International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 3, June 2025, pp. 3362~3372
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i3.pp3362-3372

Enhancing tabular data analysis for classification of airline passenger satisfaction using TabNet deep neural network

Rachid Kaidi[1], Hicham Omara[2], Mohamed Lazaar[3], Mohammed Al Achhab[4]
[1]Faculty of Sciences, Abdelmalek Essaadi University, Tetouan, Morocco
[2]Polydisciplinary Faculty, Sidi Mohammed Ben Abdellah University, Taza, Morocco
[3]National Higher School of Computer Science and Systems Analysis, Mohammed V University in Rabat, Rabat, Morocco
[4]National School of Applied Sciences, Abdelmalek Essaadi University, Tetouan, Morocco

Article Info

Article history:

Received Sep 1, 2024
Revised Jan 13, 2025
Accepted Mar 3, 2025

Keywords:

Aviation
Customer experience
Customer feedback
Deep learning
Machine learning
Passenger satisfaction
Service quality

ABSTRACT

In an era of air travel, understanding and enhancing passenger satisfaction are pivotal to the success of airlines and the overall passenger experience. Analyzing airline passenger satisfaction using tabular data can pose various challenges, both when employing classical statistical methods and when leveraging machine learning and deep learning techniques. On the one hand, statistical approaches pose various challenges including limited feature engineering techniques, the assumption of linearity of the data sets and limited predictive power. Then again, the use of machine learning and deep learning techniques may face other challenges such as the problem of overfitting, difficulty in interpreting data, intensive resource requirements, and the generalization problem in deploying machine learning-based methods. This paper presents a novel deep learning approach utilizing TabNet to classify airline passenger satisfaction. Leveraging a comprehensive dataset comprising various passenger-related attributes, our TabNet-based model demonstrates exceptional performance in distinguishing between satisfied and dissatisfied passengers. Our model's robustness in handling tabular data, underscores its power as a valuable tool for the aviation industry. Comparing out results to recent papers show that out model outperforms these studies in terms of accuracy, precision, recall and area under the curve. The results show that our TabNet Network model outperforms all implemented machine learning models by reaching respectively the following results: 96.47%, 96.41% and 96.24% for accuracy, F1-score and G-mean score.

Corresponding Author:

Rachid Kaidi
Faculty of Sciences, Abdelmalek Essaadi University
Tetouan, Morocco
Email: kaidirachid1982@gmail.com

1. INTRODUCTION

Aviation is one of the most widely employed and safest modes of transportation. The anticipated scenario was a threefold increase in global air transport demand from 2020 to 2050 [1], [2]. The International Air Transport Association (IATA) predicts that there will be a 3.7% yearly increase in air travel demand, resulting in an estimated 7.2 billion air passengers by the year 2035 [3], [4]. That is why millions of passengers around the globe prefer is because of safety and professional services. In an era defined by rapid globalization and heightened connectivity, aviation transport stands as a cornerstone of mod ern mobility, linking people and places with unprecedented speed and accessibility. That is why the aviation industry has consistently strived to ensure not only the safety and efficiency of air travel but also the satisfaction and well-being of its passengers. As the demand for air travel continues to surge, understanding and measuring passenger satisfaction has become an imperative goal for airlines and the aviation industry as a whole.

Traditionally, the measurement of passenger satisfaction in aviation has relied on conventional surveys and feedback mechanisms [5]–[9]. While these methods have provided valuable insights, they often suffer from limitations such as subjectivity, small sample sizes, and delayed responses. Moreover, the aviation landscape is dynamic, with countless factors affecting passenger satisfaction, including ticket pricing, flight punctuality, in-flight amenities, and customer service. To stay competitive and responsive to passenger needs, the aviation industry requires more sophisticated and real-time approaches to gauge passenger satisfaction comprehensively such as data mining for instance [10].

In the context of the aviation industry in the USA, Hayadi *et al.* [11] conducted an analysis of competition and customer satisfaction among airlines. They employed various classical classification models, including K-nearest neighbors (KNN), logistic regression (LR), Gaussian naïve Bayes (NB), decision trees (DT), and random forest (RF). The first step in their analysis was to clean the data, which consisted of 130,000 samples and 22 features. Out of these 22 features, 14 were obtained through a survey. After removing samples that containing either 0 values or NaN, they were left with a dataset containing 70,000 samples. Among the models tested, Random Forest yielded the best results with an area under the curve (AUC) of 0.99, precision of 0.97, and recall of 0.94. The default threshold value for random forest was set at 0.5. However, when they increased the threshold to 0.7, the precision improved from 0.97 to 0.99.

Passenger satisfaction plays a crucial role in airline selection. To assess passenger satisfaction levels, Nurdina and Puspita [12] employed two machine learning models for classification. They utilized a Kaggle dataset comprising 26,000 samples with 9 features and one label. The first model used was naïve Bayes, which leverages probability and statistical calculations to predict the class of samples.

For the second model they used KNN, which detects the input class by calculating Euclidean distances. KNN is simple to implement, easy to understand and can capture complex and nonlinear relationships in data, making it effective for classification tasks. The test results revealed a precision value of 82.25% for naïve Bayes and 67.35% for KNN, which is relatively low and indicates the need to fine-tune the models for better results.

Predicting aircraft passenger satisfaction and identifying the primary influencing factors play a vital role in helping airlines enhance their services, gain a competitive edge, and tackle challenging situations effectively. To address this, Jiang *et al.* [13] proposed an innovative model called RF-RFE-LR, which combines feature selection techniques to streamline the dataset. The dataset consisted of 23 features, categorized into basic information, flight details, and passenger satisfaction attributes. To determine the optimal number of features for analysis, they conducted cross-validation experiments, finding that utilizing 17 features resulted in the best performance. The study compared various classification models, including KNN, logistic regression, random forest, Gaussian naïve Bayes, and BP neural networks, before and after feature selection. The results indicated that the RF model, with its 17 selected features, achieved the best result in terms of accuracy of 0.963, precision of 0.973, recall of 0.942, F1 value of 0.957, and an AUC value of 0.961. This demonstrated the model's robustness in predicting passenger satisfaction and revealed its potential for guiding airlines in making data-driven decisions to enhance customer experiences and overall service quality.

Homaid and Moulitsas [14] conducted a comprehensive investigation employing various ML algorithms, including RF, LR, support vector machine (SVM), XGBoost, and naïve Bayes (NB). Their findings demonstrated that XGBoost exhibited superior performance compared to other ML models, achieving impressive results in terms of accuracy (88%), precision (85%), recall (83%), and F1-Score (84%). Kumar *et al.* [15] proposed a novel method for assessing airline passenger satisfaction. Their approach involved collecting tweets from Twitter and subsequently extracting relevant features. They employed three different machine learning models, namely artificial neural networks (ANN), SVM, and convolutional neural networks (CNN). The results indicated that CNN outperformed the other models, which is consistent with the expectation that convolutional networks excel at analyzing and extracting pertinent information from datasets.

In this context, the advent of deep learning techniques has opened up new horizons for improving the measurement of passenger satisfaction. Deep learning (DL), a subset of artificial intelligence, has demonstrated unparalleled capabilities in processing vast amounts of data, detecting subtle patterns, and making accurate predictions. By harnessing the power of deep learning, the aviation industry can transform its approach to understanding and enhancing passenger satisfaction. But the problem of deep learning approaches is the limitations of dealing with complex tabular data in order to get insights which require a dedicated deep neural for tabular data instead of implementing classical deep learning models.

This paper aims to delve into the realm of aviation passenger satisfaction measurement by leveraging cutting-edge deep learning techniques. These techniques can be applied to a variety of data sources, such as passenger reviews, social media sentiment analysis, and operational data, to provide a holistic and real time assessment of passenger satisfaction. By doing so, airlines and aviation stakeholders can gain invaluable insights into passenger preferences, identify areas for improvement, and ultimately enhance the overall passenger experience. As the aviation industry continues to evolve and adapt to changing passenger expectations, this paper stands as a vital contribution to the field, offering a roadmap for a more data-driven and responsive approach to passenger satisfaction measurement. By harnessing the potential of deep learning, we endeavor to not only improve the quality of air travel but also ensure that passengers' voices and experiences are at the forefront of aviation innovation [16]–[18].

Our research will encompass an advanced approach based on TabNets architecture, including data collection, preprocessing, and the development of deep learning models tailored to the aviation context. We will also evaluate the performance of these models against traditional satisfaction measurement methods to highlight the advantages of adopting deep learning techniques dedicated to tabular data. Numerous researchers have delved into the realm of passenger satisfaction utilizing both machine learning (ML) [5], [13]–[15] and DL [13], [15], [19], [20] techniques within the existing literature. It is noteworthy that DL-based solutions consistently outperform traditional ML models. This is not surprising, as deep learning models are renowned for their heightened accuracy and stability in extracting and analyzing key features critical for classification. Moreover, they excel in uncovering intricate patterns within datasets, translating into impressive real-world deployment results. However, a common limitation in these previous studies is the absence of a model specifically designed to handle tabular data. To address this gap, our study employs an advanced approach-TabNets meticulously crafted for tabular data analysis. Tab- Net stands as a neural network architecture explicitly tailored for processing structured data, commonly encountered in databases and spreadsheets. Its value proposition lies in its dual capability of providing high performance predictive modeling while offering interpretable insights into feature importance.

TabNet achieves this unique balance through its architecture, which combines the interpretability of decision trees with the flexibility of deep learning. Leveraging sparse attention mechanisms, TabNet selectively focuses on subsets of input features at each decision step. This efficient approach enables it to handle highdimensional data effectively, mitigating the risk of overfitting. Additionally, TabNet's robustness extends to its capability to handle missing data gracefully. The results of all our experiments, which involved employing various configurations for both the encoder and decoder steps, were aimed at attaining significant improvements in the performance of the TabNet network; In fact, we implement RF, DT, KNN, MLP, LR and NB to compare classical machine learning models performances (RF, DT, KNN, LR and NB), neural network model (MLP) performances and TabNet performances. Results show in terms of accuracy, F1-score and G-mean score how TabNet excel significantly all of these ML models, only RF and MLP give results close to TabNet. While comparing our results to the state-of-the-art papers in terms of accuracy, precision, recall and AUC, we find out that our TabNet model outperforms also all of these studies. We can say that this model can be proposed in real-life situations that require real-time analysis of passenger's satisfaction.

The rest of this paper is structures as follow: A related work section where we present state of the art papers that study airline passengers satisfaction using machine learning and deep learning techniques. A material and methods section where we describe the dataset used in this study and a theorical background of the model and metrics used in this study. A results and discussion section where we provide results of this study and the analysis and interpretation of these results, and finally a section dedicated for conclusion followed by this paper's references.

2. MATERIALS AND METHODS

2.1. Dataset description

In our work we use the "airline passenger satisfaction" dataset that encompasses 129,880 samples with 24 features. Designed to explore passenger experiences and satisfaction with airline services, these features are divided into personal information (e.g., age, gender), travel details (e.g., flight distance, arrival delay), and service ratings (e.g., WiFi quality, food, entertainment). This Kaggle-hosted dataset offers insights into factors driving passenger contentment and dissatisfaction in the airline industry [21], [22].

The dataset comprises several variables as listed below:

- Satisfaction variable, which serves as the target variable. It categorizes customers into two groups, satisfied or dissatisfied with the airline's services.
- Gender: Identifies customer gender (Male/Female).
- Customer Type: Classifies as "Loyal" or "Disloyal" based on previous interactions.
- Age: Provides insights into passenger age distribution.
- Type of Travel: Categorizes as "Business" or "Personal" travel.
- Class: Indicates service class (Economy, Eco Plus, Business).
- Flight Distance: Quantifies flight distance in miles.
- Departure Arrival Delay in Minutes: Quantifies flight delay in minutes.

– Furthermore, several variables gauge specific aspects of the customer experience, each rated on a scale from 0 to 5. How are: seat comfort, departure/arrival satisfaction with flight times, food and drink, gate location, inflight Wi-Fi service, inflight entertainment, online support, ease of online booking, on-board service, leg room service, baggage handling, check-in service, cleanliness, online boarding, ease of online booking, on-board service, leg room service, baggage handling, check-in service, cleanliness, online boarding, seat comfort, departure/arrival time convenience, food and drink, gate location, infight Wi-Fi service, Inflight Entertainment, "Departure delay in minutes" and "Arrival delay in minutes".

Table 1 showcases the class distribution across train and test datasets. This balanced distribution, with 45,025 instances of "Satisfied" and 58,879 instances of "Neutral or dissatisfied" classes, averts the need for pre-training balancing techniques. It underscores the dataset's suitability for training models, ensuring robust performance and generalization without class imbalance concerns.

Table 1. Class distribution in training and test sets

Class	Train	Test
Satisfied	45025	11403
Neutral or dissatisfied	58879	14573

2.2. TabNet classifier

Tabular data holds great importance across a wide range of industries, including healthcare, finance, banking, retail, and marketing. Arik and Pfister provide an innovative and interpretable canonical architecture, harnessing the inherent capabilities of deep neural networks. This innovative methodology combines the advantages of unsupervised pre-training, making it easier to predict hidden features, with the power of supervised learning to improve the effectiveness of classification and predictive tasks. It harnesses sequential attention mechanisms to strategically select relevant features during each decision step. This not only fosters interpretability but also promotes more efficient learning by focusing the learning capacity on the most significant features. The TabNet [23] classifier architecture consists of an encoder that uses a sequence of decision steps, encompassing feature transformations and attention mechanisms. Its main role is to discern and highlight the most informative attributes of the input data. And a decoder that takes the representation from the encoder and reconstructs the features.

2.3. TabNet encoder

Figure 1 shows the architecture of TabNet Encoder. The initial step involves utilizing the dataset without any feature engineering. Subsequently, a series of fundamental operations are applied, which include batch normalization, feature transformation, and data splitting. Following this, a repetitive sequence of steps unfolds, comprising several operations in each step: attentive transformation, masking, feature transformation, data splitting, and concluding with the application of the rectified linear unit (ReLU). The determination of the number of steps is contingent upon the available computational capacity, meaning that the number of steps is influenced by the system's computational resources

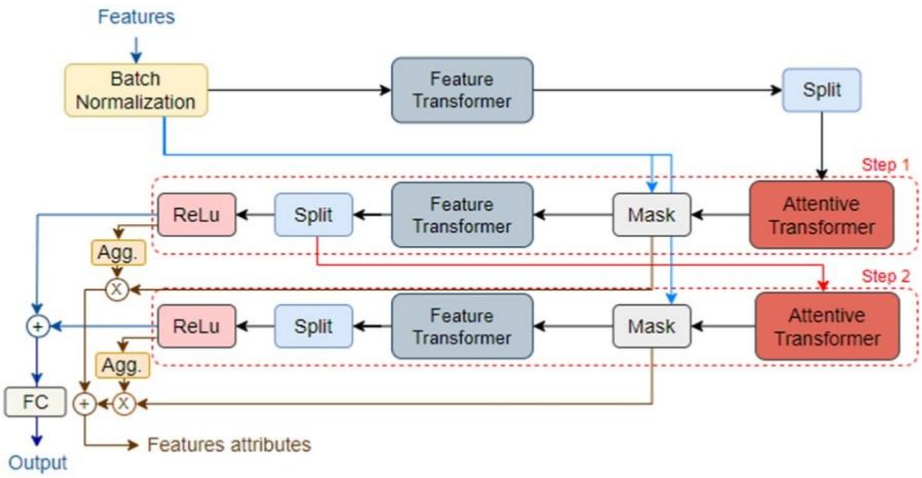

Figure 1. TabNet encoder architecture

2.3.1. Feature transformer

The inclusion of the feature transformer within the TabNet architecture plays a significant role in augmenting the model's capacity to acquire a richer and more informative data representation. It facilitates the model's capability to discern intricate, non-linear relationships inherent in the tabular data. The feature transformer comprises two parts: one that is shared across decision steps and one that is decision step-dependent. Each of these components is composed of fully connected layers, batch normalization, and gated linear units.

In Figure 2, the representation illustrates the constituents of the Feature Transformer, which comprises a total of four layers of gated linear unit (GLU) blocks, the 2 GLU blocks should be shared and 2 should be independent to enhance the robustness and efficiency of the learning process. In GLU block [24], the input is split into two parts along its last dimension, a nonlinear transformation is applied using the sigmoid function $\sigma(x_2)$ for the second part, and the output is the element-wise product between the first part and the output of the sigmoid activation:

$$GLU(x) = x_1 \otimes \sigma(x_2) \tag{1}$$

Furthermore, after each block, a normalization step with a scaling factor of $\sqrt{0.5}$ is applied. This operation contributes to stability in the training process, ensuring that the variance of the model's activations remains within reasonable bounds.

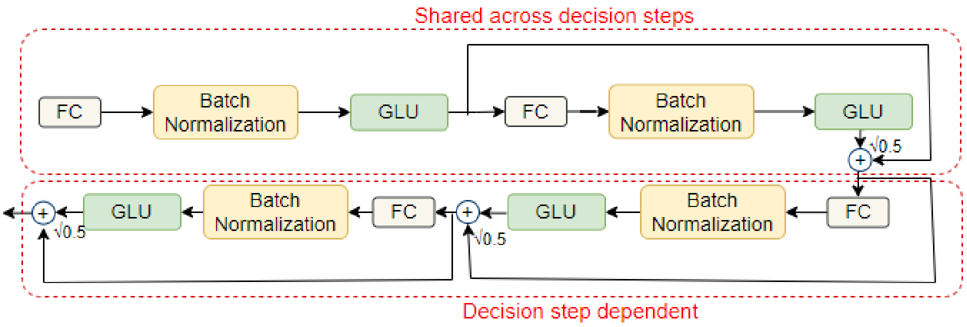

Figure 2. Feature transformer

2.3.2. Attentive transformer

Following the completion of the Feature Transformer stage, the resultant output pass to Attentive Block. This particular block plays a pivotal role in the feature selection process. In Figure 3, a representation elucidates the internal components comprising the Attentive Block. It encompasses a Fully Connected layer, a Batch Normalization layer, a layer dedicated to Prior Scales, and a Sparsemax layer [25] employed for the purpose of coefficient normalization. The employment of the Sparsemax layer facilitates the sparse selection of salient features, thereby enhancing the interpretability and efficiency of the feature selection process. The Sparsemax function is defined as (2)

$$Sparsemax(z_i) = max\{z_i - \tau, 0\} \tag{2}$$

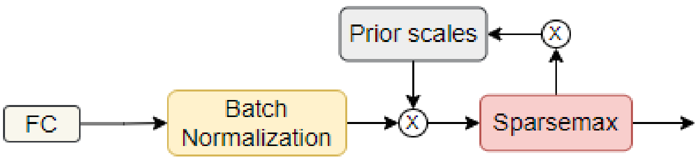

Figure 3. Attentive transformer

where the threshold τ is calculated using formula (3)

$$\tau(z) = ((\sum_{j \le k(z)} z_j) - 1)/k(z) \tag{3}$$

And k(z) is the maximum index (from the sorted set of input z) that meets this condition (4).

$$k(z) := max\{k\epsilon[K] \mid 1 + kz_{(k)} > \sum_{j\leq k} z_{(j)}\} \quad (4)$$

2.3.3 Attention mask

Subsequent to the attentive transformer step, the output is directed towards an attention mask. This mask serves a crucial role in identifying the selected features, enabling the model to quantify the overall importance of these features while also conducting a detailed analysis at each step of the process. When it becomes necessary to combine the masks from various steps, a coefficient is introduced to weigh the relative importance of each step in the decision-making process. For calculating the aggregate decision contribution for a sample b^{th} decision step i, the following formula is used (5)

$$\eta_b[i] = \sum_{c=1}^{N_d} ReLU(d_{b,c}[i]) \quad (5)$$

where $d_{b,c}[i]$ is the output of feature c for the sample b.

2.4. TabNet decoder

The Figure 4 shows the TabNet decoder architecture. It’s distinguished by its composition, beginning with a feature transformer, which is subsequently followed by fully connected layers operating within the decision step. The results of this operation are then subjected to summation with the reconstructed features. The decoder is composed of feature transformers, followed by FC layers at each decision step. The outputs are summed to obtain the reconstructed features [23].

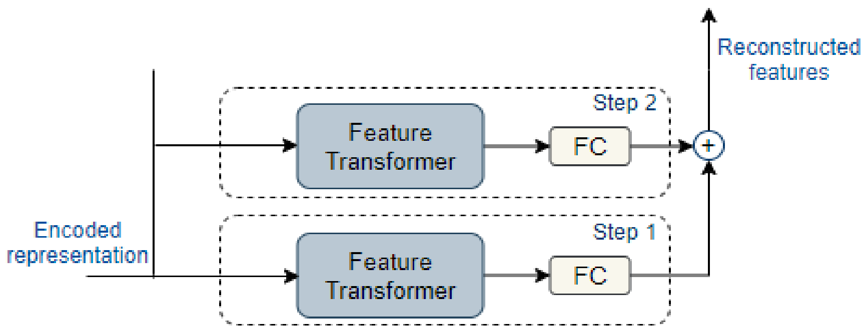

Figure 4. TabNet decoder architecture

2.5. Model evaluation metrics

Evaluating machine learning models is a crucial step to comprehensively evaluate their performance. Relying on a single metric may not provide a complete picture, as a model can perform well on one metric but poorly on another. Furthermore, the data set is imbalanced which means relying only on accuracy may lead to distorted results. For that purpose, in our study, we use a variety of metrics, including Accuracy, F1-score, g-mean score, precision, recall, and AUC, to comprehensively evaluate and demonstrate the performance of our models.

2.5.1. Accuracy

Accuracy is a common evaluation metric used to assess the performance of a classification model. It measures the proportion of correctly classified instances out of the total instances in the dataset:

$$Accuracy = (T_p + T_n)/(T_p + T_n + F_p + F_n) \quad (6)$$

where T_p and T_n are the true positive and negative; and F_p and F_n are the false positive and negative.

2.5.2. F1-score

The F1 score is an alternative evaluation metric in machine learning. It offers a more detailed perspective on a model’s performance by considering its performance on individual classes rather than providing an overall assessment, as accuracy does.

$$F_1 = 2.(Precision \times Recall)/(Precision + Recall) \tag{7}$$

where precision is:

$$Precision = T_p/(T_p + F_p) \tag{8}$$

and recall is:

$$Recall = T_p/(T_p + F_n) \tag{9}$$

2.5.3. G-Mean score

The G-Mean, also known as the Geometric Mean score, is an assessment metric that takes into account both sensitivity and specificity in a classification model. It plays a crucial role in achieving a balance between accurately identifying positive and negative cases, making it particularly useful when dealing with imbalanced datasets.

$$G - Mean = \sqrt{Recall \times Specificity} \tag{10}$$

where

$$Specificity = T_n/(F_p + T_n) \tag{11}$$

2.5.4. AUC

The AUC is a vital metric in binary classification tasks. It quantifies a model's ability to discriminate between positive and negative classes by measuring the area under the receiver operating characteristic (ROC) curve. A higher AUC signifies better classification accuracy, making it a valuable tool for model evaluation and comparison. It provides a concise summary of a classifier's overall performance across various decision thresholds, simplifying the assessment process in machine learning.

3. RESULTS AND DISCUSSION

3.1. Performances analysis of TabNet and implemented machine learning models

To demonstrate the performance of TabNet models, we conducted a comprehensive comparison of TabNet's results against several classical machine learning models, each evaluated using the following key metrics: accuracy, F1 score, and g-mean score. Table 2 shows the result obtained. Among the models tested, the TabNetClassifier stands out as an exceptional contender, boasting the highest scores across all three measures. Notably, the TabNetClassifier achieves an impressive precision rate of 96.47%, signaling that a substantial portion of instances are correctly classified. This high precision underscores the model's ability to effectively discern the various classes within the dataset. Furthermore, the TabNetClassifier maintains its superiority in terms of the F1 score, a metric harmonizing precision and recall, reaching a score of 96.41%. This highlights the model's equilibrium between accurate positive predictions and the comprehensive capture of actual positive instances. Similarly, the TabNetClassifier's geometric score of 96.24% reflects its capacity to achieve a harmonious balance between precision and recall through the geometric mean. The consistent performances of the RF and MLP models across all three metrics also underscore their robustness, with the RF achieving an accuracy of 96.28%, an F1 score of 96.21%, and a geometric score of 96.05%, while the MLP attains 96.20%, 96.13%, and 96.04%, respectively. Although the DT, KNN, and Naïve Bayes models display slightly lower scores, they still present credible outcomes, showcasing their ability to predict with precision. Notably, LR lags behind other models with an accuracy of 81.70%. This relatively lower accuracy might arise from the inherent simplicity and linearity of LR, suggesting that for this specific dataset, more complex models like TabNetClassifier, RF, and MLP are better suited to capture intricate patterns and complexities. In summary, the collective results collectively imply that the TabNetClassifier shines as a versatile and dependable model for predicting airline passenger satisfaction.

Figure 5 illustrates the comparative performance of seven machine learning models across the accuracy, F1-Score, and G-Mean score for the classification of passenger's satisfaction. This figure reinforces the findings from the previous tables, emphasizing that TabNet is the clear leader in terms of predictive performance for this particular classification task. Its consistently higher values across all metrics indicate its superiority over the other models considered in this study.

TabNet's superior performance in this study can be attributed to several key factors. Its architectural complexity, the featuring attention mechanisms, enables it to discern intricate patterns and relationships within the dataset, allowing it to prioritize relevant information and capture non-linear dependencies effectively. We also have the regularization techniques such as feature dropout and sparsity constraints prevent overfitting, contributing to the model's robust generalization. Furthermore, the possibility of ensemble learning amplifies its capabilities. Furthermore, effective feature engineering tailored to the dataset's characteristics likely played a role in optimizing TabNet's performance.

Table 2. Performances comparison between TabNet and machine learning models

	Accuracy	F1-Score	G-Mean Score
Random Forest	96.28	96.21	96.05
Decision Tree	94.73	94.65	94.67
KNN	92.55	92.38	92.04
MLP	96.20	96.13	96.04
LR	81.70	86.84	86.66
Naïve Bayes	86.12	85.81	85.61
TabNet Classifier	96.47	96.41	96.24

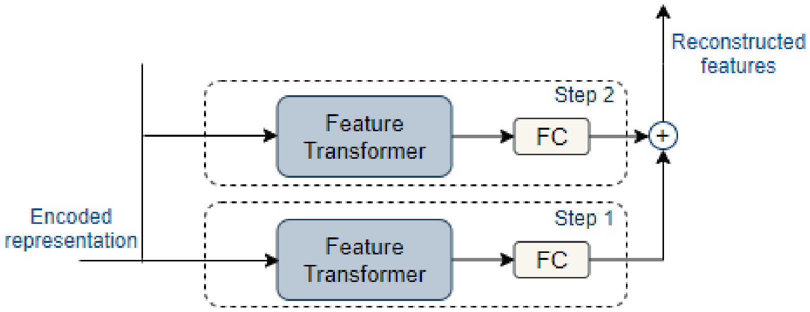

Figure 5. Visualization depicting TabNet alongside other implemented machine learning models

3.2. Comparative performances analysis of TabNet and the state-of-the-art studies

Table 3 shows the results of the established comparative study between our developed model and state of the art studies that treated the problem of analyzing and classifying the satisfaction of airline passengers. Results show that TabNet outperforms all other models in terms of accuracy, precision, recall, and AUC. It achieves the highest accuracy and precision while maintaining a high recall rate and AUC score, indicating strong overall performance. RF also demonstrates strong performance, with high accuracy, precision, and a respectable AUC score. However, TabNet slightly outperforms it in most metrics. PSO + NB performs better than GA + NB in terms of Accuracy, Precision, and Recall. Additionally, it provides an AUC score of 92.3%, indicating good discriminatory power. GA + NB achieves the lowest performance among the models, with the lowest Accuracy, Precision, and Recall scores. We can say that the results suggest that TabNet is the top-performing model in this comparative study, offering the highest accuracy and precision while maintaining a good balance between recall and AUC. In summary, we can say that the results suggest that TabNet appears to be the preferred choice for the classification of passenger's satisfaction, as it consistently achieves the highest performance across multiple metrics compared to the six implemented machine learning models. Furthermore, TabNet is the best model in this comparative study with highest values of the use metrics.

Table 3. Performances comparison between TabNet and state of art studies

	Accuracy	Precision	Recall	AUC
GA + NB [26]	85.99	87.91	87.43	-
PSO + NB [27]	86.13	87.90	87.29	92.3
RF	96.28	97.23	94.21	96.05
TabNet	96.44	97.57	94.31	96.24

4. CONCLUSION

In conclusion, this comprehensive study comparing the performance of various machine learning models for the classification of passenger's satisfaction shows the robustness of TabNet neural network in dealing with tabular data. Its consistent superiority in terms of Accuracy, F1-Score, and G-Mean score is attributed to its architectural complexity, attention mechanisms and non-linearity capturing abilities. TabNet's robustness to handle complex, high-dimensional datasets make it a compelling choice for a complex classification such as classifying airline passengers satisfaction. Having a dedicated deep learning model for tabular data helped in improving the performances of our classification task compared to other machine learning models and state of the art studies. In this study we only on one dataset, we propose for in our future work to add other data sets in order to test if out model is data independent and can deal with different kind of data sets. As a perspective arising from this study, we suggest the development of a more robust deep learning model for handling tabular data in the context of satisfaction classification that addresses the challenges associated with training and testing time efficiency to ensure the creation of a high-performance model that optimizes resource consumption. Striking a balance between model performance and computational efficiency will be crucial for practical applications and scalability.

FUNDING INFORMATION

We declare on our honor that this study was carried out without any external funding.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Rachid Kaidi	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				✓
Hicham Omara		✓		✓	✓	✓		✓	✓	✓	✓			
Mohamed Lazaar						✓			✓			✓	✓	
Mohammed Al Achhab												✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study will be available in Airlines customer satisfaction. *https://www.kaggle.com/datasets/sjleshrac/airlines-customer-satisfaction*.

REFERENCES

[1] "Carbon offsetting and reduction scheme for international aviation (CORSIA)," *ICAO*. https://www.icao.int/environmental-protection/CORSIA/Pages/default.aspx (accessed Sep. 01, 2024).
[2] S. Gössling and A. Humpe, "The global scale, distribution and growth of aviation: Implications for climate change," *Global Environmental Change*, vol. 65, Nov. 2020, doi: 10.1016/j.gloenvcha.2020.102194.
[3] X. Zhang and S. Mahadevan, "Aircraft re-routing optimization and performance assessment under uncertainty," *Decision Support Systems*, vol. 96, pp. 67–82, Apr. 2017, doi: 10.1016/j.dss.2017.02.005.
[4] X. Zhang and S. Mahadevan, "Ensemble machine learning models for aviation incident risk prediction," *Decision Support Systems*, vol. 116, pp. 48–63, Jan. 2019, doi: 10.1016/j.dss.2018.10.009.
[5] H. Jiang and Y. Zhang, "An assessment of passenger experience at Melbourne Airport," *Journal of Air Transport Management*, vol. 54, pp. 88–92, Jul. 2016, doi: 10.1016/j.jairtraman.2016.04.002.
[6] M. Arif, A. Gupta, and A. Williams, "Customer service in the aviation industry – an exploratory analysis of UAE airports," *Journal of Air Transport Management*, vol. 32, pp. 1–7, Sep. 2013, doi: 10.1016/j.jairtraman.2013.05.001.
[7] W.-H. Tsai, W. Hsu, and W.-C. Chou, "A gap analysis model for improving airport service quality," *Total Quality Management & Business Excellence*, vol. 22, no. 10, pp. 1025–1040, Oct. 2011, doi: 10.1080/14783363.2011.611326.
[8] R. Merkert and A. G. Assaf, "Using DEA models to jointly estimate service quality perception and profitability – Evidence from international airports," *Transportation Research Part A: Policy and Practice*, vol. 75, pp. 42–50, May 2015, doi: 10.1016/j.tra.2015.03.008.
[9] E. Ruiz, W. F. Yushimito, L. Aburto, and R. de la Cruz, "Predicting passenger satisfaction in public transportation using machine learning models," *Transportation Research Part A: Policy and Practice*, vol. 181, Mar. 2024, doi: 10.1016/j.tra.2024.103995.
[10] T. Noviantoro and J.-P. Huang, "Investigating airline passenger satisfaction: Data mining method," *Research in Transportation*

Business & Management, vol. 43, Jun. 2022, doi: 10.1016/j.rtbm.2021.100726.
[11] B. H. Hayadi, J.-M. Kim, K. Hulliyah, and H. T. Sukmana, "Predicting airline passenger satisfaction with classification algorithms," *IJIIS: International Journal of Informatics and Information Systems*, vol. 4, no. 1, pp. 82–94, Mar. 2021, doi: 10.47738/ijiis.v4i1.80.
[12] A. Nurdina and A. B. I. Puspita, "Naïve Bayes and KNN for airline passenger satisfaction classification: Comparative analysis," *Journal of Information System Exploration and Research*, vol. 1, no. 2, Jul. 2023, doi: 10.52465/joiser.v1i2.167.
[13] X. Jiang, Y. Zhang, Y. Li, and B. Zhang, "Forecast and analysis of aircraft passenger satisfaction based on RF-RFE-LR model," *Scientific Reports*, vol. 12, no. 1, Jul. 2022, doi: 10.1038/s41598-022-14566-3.
[14] M. S. Homaid and I. Moulitsas, "Measuring airport service quality using machine learning algorithms," in *2022 The 6th International Conference on Advances in Artificial Intelligence*, Oct. 2022, pp. 8–14. doi: 10.1145/3571560.3571562.
[15] S. Kumar and M. Zymbler, "A machine learning approach to analyze customer satisfaction from airline tweets," *Journal of Big Data*, vol. 6, no. 1, Dec. 2019, doi: 10.1186/s40537-019-0224-1.
[16] S. Deb, M. Ali Ahmed, and D. Das, "Service quality estimation and improvement plan of bus Service: A perception and expectation based analysis," *Case Studies on Transport Policy*, vol. 10, no. 3, pp. 1775–1789, Sep. 2022, doi: 10.1016/j.cstp.2022.07.008.
[17] L. Dell'Olio, A. Ibeas, and P. Cecin, "The quality of service desired by public transport users," *Transport Policy*, vol. 18, no. 1, pp. 217–227, Jan. 2011, doi: 10.1016/j.tranpol.2010.08.005.
[18] C. Zhang, Y. Liu, W. Lu, and G. Xiao, "Evaluating passenger satisfaction index based on PLS-SEM model: Evidence from Chinese public transport service," *Transportation Research Part A: Policy and Practice*, vol. 120, pp. 149–164, Feb. 2019, doi: 10.1016/j.tra.2018.12.013.
[19] S.-H. Park, M.-Y. Kim, Y.-J. Kim, and Y.-H. Park, "A deep learning spproach to analyze sirline customer propensities: The case of South Korea," *Applied Sciences*, vol. 12, no. 4, Feb. 2022, doi: 10.3390/app12041916.
[20] H. B. Sankaranarayanan, B. V Vishwanath, and V. Rathod, "An exploratory analysis for predicting passenger satisfaction at global hub airports using logistic model trees," in *2016 Second International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, Sep. 2016, pp. 285–290. doi: 10.1109/ICRCICN.2016.7813672.
[21] C. C. E. van Geest, Y. Wan Yit, Z. T. Gouliev, and K. Quille, "Predictive models with XAI: A comparative study of enhancing airline customer satisfaction," in *Proceedings of the 2023 Conference on Human Centered Artificial Intelligence: Education and Practice*, Dec. 2023, pp. 36–41. doi: 10.1145/3633083.3633189.
[22] S. Jana, "Airlines customer satisfaction," *Kaggle*. https://www.kaggle.com/datasets/sjleshrac/airlines-customer-satisfaction (accessed Dec. 02, 2024).
[23] S. Ö. Arik and T. Pfister, "TabNet: Attentive interpretable tabular learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 8, pp. 6679–6687, May 2021, doi: 10.1609/aaai.v35i8.16826.
[24] Y. N. Dauphin, A. Fan, M. Auli, and D. Grangier, "Language modeling with gated convolutional networks," *Prepr. arXiv.1612.08083*, Dec. 2016.
[25] A. F. T. Martins and R. F. Astudillo, "From softmax to sparsemax: A sparse model of attention and multi-label classification," *Prepr. arXiv.1602.02068*, Feb. 2016.
[26] Y. Religia and D. Maulana, "Genetic algorithm optimization on Nave Bayes for airline customer satisfaction classification," *JISA(Jurnal Informatika dan Sains)*, vol. 4, no. 2, pp. 121–126, Dec. 2021, doi: 10.31326/jisa.v4i2.925.
[27] A. Nurdina and A. B. I. Puspita, "Naive Bayes and KNN for airline passenger satisfaction classification: comparative analysis," *Journal of Information System Exploration and Research*, vol. 1, no. 2, pp. 527–533, Jul. 2023, doi: 10.52465/joiser.v1i2.167.

BIOGRAPHIES OF AUTHORS

Rachid Kaidi is currently a doctoral student at the Faculty of Sciences of Tetouan, Abdelmalek Essaâdi University, Morocco. He received his master's degree from Mohammed VI International Academy of Civil Aviation Casablanca Morocco. His research focuses on artificial intelligence. He can be contacted at email: kaidirachid1982@gmail.com.

Hicham Omara is currently a Professor of Computer Science in the Department of Computer Science at the Polydisciplinary Faculty of Taza, Sidi Mohamed Ben Abdellah University, Morocco. He received his PhD from the Faculty of Science at Abdelmalek Essaâdi University, Morocco. He has participated in numerous national and international scientific and organizational committees. His research interests include artificial intelligence, image processing, big data, and other areas. He can be contacted at email: h.omara@uae.ac.ma.

Mohamed Lazaar currently works at the Department of Computer Sciences, National Higher School of Computer Science and Systems Analysis, Mohammed V University. Mohamed does research in applied mathematics, statistics and computer engineering. He has participated in numerous national and international scientific and organizational committees. Their current project is ‘machine learning, big data and other areas’. He can be contacted at email: mohamed.lazaar@ensias.um5.ac.ma.

Mohammed Al Achhab is currently a professor at the National School of Applied Sciences in Tetouan. He is a member of numerous scientific associations, including IEEE, innove, and mocit. He has participated in numerous national and international scientific and organizational committees. His research focuses on business process analysis and validation, natural language processing, case study Arabic, and adaptive e-learning. He can be contacted at email: m.alachhab@uae.ac.ma