9184

Legend: AB (AdaBoost); DT (Decision Tree); GB (Gradient Boosting); HGB (Histogram Gradient Boosting); KNN (k-Nearest Neighbors); LR (Logistic Regression); RF (Random Forest); ANN (Artificial Neural Networks); SVM (Support Vector Machine).

To classify customers as satisfied or dissatisfied using ML, reducing the false-positive rate is key to correctly identifying dissatisfied customers. In Model 1, AB, HGB, and LR had the highest precision for dissatisfied customers. All techniques averaged AUC scores near 0.92. HGB and GB yielded the best F1 scores.

Figure 8 presents cross-validation accuracy. SVM and HGB showed the lowest metric variance and the smallest range. RF, DT, and LR had intermediate variance and wide ranges. Also, KNN had the lowest accuracy, averaging 0.8132.

Table 9 shows feature importance for RF, DT, and GB. All three identify the same top five attributes, but the rankings differ. The customer's view of the company's ability to deliver on promises (B1_2) ranks highest for RF (21.00%) and GB (27.17%). For DT, customer satisfaction with internet browsing speed (C1_3) ranks highest at 65.61%. This attribute also ranks second and third for RF and GB, respectively. Pricing (C1_1) is the third most important for RF and second for GB.

The SHAP values in Figure 9 for predicting dissatisfied values show a similar relationship to those in Table 9, as the top five resulting attributes were the same. In the case of SHAP, attribute B1_2 ranked first in terms of impact on the output, and, as shown in

Figure 9, high customer scores for this attribute result in a negative impact on classifying them as dissatisfied. The same holds for attributes C1_3 and C1_1.

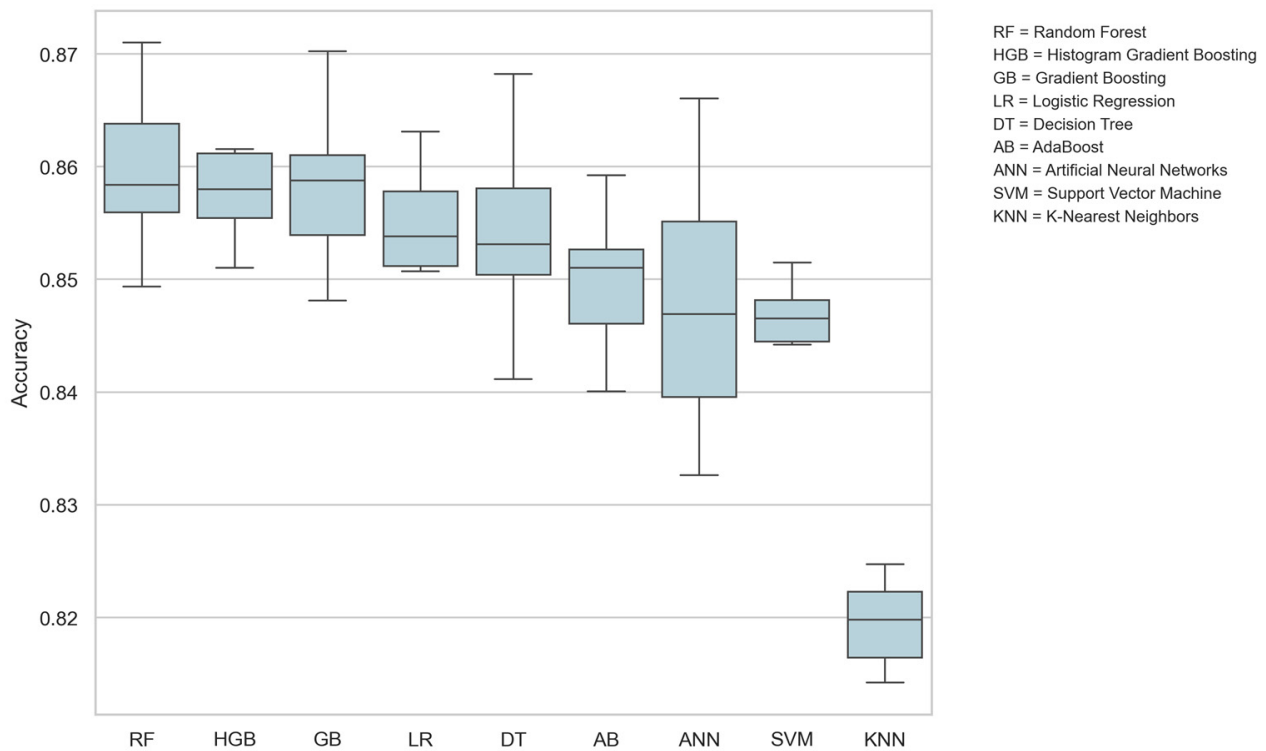


Figure 8. Box plot of accuracy in Model 1.

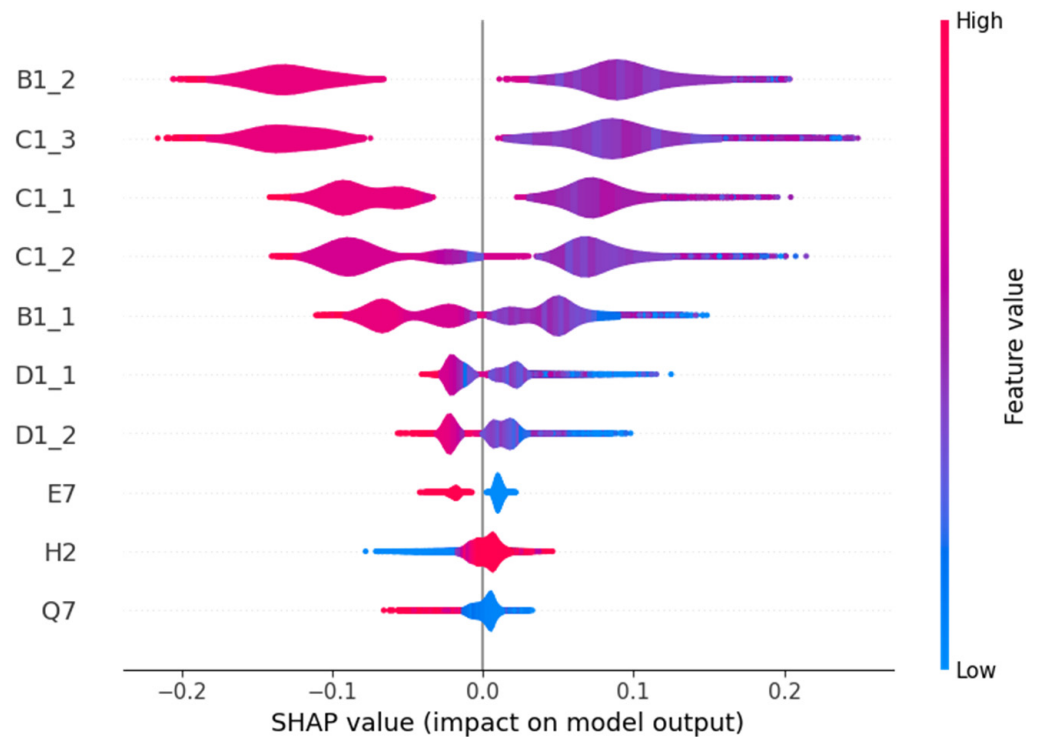


Figure 9. SHAP beeswarm plot for predicting dissatisfied customers in Model 1. Feature descriptions are provided in Appendix B.

Table 9. Most relevant attributes according to RF, DT, and GB in Model 1.

Code	Feature	RF	DT	GB
B1_2	Rating assigned regarding the operator's commitment to fulfill what was promised and advertised in their publicity.	21.00%	22.46%	27.17%
C1_3	Ratings assigned to browsing speed	20.44%	65.61%	23.95%
C1_1	Ratings assigned to billing accuracy according to the contracted services.	14.58%	3.34%	25.38%
C1_2	Ratings assigned to the ability to maintain a stable connection without drops	14.05%	5.57%	8.78%
B1_1	Rating assigned regarding the ease of understanding the contracted plans and services.	9.07%	1.71%	6.90%

Note: The heatmap shows the most important features in red, with less important features moving to green.

5.3. Model 2

Table 10 presents the performance results from applying ML techniques in Model 2. The results are better than those presented in Model 1. As observed, the techniques achieve an average accuracy of approximately 0.94, whereas the first model achieved around 0.85. One possible explanation for this is the removal of neutral customers during training and testing. As these customers belong to a region considered uncertain by NPS, their assignment can confuse the pattern of the two classes. In the case of Model 2, the techniques were applied and tested without these neutral customers, which may have favored the results, unlike the algorithms trained in Model 1.

Table 10. Performance of techniques for Model 2.

ML	Accuracy	Precision		Recall		F1		AUC-Score
		Dissatisfied	Satisfied	Dissatisfied	Satisfied	Dissatisfied	Satisfied	
AB	0.9479	0.9625	0.9175	0.9606	0.9212	0.9616	0.9194	0.9836
DT	0.9474	0.9598	0.9210	0.9627	0.9151	0.9613	0.9181	0.9805
GB	0.9523	0.9551	0.9461	0.9755	0.9035	0.9652	0.9243	0.9862
HGB	0.9537	0.9614	0.9369	0.9706	0.9179	0.9660	0.9273	0.9867
KNN	0.9183	0.9357	0.8806	0.9444	0.8634	0.9400	0.8719	0.9619
LR	0.9513	0.9625	0.9275	0.9658	0.9209	0.9642	0.9242	0.9800
RF	0.9541	0.9592	0.9428	0.9737	0.9128	0.9664	0.9275	0.9868
ANN	0.9488	0.9545	0.9361	0.9707	0.9026	0.9625	0.9190	0.9810
SVM	0.9377	0.9517	0.9078	0.9567	0.8979	0.9542	0.9028	0.9761

Legend: AB (AdaBoost); DT (Decision Tree); GB (Gradient Boosting); HGB (Histogram Gradient Boosting); KNN (k-Nearest Neighbors); LR (Logistic Regression); RF (Random Forest); ANN (Artificial Neural Networks); SVM (Support Vector Machine).

For Model 2, as shown in Table 10, the techniques RF, HGB, and GB achieved higher accuracies of 0.9541, 0.9537, and 0.9523, respectively, compared to other techniques. The KNN technique, in contrast, exhibited the lowest accuracy at approximately 0.9183. When comparing precision in identifying dissatisfied customers, AB, LR, and HGB yielded the most favorable results, each with values of 0.9625, 0.9625, and 0.9614, respectively. Regarding AUC, all techniques demonstrated similar results, averaging around 0.98, indicating consistent model discrimination performance across techniques.

Building on these results, Figure 10 presents the accuracy performance of all techniques with cross-validation. The graph shows that the ANN exhibited greater variation in

accuracy compared to GB, SVM, KNN, and RF, indicating lower performance in classifying customer satisfaction in Model 2. In contrast, GB, SVM, KNN, and RF demonstrated more consistent accuracy, with GB achieving the lowest standard deviation of 0.004123.

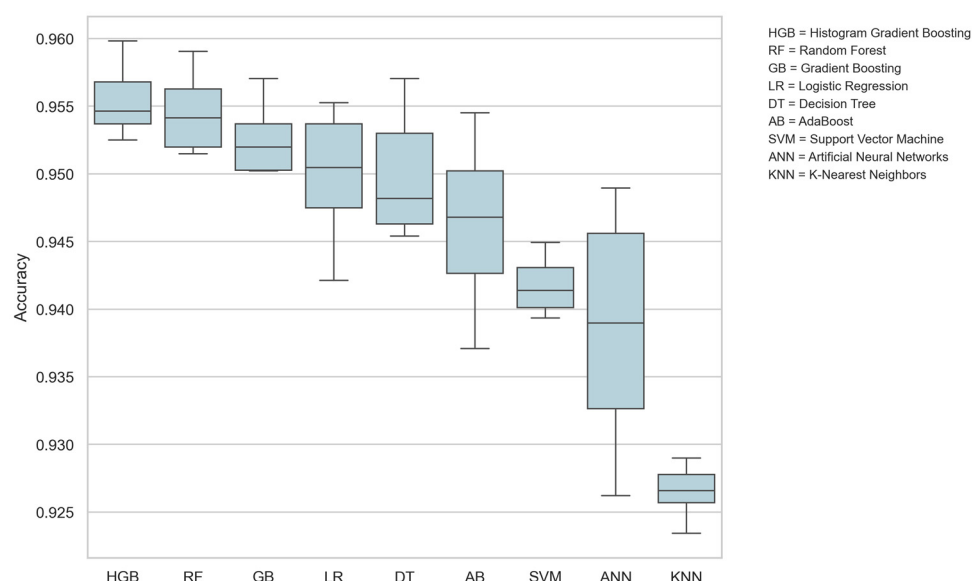


Figure 10. Box plot of accuracy in Model 2.

Table 11 presents the relationship of the most important features according to the RF, DT, and GB techniques. The survey showed that the first five features are the same across all techniques, only in different orders for each technique. RF resulted in a relevance of 21.42% for feature C1_3. This feature also had a high relative importance result for DT (20.75%) and GB (11.98%), but with less impact than in RF. The rating given regarding the operator's commitment to fulfilling what was promised (B1_2) is also a highlight for the techniques, ranking first for DT (68.05%) and GB (41.51%), and second for RF (21.15%). Another notable feature is C1_2, which ranked second for GB (19.36%).

Table 11. Most relevant features according to RF, DT, and GB in Model 2.

Code	Feature	RF	DT	GB
C1_3	Ratings assigned to browsing speed.	21.42%	20.75%	11.98%
B1_2	Rating assigned regarding the operator's commitment to fulfill what was promised and advertised in their publicity	21.15%	68.05%	41.51%
C1_2	Ratings assigned to the ability to maintain a stable connection without drops.	14.37%	5.11%	19.36%
C1_1	Ratings assigned to billing accuracy according to the contracted services.	13.91%	3.55%	11.44%
B1_1	Rating assigned regarding the ease of understanding the contracted plans and services.	8.05%	0.94%	7.23%

Note: The heatmap shows the most important features in red, with less important features moving to green.

Figure 11 presents the Beeswarm plot with SHAP values for predicting satisfied and dissatisfied customers in Model 2. The top three features with the greatest impact for this output are C1_3, B1_2, and C1_2. In all cases, high scores for these questions indicate a negative impact on classifying the customer as dissatisfied. In other words, as the consumer is satisfied with these features, the impact on their classification as satisfied is greater. Other

features, such as scores featured to billing values according to service performance (C1_1), ease of understanding plans, among others, have lower impacts (B1_1).

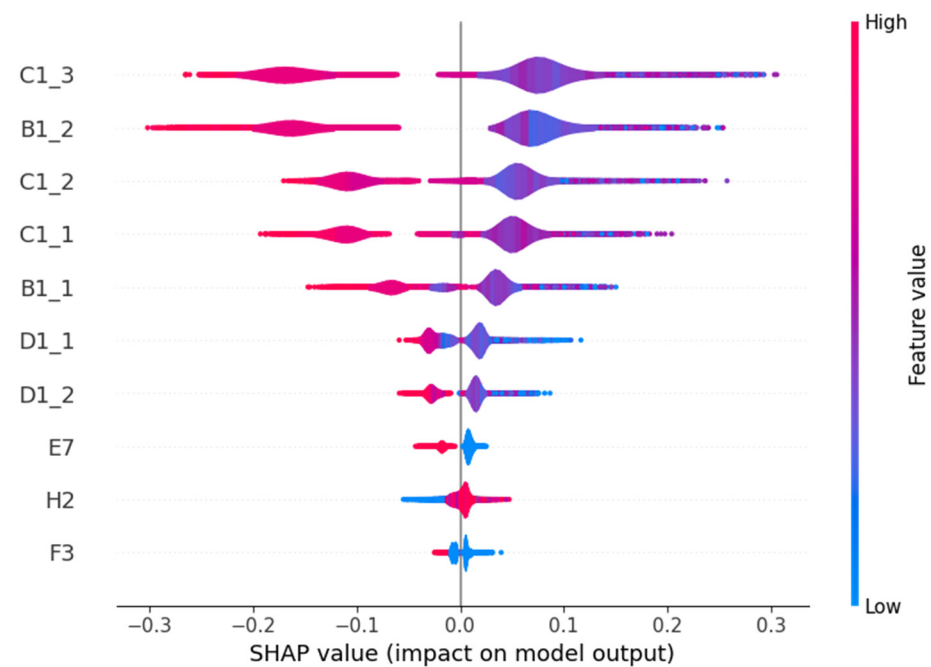


Figure 11. SHAP beeswarm plot for predicting dissatisfied customers in Model 2. Feature descriptions are provided in Appendix B.

5.4. Model 3

Model 3 aimed to apply classification techniques to predict whether a customer is a ‘Detractor,’ ‘Neutral,’ or ‘Promoter,’ making it a multiclass model. Table 12 shows the accuracy results for each technique. The performance on this criterion in multiclass classification models is lower than that of Models 1 and 2. As demonstrated in previous studies, multiclass classification is more challenging due to factors such as class overlap, imbalance rates, and varying class sizes [86]. The average result for this indicator is 76%, with RF, GB, and HGB emerging as the best-performing techniques. Similar to the other models, KNN also performed the worst in this aspect.

Table 12. Performance of ML techniques for Model 3.

ML	Accuracy	Precision			Recall			F1		
		Promoter	Neutral	Detractor	Promoter	Neutral	Detractor	Promoter	Neutral	Detractor
AB	0.7420	0.7114	0.6747	0.8181	0.6653	0.6895	0.8260	0.6876	0.6820	0.8220
DT	0.7617	0.7608	0.6882	0.8368	0.6299	0.7526	0.8311	0.6892	0.7189	0.8340
GB	0.7708	0.7869	0.6930	0.8462	0.6301	0.7740	0.8330	0.6998	0.7313	0.8395
HGB	0.7739	0.7599	0.7060	0.8476	0.6821	0.7527	0.8360	0.7189	0.7286	0.8418
KNN	0.6910	0.6639	0.5927	0.8306	0.4992	0.7420	0.7328	0.5699	0.6590	0.7786
LR	0.7649	0.7598	0.6840	0.8545	0.6577	0.7638	0.8155	0.7051	0.7217	0.8345
RF	0.7730	0.7877	0.6961	0.8475	0.6425	0.7718	0.8347	0.7078	0.7320	0.8410
ANN	0.7545	0.7076	0.6693	0.8856	0.7416	0.7583	0.7571	0.7242	0.7110	0.8163
SVM	0.7446	0.7253	0.6666	0.8354	0.6323	0.7362	0.8044	0.6756	0.6997	0.8196

Legend: AB (AdaBoost); DT (Decision Tree); GB (Gradient Boosting); HGB (Histogram Gradient Boosting); KNN (k-Nearest Neighbors); LR (Logistic Regression); RF (Random Forest); ANN (Artificial Neural Networks); SVM (Support Vector Machine).

Considering the detractors class, RF, HGB, GB, and LR obtained results above 0.84, as shown in Table 12. These techniques also presented noteworthy results in Recall and

F1 score. It should be noted that the model is better at classifying detractors and worse at classifying neutrals and promoters. For example, KNN has a precision of 0.8306 for classifying customers as detractors, 0.6639 for promoters, and 0.5927 for neutrals.

The results of cross-validation are presented in a box plot in Figure 12. As with the other models, the ANN showed higher variance and amplitude, with a standard deviation of 0.1016. This result is lower than other techniques, such as SVM (0.005657), KNN (0.005663), HGB (0.007913), and LR (0.006274). Consequently, HGB achieved the highest average accuracy, low variation, and consistent precision and recall in detecting detractor clients.

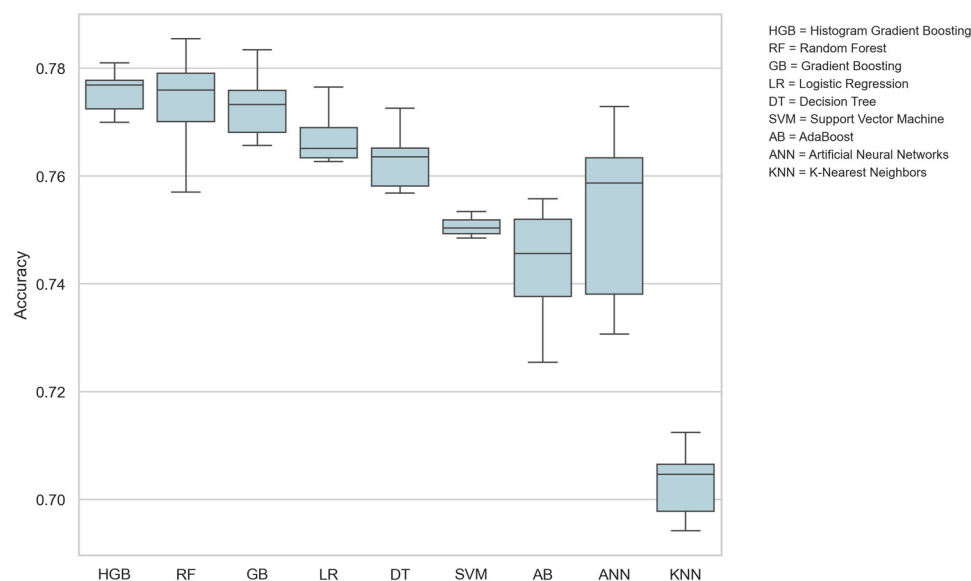


Figure 12. Box plot of accuracy in Model 3.

The most relevant features extracted from the RF, DT, and GB techniques are presented in Table 13. The data in the table align with the results from Models 1 and 2, as they use the same features, with changes in order and weights. In Model 3, the feature C1_3 was the most relevant for all three techniques, followed by B1_2. Other features, such as satisfaction with billing, ability to maintain a stable connection, and plan communication, also contribute to the rankings, but with lower weights than the top two.

Table 13. Most Important Features for RF, DT, and GB.

Código	Atributo	RF	DT	GB
C1_3	Ratings assigned to browsing speed	20.53%	53.55%	24.21%
B1_2	Rating assigned regarding the operator's commitment to fulfill what was promised and advertised in their publicity.	17.36%	31.34%	22.11%
C1_1	Ratings assigned to billing accuracy according to the contracted services.	13.89%	5.24%	20.72%
C1_2	Ratings assigned to the ability to maintain a stable connection without drops.	12.62%	7.56%	10.86%
B1_1	Rating assigned regarding the ease of understanding the contracted plans and services.	9.09%	2.09%	9.10%

Note: The heatmap shows the most important features in red, with less important features moving to green.

The Beeswarm plots with SHAP values for classifying Detractors, Neutrals, and Promoters are shown in Figure 13. For Detractors, customer satisfaction with fulfilling promises (B1_2), internet speed (C1_3), and billing (C1_1) have high importance indices, where higher scores on these features decrease the likelihood of being classified as a Detractor, and lower scores increase this likelihood. In contrast, for Neutrals and Promoters, higher scores on these same features increase the probability of classification, but the separation between positive and negative impact is less distinct than for Detractors.

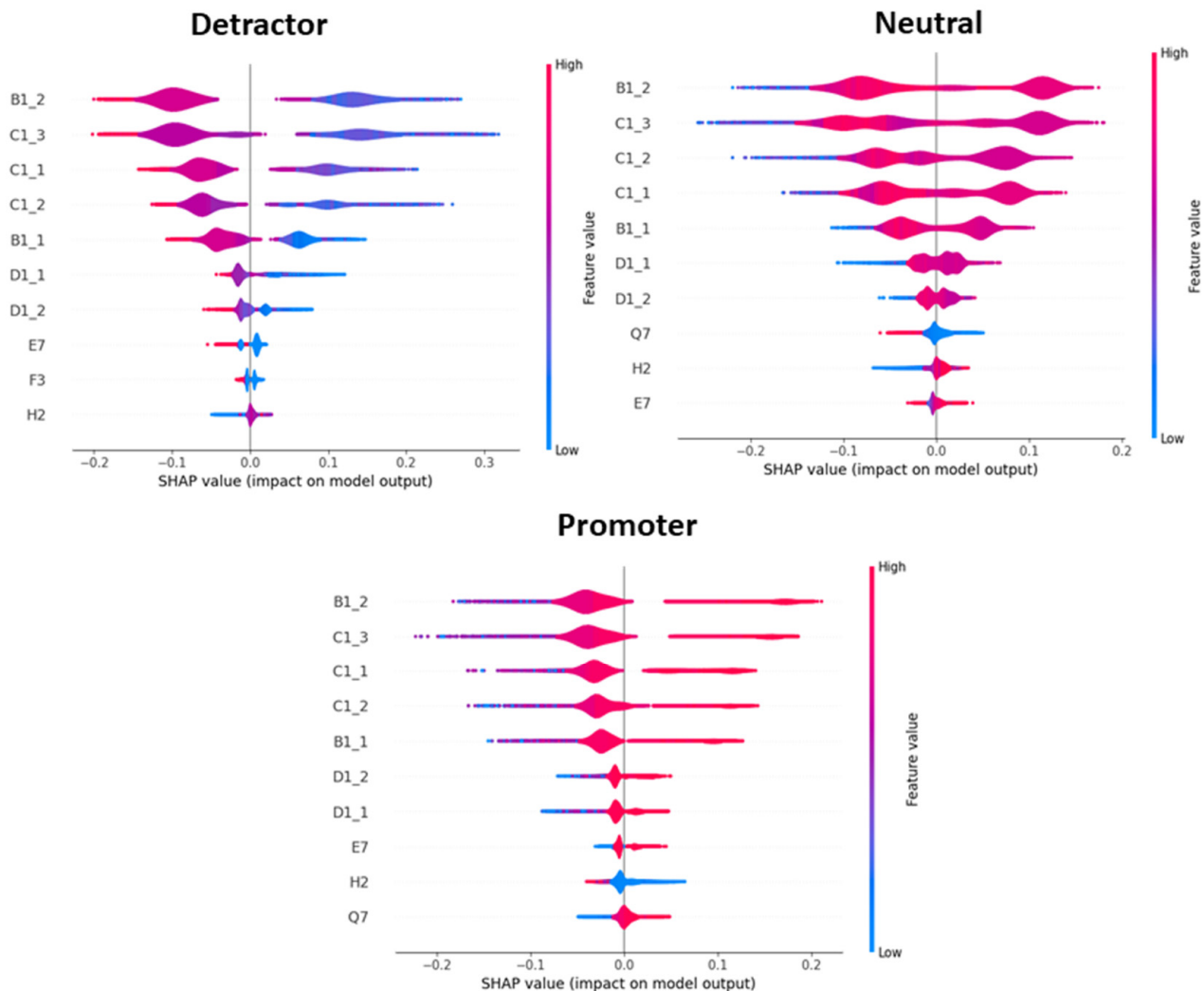


Figure 13. SHAP beeswarm plot for predicting dissatisfied customers in Model 3. Feature descriptions are provided in Appendix B.

5.5. Prediction of Neutral Customers

Neutral customers show indecision about their satisfaction with the services. To clarify, we must assess whether they show more traits of satisfaction or dissatisfaction. Model 1 and Model 2 used different labeling strategies. Applying both models to neutral customers helps us compare the results and understand their characteristics.

Figure 14 shows heatmaps of the classification results for neutral customers in Model 2, for each technique. For instance, in the case of RF in Model 2, among respondents who rated “7” for overall satisfaction, the algorithm classified 79.59% as dissatisfied and 20.41% as satisfied. For those who responded with a rating of “8”, 61.25% were classified as satisfied and 38.75% as dissatisfied. The other techniques showed similar classification proportions: customers who gave a rating of “7” (69.22% to 80.79%) are considered dissatisfied, whereas

those who gave a rating of “8” (59.28% to 66.88%) are considered satisfied. The results indicate that the algorithms trained in Model 2 treat ratings of “7” as closer to dissatisfaction and ratings of “8” as related to satisfaction.

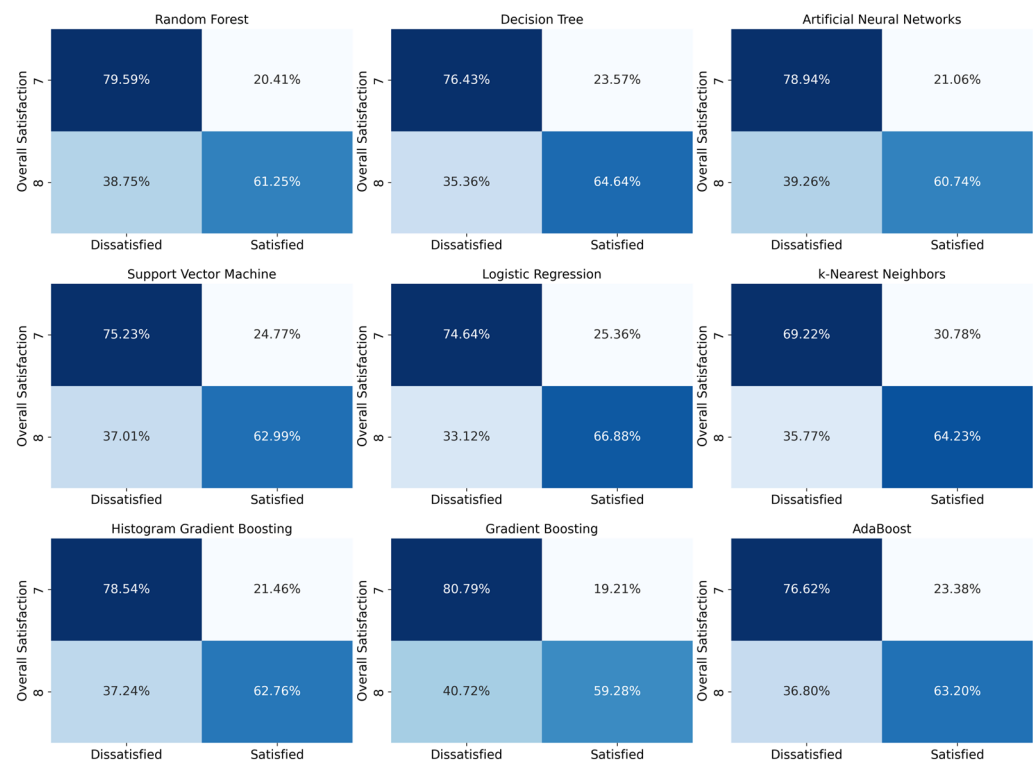


Figure 14. Neutral customers classified by algorithms trained in Model 2.

This same experiment was also conducted using the classifiers trained in Model 1, as shown in Figure 15. The results also show a pattern similar to that of the classifiers in Model 2. However, in this case, it is important to note that the classifiers were trained to treat customers who rated “7” as dissatisfied and “8” as satisfied, thereby increasing the proportions of satisfied and dissatisfied customers for ratings “7” and “8”. This led to increases in the proportions of respondents with ratings of “8” (satisfied), ranging from 53.03% to 84.85%, and “7” (dissatisfied), ranging from 67.95% to 80.74%.

The results from Model 2 indicate that neutral customers can cause uncertainty in the classification and performance of the techniques. First, the algorithm’s performance improved relative to Model 1, as it only used extreme values for classification, which favored the training process. Second, the classification of customers who responded with ratings “7” and “8” for overall satisfaction shows uncertainty about whether they are satisfied or dissatisfied, for both the algorithms in Model 1 and Model 2. Despite most techniques identifying “7” as dissatisfied and “8” as satisfied, a considerable number of instances are still assigned to the opposite class. The results indicate that the 10-point Likert scale amplifies uncertainty for classification techniques.

According to Reichheld [81], customers who are considered neutral are also within the spectrum of satisfaction with the service. However, they are in a region where their level of satisfaction does not significantly affect future company results, as there is a low likelihood that these customers will promote the service. However, ML predictions indicate that neutral customers may have profiles of both satisfied and dissatisfied customers. Therefore, treating neutral customers as those who do not affect the business can lead to erroneous conclusions. As Kristensen and Eskildsen [82] have shown, the NPS classes are imprecise and may be insufficient for service quality management.

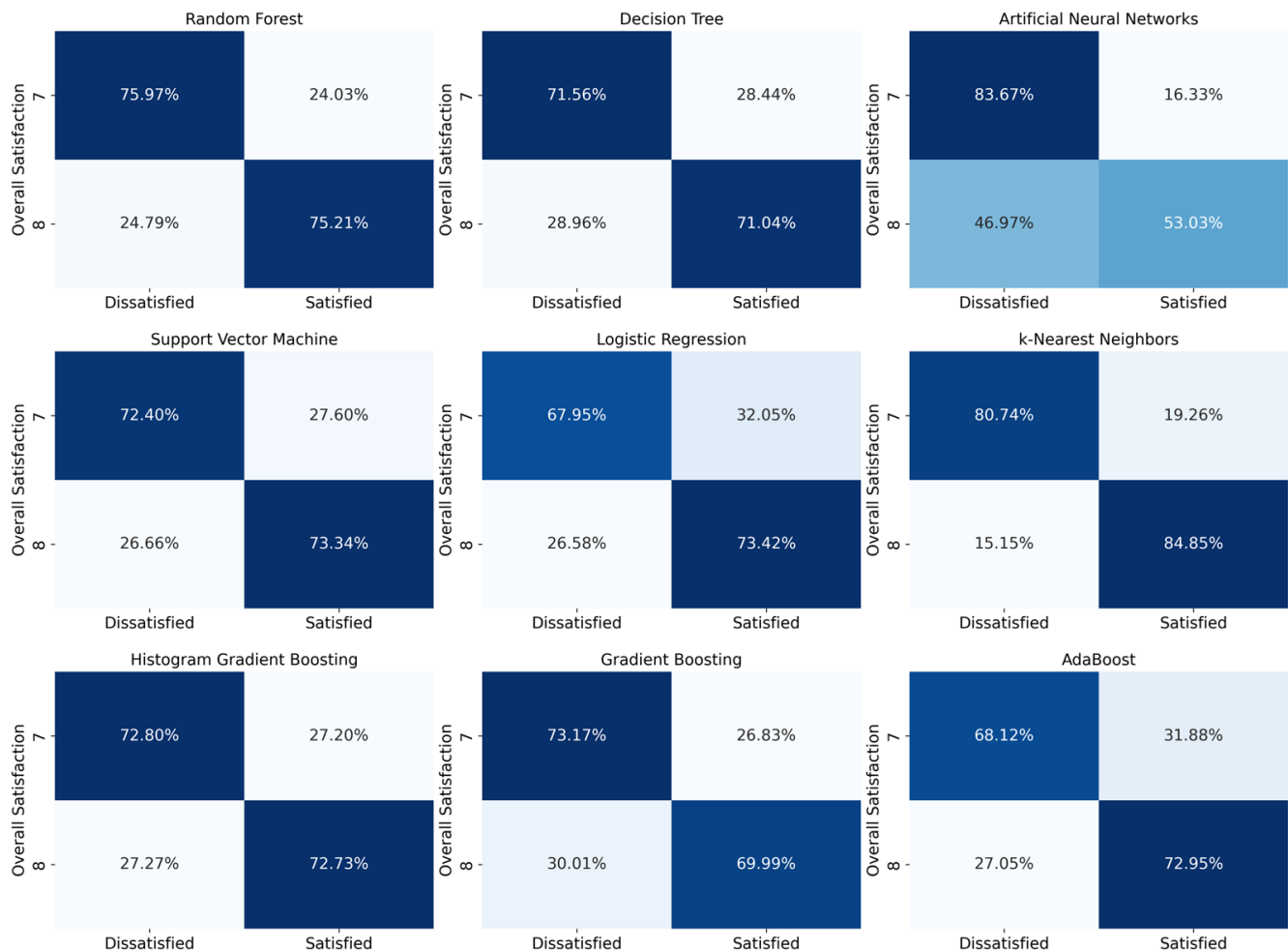


Figure 15. Neutral customers classified by algorithms trained in Model 1.

6. Discussion

In general, supervised ML techniques were applied to classify customers based on their satisfaction, using data from the ANATEL customer satisfaction survey. The Net Promoter Score (NPS), commonly used by service providers to categorize customers as detractors, neutrals, and promoters, was used as the basis for the classification.

Based on the “Overall Satisfaction with Service” output variable, three models were built with different thresholds to segment customers according to NPS benchmarks. Model 1 used binary classification, with satisfied customers defined as those who rated the service “8” or higher, and dissatisfied customers as those who rated below “8.” Model 2 removed customers who rated “7” or “8”; here, satisfied customers had ratings of “9” or above, while dissatisfied customers had ratings of “6” or below. Model 3 used multiclass classification based on the NPS classes: Detractors, Neutrals, and Promoters. For all models presented in the Section 5, preliminary analyses were conducted to assess the risk of overfitting. No discrepancies were observed between the results on the training and test datasets.

Table 14 summarizes the most relevant features and their positions based on the rankings in Tables 9, 11 and 13, generated using the ML techniques RF, DT, and GB, as well as the SHAP method. The top five features in all methods were: rating of the operator’s commitment to fulfill promises (B1_2), navigation speed (C1_3), billing accuracy with the contracted plan (C1_1), ability to maintain connection without drops (C1_2), and ease of understanding plans and services (B1_1). For RF, DT, and GB, C1_3, B1_2, and C1_1 showed the highest importance scores.

The SHAP results also indicated that features B1_2, C1_3, and C1_1 had higher average absolute impact values on the output. The SHAP values also enabled a more detailed identification of each feature's impact. In all cases, the importance of the features remained the same, with only changes in their positions. This result shows alignment with the findings from other methods.

Table 14. Summary of the order of the most relevant features obtained by ML techniques (RF, DT, and GB) and SHAP.

Code	Feature	Order of Importance by Model					
		Model 1		Model 2		Model 3	
		ML	SHAP	ML	SHAP	ML	SHAP
C1_3	Ratings assigned to browsing speed.	2	2	1	1	1	2
B1_2	Rating assigned regarding the operator's commitment to fulfill what was promised and advertised in their publicity.	1	1	2	2	2	1
C1_1	Ratings assigned to billing accuracy according to the contracted services.	3	3	4	4	3	3
C1_2	Ratings assigned to the ability to maintain a stable connection without drops.	4	4	3	3	4	4
B1_1	Rating assigned regarding the ease of understanding the contracted plans and service.	5	5	5	5	5	5

The findings from Table 14 align with related work by Abdullah et al. [79], who found that network quality and connectivity were among the main drivers of satisfaction in the telecommunications industry. In our results, browsing speed and connection stability had a strong impact on customer satisfaction classification. These aspects directly impact the customer experience and are related to service reliability, as connection failures and low speeds can disrupt customers' communications and daily activities.

Other factors identified as relevant in Table 14, such as billing accuracy, fulfillment of contract clauses as advertised, and ease of understanding the contracted plans, can be viewed as the gap between what the customer expected when hiring the services and the perceived quality after the services were provided. This is supported by Suchanek and Bucicova [78], who detected a negative relationship between customer expectations and perceived service value. This finding indicates that companies should align their marketing and communication efforts with the services they actually provide. Customers seem concerned about billing and contract transparency to avoid unexpected charges.

Regarding Brazil, the most important features are browsing speed, the delivery of the advertised service, and connection stability, which are key drivers of customer satisfaction. These findings are consistent with results reported by the Institute for Applied Economics [87], which showed that download rates, their reference values (contracted download speed), and customer satisfaction are closely related. This evidence suggests that, for most Brazilian consumers, the alignment between what is paid for at the time of contracting and what is actually delivered is critical, reinforcing the need for service providers to deliver what is contractually promised. Moreover, ANATEL requires that the average broadband internet speed delivered to consumers be at least 80% of the contracted speed and not fall below 40% at any time during usage. If these conditions are not met, consumers may file a complaint with the competent authority to enforce their rights and, as a last resort, pursue legal action against the service provider.

The research conducted by Markoulidakis et al. [77] also explored ML using NPS, but they used data from postpaid mobile telecom customers in Greece. The authors identified

several important features for classification using SHAP values in a linear regression model. The authors highlighted four features: “Network Voice,” which indicates the network quality for voice services; “Tariff Plan,” the billing plan used; “Billing,” the billing process; and “Network Data,” the data network quality. Despite the companies offering different services within the same industry, it is possible to draw some connections. Their results indicated that the most important features are related to billing and network quality, which aligns with the findings presented in this article, showing that some of the most relevant features are also related to payment and connection quality.

Another study that analyzed the most relevant features was conducted by Hosseini and Ziaei Bideh [47], which applied a satisfaction questionnaire to mobile telecom customers in Yazd province, Iran. The authors identified the most relevant features for satisfaction as value-added services, customer support, network quality, and plan prices. Similar to the findings by Markoulidakis et al. [77], their publication also indicated features related to payment and network quality.

Indeed, despite differences in terminology, the services provided, and the methodology used to calculate the features, this research aligns with the studies by Markoulidakis et al. [77] and Hosseini and Ziaei Bideh [47]. This study identified the five most important features for classification, among which one is related to billing (ratings assigned to billing based on the contract), and two are related to network quality, such as ratings for browsing speed and the ability to maintain a stable connection. Therefore, the most relevant features identified in this study indicate a connection to two research works in the literature.

7. Conclusions

Since 2016, there has been a significant growth in publications on the use of ML in service quality management. The available literature varies across various aspects, including scientific objectives, sectors, databases, methods, types of learning, and techniques. Given that the telecommunications sector still has few publications in the international literature and none in the Brazilian literature, this research aimed to apply ML techniques to analyze satisfaction data on telecommunications services in Brazil.

Beyond analyzing classifier performance metrics and XAI-based models for a deeper understanding of the underlying features, this study has both theoretical and practical implications. From a theoretical perspective, the findings demonstrate that the Net Promoter Score (NPS) can be integrated as a variable of interest in applications involving nonlinear ML-based models. Moreover, the XAI-based models identified several relevant variables that can be assessed or prioritized in any context involving broadband internet services, serving as references for customer satisfaction and loyalty strategies, as well as for marketing direction regarding product and service delivery.

From a practical standpoint, the results indicate that “browsing speed” and “billing” are priority variables in planning and managing critical factors to increase the likelihood of customer satisfaction and potential loyalty while simultaneously reducing the probability of churn. To this end, infrastructure stability and transparency in billing processes are essential. In addition, the nationwide survey conducted by ANATEL provides frequent, quantitative evidence to support shifts in broadband users’ priorities, serving as a reference for management practices, public policy formulation, and updates to regulations governing the provision of this service.

Therefore, the results presented in this article also reveal challenges in applying ML techniques to classify customer satisfaction. As a study limitation, the dataset’s use of a 10-point Likert scale as the output introduces uncertainty when transforming it into discrete classes, given the diversity of customer choices, intermediate levels, and neutral satisfaction states. Additionally, there are ongoing debates about the relationship between

customer satisfaction and loyalty, with no definitive evidence of causality. This study used NPS-based classification as a reference to facilitate the interpretation of classification algorithms; however, no in-depth analysis of this categorization was conducted. It is also worth noting that the data used (2017–2020) were selected based on availability at the beginning of the research, characterizing the study as retrospective. While more recent data were not analyzed, the proposed methodology can be extended to contemporary datasets.

Subsequent research could investigate additional classification techniques not explored in this study, as well as further preprocessing strategies such as linearization, data imputation, and class balancing. Regarding XAI-based interpretations, the use of alternative techniques, such as Local Interpretable Model-Agnostic Explanations (LIME), DeepSHAP, and Accelerated Model-Agnostic Explanations (ACME), would be valuable to complement analyses of feature importance across models. Furthermore, the methodology may be replicated using more recent data to compare the impact of predictor variables across NPS categories and observe changes in their importance for detractors and promoters. Finally, qualitative studies derived from this research could explore the telecommunications market more broadly, addressing service levels, oligopolistic structures, and emerging trends in customer requirements.

Author Contributions: Conceptualization, M.R.E. and R.H.P.L.; methodology, M.R.E., R.H.P.L. and B.S.d.S.; software, M.R.E.; validation, R.H.P.L., B.S.d.S. and G.C.L.L.; formal analysis, M.R.E. and R.H.P.L.; data curation, M.R.E.; writing—original draft preparation, M.R.E. and R.H.P.L.; writing—review and editing, B.S.d.S. and G.C.L.L.; visualization, R.H.P.L., B.S.d.S. and G.C.L.L.; supervision, R.H.P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are openly available in [github.com/mrelero at https://github.com/mrelero/machine-learning-anatel-database/tree/main/raw_data](https://github.com/mrelero/machine-learning-anatel-database/tree/main/raw_data), accessed on 2 October 2025.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AB	AdaBoost
ANATEL	Agência Nacional de Telecomunicações
ANN	Artificial Neural Networks
API	Application Programming Interface
AUC	Area Under the Curve
DT	Decision Tree
GB	Gradient Boosting
HGB	Histogram Gradient Boosting
KNN	k-Nearest Neighbors
LR	Logistic Regression
ML	Machine Learning
NPS	Net Promoter Score
RF	Random Forest
ROC	Receiver Operating Characteristic
SHAP	SHapley Additive exPlanations
SVM	Support vector machine

Appendix A. Pseudocode

Below is the pseudocode for applying machine learning models.

```

### Comments ###

# classifiers_GRID_Output: List of machine learning techniques and the
hyperparameters obtained from the grid search
# classifiers_GRID_Result: List of ML techniques with the hyperparameters obtained
from the configured classifiers_GRID_Output. Note: Different variables were used to
optimize code performance.

for number, i in enumerate (ML_ParamGridList):
    ## Run Grid Search
    classifiers_GRID_Output[i] = GridSearchCV (estimator = classifiers_GRID_Input[i],
param_grid = ML_ParamGridList[i], cv = 5)
    classifiers_GRID_Output[i].fit (X_train, y_train)

    ## Set the params resulted by Grid Search
    classifiers_GRID_Result[i].set_params(**classifiers_GRID_Output[i].best_params_)

    ## Train
    classifiers_GRID_Result[i].fit (X_train, y_train)

    ## Predict
    y_pred = classifiers_GRID_Result[i].predict (X_test)
    acc = metrics.accuracy_score (y_test, y_pred)
    y_pred_prob = classifiers_GRID_Result[i].predict_proba (X_test)
    y_pred_prob = y_pred_prob[:, 1]

    ## Get the results
    cm = confusion_matrix (y_test, y_pred)
    pr = precision_score (y_test,y_pred)
    rec = recall_score (y_test,y_pred)
    F1 = f1_score (y_test, y_pred)
    fpr, tpr, threshold = roc_curve (y_test,y_pred_prob)
    plot_roc_curve (fpr,tpr, i)
    auc = roc_auc_score (y_test,y_pred_prob)

```

Appendix B. Relevant Variables and Descriptions

Description	Code
Rating assigned to the ease of understanding the contracted plans and services.	B1_1
Rating assigned to the provider's commitment to delivering what was promised and advertised.	B1_2
Rating assigned to billing accuracy according to the contracted services.	C1_1
Rating assigned to the ability to maintain a stable connection without outages.	C1_2

Description	Code
Rating assigned to browsing speed.	C1_3
Rating assigned to billing accuracy according to the contracted services.	D1_1
Rating assigned to the clarity of information on the bill.	D1_2
Indicates whether the respondent contacted the provider regarding fixed-internet performance issues.	E7
Indicates whether the respondent requested any repair for fixed-internet service in the 6 months prior to the survey.	F3
Monthly household income (recorded as reported).	H2
Individual or legal entity? (1 for individual; 2 for legal entity).	Q7

References

1. Abimanyu, D.P.S. Indonesian online travel agent sentiment analysis using machine learning methods. *Indones. J. Electr. Eng. Comput. Sci.* **2019**, *14*, 113–117. [\[CrossRef\]](#)
2. Siebert, L.C.; Bianchi, J.F.; Da Silva, E.J.; Yamakawa, E.K.; Catapan, A. Predicting customer satisfaction for distribution companies using machine learning. *Int. J. Energy Sect. Manag.* **2021**, *15*, 743–764. [\[CrossRef\]](#)
3. Kowalski, R.; Esteve, M.; Mikhaylov, S.J. Improving public services by mining citizen feedback: An application of natural language processing. *Public Adm.* **2020**, *98*, 1011–1026. [\[CrossRef\]](#)
4. Zhang, C.; Moreira, M.R.A.; Sousa, P.S.A. A bibliometric view on the use of total quality management in services. *Total Qual. Manag. Bus. Excell.* **2021**, *32*, 1466–1493. [\[CrossRef\]](#)
5. Altuntas, S.; Dereli, T.; Erdoğan, Z. Evaluation of service quality using SERVQUAL scale and machine learning algorithms: A case study in health care. *Kybernetes* **2021**, *51*, 846–875. [\[CrossRef\]](#)
6. Berry, L.L.; Parasuraman, A.; Zeithaml, V.A. Improving service quality in America: Lessons learned. *Acad. Manag. Perspect.* **1994**, *8*, 32–45. [\[CrossRef\]](#)
7. Prihatna, K.A.; So, I.G.; Saroso, H.; Kosasih, W. What about service quality, satisfaction, and loyalty in e-business? A systematic review of literature. In Proceedings of the 2021 International Conference on Information Management and Technology (ICIMTech), Jakarta, Indonesia, 19–20 August 2021; pp. 767–772.
8. Torrão, J.; Teixeira, S. The Antecedents of Customer Satisfaction in the Portuguese Telecommunications Sector. *Sustainability* **2023**, *15*, 2778. [\[CrossRef\]](#)
9. Abd-Elrahman, A.-E.H.; Hassan, S.A.; El-Borsaly, A.A.-E.; Hafez, E.A.-E. A critical review of alternative measures of telecommunications service quality. *Int. J. Qual. Serv. Sci.* **2020**, *12*, 247–263. [\[CrossRef\]](#)
10. Owusu-Kyei, M.; Kong, Y.; Owusu Akomeah, M.; Owusu Afriyie, S. Assessing the Influence of Total Quality Management on Customer Satisfaction in the Telecom Industry: A TQM–SERVQUAL Perspective. *Businesses* **2023**, *3*, 251–271. [\[CrossRef\]](#)
11. Halpern, N.; Mwesiumo, D. Airport service quality and passenger satisfaction: The impact of service failure on the likelihood of promoting an airport online. *Res. Transp. Bus. Manag.* **2021**, *41*, 100667. [\[CrossRef\]](#)
12. Norouzi, V. Predicting E-Commerce CLV with Neural Networks: The Role of NPS, ATV, and CES. *J. Econ. Technol.* **2024**, *2*, 174–189. [\[CrossRef\]](#)
13. Wang, J.; Wu, J.; Sun, S.; Wang, S. The relationship between attribute performance and customer satisfaction: An interpretable machine learning approach. *Data Sci. Manag.* **2024**, *7*, 164–180. [\[CrossRef\]](#)
14. Aftan, S.; Shah, H. Using the AraBERT Model for Customer Satisfaction Classification of Telecom Sectors in Saudi Arabia. *Brain Sci.* **2023**, *13*, 147. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Park, E.; Kim, S.-B. Globalization and the Korean Wave: How Online Reviews Reflect Customer Satisfaction in U.S. Korean Restaurants. *Sage Open* **2025**, *15*, 21582440251. [\[CrossRef\]](#)
16. Liu, P.; Shi, X.; Xu, Y.; Dang, R.; Wu, Y. Integration of Machine Learning with Comprehensive IVIF-QFD-MCDM Framework for Enhancing Online Hotel Operations. *Inf. Sci.* **2025**, *720*, 122493. [\[CrossRef\]](#)
17. Mohanty, A.; Mohapatra, A.G.; Tripathy, P.K.; Bal, P.K.; Tripathy, A.K.; Nayak, S.; Mohanty, S.K.; Jena, B. Smart Hospitality Using IoT Enabled Integrated Face Recognition, Machine Learning, and Fuzzy AHP for Analyzing Customer Satisfaction Measurements. *Int. J. Inf. Technol.* **2025**, *17*, 1597–1605. [\[CrossRef\]](#)

18. Mathebula, M.; Modupe, A.; Marivate, V. ChatGPT as a Text Annotation Tool to Evaluate Sentiment Analysis on South African Financial Institutions. *IEEE Access* **2024**, *12*, 144017–144043. [[CrossRef](#)]
19. Alrizq, M.; Alghamdi, A. Customer Satisfaction Analysis with Saudi Arabia Mobile Banking Apps: A Hybrid Approach Using Text Mining and Predictive Learning Techniques. *Neural Comput. Appl.* **2024**, *36*, 6005–6023. [[CrossRef](#)]
20. Thambirajan, S.; Devaraj, K. Conceptual Study on E-Banking Systems and Customer Satisfaction Using Deep Learning and Blockchain. *Int. J. Knowl. Manag. Stud.* **2025**, *16*, 77–94. [[CrossRef](#)]
21. Ghallab, H.; Nasr, M.; Fahmy, H. Enhancing Customer Churn Analysis by Using Real-Time Machine Learning Model. *Int. J. Adv. Comput. Sci. Appl.* **2025**, *16*, 388–396. [[CrossRef](#)]
22. Li, Q.; Jing, R.; Zhu, X. Determinants of Travel Satisfaction for Commercial Airlines: A Data Mining Approach. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108597. [[CrossRef](#)]
23. Jethwa, S.; Vora, N. A Comprehensive Analysis of Airline Passenger's Satisfaction through Classification Model. *J. Trends Comput. Sci. Smart Technol.* **2024**, *6*, 324–337. [[CrossRef](#)]
24. Dinh, D.H.; Lap Le, S.N. Using Ensemble Classification Algorithms to Predict Airline Customer Satisfaction. *Appl. Mark. Anal. Peer-Rev. J.* **2025**, *11*, 175. [[CrossRef](#)]
25. Lucini, F.R.; Tonetto, L.M.; Fogliatto, F.S.; Anzanello, M.J. Text Mining Approach to Explore Dimensions of Airline Customer Satisfaction Using Online Customer Reviews. *J. Air Transp. Manag.* **2020**, *83*, 101760. [[CrossRef](#)]
26. Ponnusamy, M.; Sonvani, T.; Parida, R.; Nanda, G. A Machine Learning Based Semi-Automated Framework for House of Quality Analysis. *Comput. Ind. Eng.* **2025**, *208*, 111371. [[CrossRef](#)]
27. Imtiaz, N.; Ben Islam, K. Identifying Significance of Product Features on Customer Satisfaction Recognizing Public Sentiment Polarity: Analysis of Smart Phone Industry Using Machine-Learning Approaches. *Appl. Artif. Intell.* **2020**, *34*, 832–848. [[CrossRef](#)]
28. Zhou, F.; Ayoub, J.; Xu, Q.; Yang, X.J. A Machine Learning Approach to Customer Needs Analysis for Product Ecosystems. *J. Mech. Des.* **2020**, *142*, 011101. [[CrossRef](#)]
29. Le, H.-S.; Do, T.-V.H.; Nguyen, M.H.; Tran, H.-A.; Pham, T.-T.T.; Nguyen, N.T.; Nguyen, V.-H. Predictive Model for Customer Satisfaction Analytics in E-Commerce Sector Using Machine Learning and Deep Learning. *Int. J. Inf. Manag. Data Insights* **2024**, *4*, 100295. [[CrossRef](#)]
30. Duong, Q.H.; Zhou, L.; Meng, M.; Dang, L.T.A.; Nguyen, T.D. Striking the Right Balance: Customising Return Policy Leniency for Managing Customer Online Return Proclivity and Satisfaction. *J. Retail. Consum. Serv.* **2025**, *85*, 104315. [[CrossRef](#)]
31. Beyari, H. Artificial Intelligence's Effect on Customer Loyalty in the Context of Electronic Commerce. *J. Umm Al-Qura Univ. Eng. Archit.* **2025**, *16*, 617–626. [[CrossRef](#)]
32. Zaghloul, M.; Barakat, S.; Rezk, A. Enhancing Customer Retention in Online Retail through Churn Prediction: A Hybrid RFM, K-Means, and Deep Neural Network Approach. *Expert. Syst. Appl.* **2025**, *290*, 128465. [[CrossRef](#)]
33. Tian, Z.; Zhong, R.Y.; Barenji, A.V.; Wang, Y.T.; Li, Z.; Rong, Y. A Blockchain-Based Evaluation Approach for Customer Delivery Satisfaction in Sustainable Urban Logistics. *Int. J. Prod. Res.* **2021**, *59*, 2229–2249. [[CrossRef](#)]
34. Aljohani, K. The Role of Last-Mile Delivery Quality and Satisfaction in Online Retail Experience: An Empirical Analysis. *Sustainability* **2024**, *16*, 4743. [[CrossRef](#)]
35. Pacheco, F.; Exposito, E.; Gineste, M. A Framework to Classify Heterogeneous Internet Traffic with Machine Learning and Deep Learning Techniques for Satellite Communications. *Comput. Netw.* **2020**, *173*, 107213. [[CrossRef](#)]
36. Wassouf, W.N.; Alkhatib, R.; Salloum, K.; Balloul, S. Predictive Analytics Using Big Data for Increased Customer Loyalty: Syriatel Telecom Company Case Study. *J. Big Data* **2020**, *7*, 29. [[CrossRef](#)]
37. Namli, Ö.H.; Yanık, S.; Nouri, F.; Serap Şengör, N.; Koyuncu, Y.M.; Uçar, Ö.B. A Neural Networks Approach to Predict Call Center Calls of an Internet Service Provider1. *J. Intell. Fuzzy Syst.* **2021**, *42*, 503–515. [[CrossRef](#)]
38. Espadinha-Cruz, P.; Fernandes, A.; Grilo, A. Lead Management Optimization Using Data Mining: A Case in the Telecommunications Sector. *Comput. Ind. Eng.* **2021**, *154*, 107122. [[CrossRef](#)]
39. Yayah, F.C.; Ghauth, K.I.; Ting, C.-Y. The Performance of Classification Method in Telco Customer Trouble Ticket Dataset. *Int. J. Comp. Sci.* **2022**, *49*, 1–10.
40. Rintyarna, B.S.; Kuswanto, H.; Sarno, R.; Rachmaningsih, E.K.; Rachman, F.H.; Suharso, W.; Cahyanto, T.A. Modelling Service Quality of Internet Service Providers during COVID-19: The Customer Perspective Based on Twitter Dataset. *Informatics* **2022**, *9*, 11. [[CrossRef](#)]
41. Haq, I.; Soomro, J.A.; Mazhar, T.; Ullah, I.; Al Shloul, T.; Ghadi, Y.Y.; Ullah, I.; Saad, A.; Tolba, A. Impact of 3G and 4G Technology Performance on Customer Satisfaction in the Telecommunication Industry. *Electronics* **2023**, *12*, 1697. [[CrossRef](#)]
42. Saleh, S.; Saha, S. Customer Retention and Churn Prediction in the Telecommunication Industry: A Case Study on a Danish University. *SN Appl. Sci.* **2023**, *5*, 173. [[CrossRef](#)] [[PubMed](#)]
43. Afzal, M.; Rahman, S.; Singh, D.; Imran, A. Cross-Sector Application of Machine Learning in Telecommunications: Enhancing Customer Retention Through Comparative Analysis of Ensemble Methods. *IEEE Access* **2024**, *12*, 115256–115267. [[CrossRef](#)]

44. Chajia, M.; Nfaoui, E.H. Customer Churn Prediction Approach Based on LLM Embeddings and Logistic Regression. *Future Internet* **2024**, *16*, 453. [CrossRef]
45. AbdelAziz, N.M.; Bekheet, M.; Salah, A.; El-Saber, N.; AbdelMoneim, W.T. A Comprehensive Evaluation of Machine Learning and Deep Learning Models for Churn Prediction. *Information* **2025**, *16*, 537. [CrossRef]
46. Satisfação e Qualidade Percebida. Available online: <https://www.gov.br/anatel/pt-br/consumidor/compare-as-prestadoras/pesquisa-de-satisfacao-e-qualidade> (accessed on 5 March 2023).
47. Hosseini, S.Y.; Bideh, A.Z. A data mining approach for segmentation-based importance-performance analysis (SOM-BPNN-IPA): A new framework for developing customer retention strategies. *Serv. Bus.* **2014**, *8*, 295–312. [CrossRef]
48. Li, C.; Wang, Y.; Li, D.; Chu, D.; Ma, M. An Effective Method of Evaluating Pension Service Quality Using Multi-Dimension Attention Convolutional Neural Networks. *Int. J. Softw. Eng. Knowl. Eng.* **2021**, *31*, 533–543. [CrossRef]
49. Tong, L.; Wang, Y.; Wen, F.; Li, X. The research of customer loyalty improvement in telecom industry based on NPS data mining. *China Commun.* **2017**, *14*, 260–268. [CrossRef]
50. Gewaly, M.M.; Saeed, M.S.; Ammar, A.M. Machine Learning Approach for Complaint Prediction in the Telecom Industry. In Proceedings of the Intelligent Methods, Systems, and Applications (IMSA), Giza, Egypt, 13–14 July 2024; pp. 12–17.
51. Rožman, M.; Kišić, A.; Oreški, D. Comparative Analysis of Nonlinear Models Developed Using Machine Learning Algorithms. *WSEAS Trans. Inf. Sci. Appl.* **2024**, *21*, 303–307. [CrossRef]
52. Grebovic, M.; Filipovic, L.; Katnic, I.; Vukotic, M.; Popovic, T. Machine Learning Models for Statistical Analysis. *Int. Arab. J. Inf. Technol.* **2023**, *20*, 505–514. [CrossRef]
53. Albuquerque, E.; Eduardo, E. Evaluation Of The Capability Of The Mobile Phone Service In Southern Brazil Through The Perception Of Its Customers. *Int. J. Qual. Res.* **2018**, *12*, 441–458.
54. Bernardes, B.V.C. Caracterização da Pesquisa de Satisfação e Qualidade Percebida da Anatel Para Telefonia Móvel. Ph.D. Thesis, Fluminense Federal University, Niterói, Brazil, 2020.
55. Hermosilla, P.; Díaz, M.; Berríos, S.; Allende-Cid, H. Use of Explainable Artificial Intelligence for Analyzing and Explaining Intrusion Detection Systems. *Computers* **2025**, *14*, 160. [CrossRef]
56. Srirahayu, D.P.; Anugrah, E.P.; Layyindah, K. Influence of Satisfaction and Loyalty on Net Promoter Score (NPS) in Academic Libraries in Indonesia. *Libr. Manag.* **2021**, *42*, 325–339. [CrossRef]
57. Zaghloul, M.; Barakat, S.; Rezk, A. Predicting E-Commerce Customer Satisfaction: Traditional Machine Learning vs. Deep Learning Approaches. *J. Retail. Consum. Serv.* **2024**, *79*, 103865. [CrossRef]
58. Ferreyra, L.F.M.; Tauil, Y.B.; Adigneri, H.M.; dos Santos, B.S.; Lima, R. Analysis of Economic Satisfaction Using Machine Learning Models and Explainable Artificial Intelligence. *Navus-Rev. De. Gestão E Tecnol.* **2024**, *15*, 1–16. [CrossRef]
59. Bortolotti, S.L.V.; Junior, F.d.J.M.; Bornia, A.C.; Júnior, A.F.d.S.; de Andrade, D.F. Consumer Satisfaction and Item Response Theory: Creating a Measurement Scale. *Gestão Produção* **2012**, *19*, 287–302. [CrossRef]
60. Monfort, A.; López-Vázquez, B.; Sebastián-Morillas, A. Building Trust in Sustainable Brands: Revisiting Perceived Value, Satisfaction, Customer Service, and Brand Image. *Sustain. Technol. Entrep.* **2025**, *4*, 100105. [CrossRef]
61. Drosos, D.; Skordoulis, M.; Arabatzis, G.; Tsotsolas, N.; Galatsidas, S. Measuring Industrial Customer Satisfaction: The Case of the Natural Gas Market in Greece. *Sustainability* **2019**, *11*, 1905. [CrossRef]
62. Leiba, Y.; Shirokopetleva, M.; Gruzdo, I. Research on methods of determining customer loyalty and assessing their level of satisfaction. *Innov. Technol. Sci. Solut. Ind.* **2023**, *24*, 104–117. [CrossRef]
63. Czajkowska, A.; Manuela, I. Application of Servqual and Servperf Methods to Assess the Quality of Teaching Services—Comparative Analysis. *Manuf. Technol.* **2021**, *21*, 294–305. [CrossRef]
64. Baquero, A. Net Promoter Score (NPS) and Customer Satisfaction: Relationship and Efficient Management. *Sustainability* **2022**, *14*, 2011. [CrossRef]
65. Patti, C.H.; van Dessel, M.M.; Hartley, S.W. Reimagining Customer Service through Journey Mapping and Measurement. *Eur. J. Mark.* **2020**, *54*, 2387–2417. [CrossRef]
66. D’Avila, O.P.; Perrone, L.R.; Ribeiro, M.C.; Pinto, L.F. The Use of Net Promoter Score for Evaluating Primary Health Care: Results from Population-Based Surveys. *Cien Saude Colet.* **2024**, *29*, e01452024. [CrossRef]
67. Hallowell, R. The Relationships of Customer Satisfaction, Customer Loyalty, and Profitability: An Empirical Study. *Int. J. Serv. Ind. Manag.* **1996**, *7*, 27–42. [CrossRef]
68. Aggarwal, C.C. *Data Mining: The Textbook*; Springer International Publishing: Cham, Switzerland, 2015; ISBN 978-3-319-14141-1.
69. Géron, A. *Hands-on: Machine Learning with Scikit-Learn, Keras & Tensorflow*, 2nd ed.; O’Reilly Media: Sebastopol CA, USA, 2019; ISBN 978-1492032649.
70. Mohamed, A.; Abdelqader, K.; Shaalan, K. Explainable Artificial Intelligence: A Systematic Review of Progress and Challenges. *Intell. Syst. Appl.* **2025**, *28*, 200595. [CrossRef]
71. Vishwarupe, V.; Joshi, P.M.; Mathias, N.; Maheshwari, S.; Mhaisalkar, S.; Pawar, V. Explainable AI and Interpretable Machine Learning: A Case Study in Perspective. *Procedia Comput. Sci.* **2022**, *204*, 869–876. [CrossRef]

72. Asif, D.; Arif, M.S.; Mukheimer, A. A Data-Driven Approach with Explainable Artificial Intelligence for Customer Churn Prediction in the Telecommunications Industry. *Results Eng.* **2025**, *26*, 104629. [CrossRef]
73. Dandolo, D.; Masiero, C.; Carletti, M.; Dalle Pezze, D.; Susto, G.A. AcME—Accelerated Model-Agnostic Explanations: Fast Whitening of the Machine-Learning Black Box. *Expert. Syst. Appl.* **2023**, *214*, 119115. [CrossRef]
74. Amin, M.N.; Ahmad, W.; Khan, K.; Nazar, S.; Arab, A.M.A.; Deifalla, A.F. Evaluating the Relevance of Eggshell and Glass Powder for Cement-Based Materials Using Machine Learning and SHapley Additive ExPlanations (SHAP) Analysis. *Case Stud. Constr. Mater.* **2023**, *19*, e02278. [CrossRef]
75. Cox, T.; Bell, G. A machine learning approach to process improvement in a telecommunications company. *Ann. Oper. Res.* **1996**, *65*, 21–34. [CrossRef]
76. Cox, T.; Bell, G.; Glover, F. A New learning approach to process improvement in a telecommunications company. *Prod. Oper. Manag.* **1995**, *4*, 217–227. [CrossRef]
77. Markoulidakis, I.; Rallis, I.; Georgoulas, I.; Kopsiaftis, G.; Doulamis, A.; Doulamis, N. A Machine Learning Based Classification Method for Customer Experience Survey Analysis. *Technologies* **2020**, *8*, 76. [CrossRef]
78. Suchanek, P.; Bucicova, N. The Customer Satisfaction Model in the Mobile Telecommunications Sector after Covid-19 Pandemic. *PLoS ONE* **2025**, *20*, e0317093. [CrossRef] [PubMed]
79. Abdullah, N.N.; Prabhu, M.; Othman, M.B. Analysing Driving Factors of Customer Satisfaction among Telecommunication Service Providers in Kurdistan Region. *Int. J. Eng. Bus. Manag.* **2022**, *14*, 184797902211114. [CrossRef]
80. Joia, L.A.; Oliveira, A.F. Call Center Key Performance Indicators and Customer Satisfaction. In Proceedings of the AMCIS 2010 Proceedings, Lima, Peru, 12–15 August 2010; pp. 1–10.
81. Reichheld, F.F. The one number you need to Grow. *Harv. Bus. Rev.* **2003**, *81*, 46.
82. Kristensen, K.; Eskildsen, J. The validity of the Net Promoter Score as a business performance measure. In Proceedings of the 2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering, Guiyang, China, 12–15 June 2011; pp. 970–974.
83. An Introduction to Explainable AI with Shapley Values. Available online: <https://www.scribd.com/document/840971119/An-Introduction-to-Explainable-AI-With-Shapley-Values-SHAP-Latest-Documentation> (accessed on 3 December 2022).
84. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30.
85. Lundberg, S.M.; Erion, G.; Chen, H.; Degrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef]
86. Lango, M.; Stefanowski, J. What Makes Multi-Class Imbalanced Problems Difficult? An Experimental Study. *Expert. Syst. Appl.* **2022**, *199*, 116962. [CrossRef]
87. Santos, P.S.L.; Corseuil, C.H.L.; Borges, C.P. TD 2810—Percepção e Qualidade: O Acesso à Internet Banda Larga e a Avaliação Dos Consumidores à Luz Da Teoria Da Perspectiva. In *Texto Para Discussão*; Instituto de Pesquisa Econômica Aplicada—Ipea: Brasília, Brazil, 2022; pp. 1–42. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.