https://doi.org/10.31449/inf.v50i15.7339

Developing a Predictive Model Predicting the Missing Appointment Using Deep Learning Algorithms

Dalia Fadl
Business technology Department, Canadian international College, Giza, Egypt
Dalia.sayed.43@gmail.com,dalia_fadl@cic-cairo.com

Keywords: Deep learning, demographics, predictive modelling, neural networks, appointment

Recieved: October 15, 2024

In recent years, the use of predictive modelling techniques has gained significant attention in healthcare systems, aiming to improve operational efficiency and optimize resource allocation. One crucial area where predictive modelling can be applied is in predicting missing appointments, as missed appointments can lead to increased costs and inefficiencies in healthcare settings. This study focuses on developing a predictive model that uses deep learning algorithms to forecast the likelihood of a patient missing their scheduled appointment. Deep learning, a subset of machine learning, has shown promising results in various domains due to its ability to automatically extract relevant features from complex data. By using deep learning techniques, we aim to enhance the accuracy and reliability of predicting missed appointments. The predictive model is trained on a dataset containing patient demographics, historical appointment records, and other relevant features. Through the use of deep learning architectures such as recurrent neural networks (RNNs) or long short-term memory networks (LSTMs), the model learns intricate temporal dependencies and subtle patterns in the data, enabling the correct prediction of missed appointments. The study contributes to the advancement of predictive analytics in healthcare by proving the potential of deep learning algorithms in improving appointment scheduling efficiency and resource allocation.

Povzetek: Študija obravnava uporabo globokega učenja za napovedovanje izostankov pacientov z zdravstvenih pregledov ter izboljšanje učinkovitosti naročanja in razporejanja virov.

1 Introduction

Predicting and managing appointments is crucial in various domains, including healthcare, business, and service industries. However, missed appointments pose significant challenges, leading to inefficiencies, lost revenue, and compromised patient or customer satisfaction. To address this issue, developing a correct predictive model that can estimate the likelihood of appointment no-shows has become an area of active research. Traditional statistical and machine learning approaches have been employed to predict appointment attendance. However, these methods often struggle to capture the complex patterns and dependencies present in appointment data. Deep learning algorithms, on the other hand, have appeared as powerful tools for modelling intricate relationships in large-scale datasets and have shown promising results in various prediction tasks [1].

This research aims to develop a predictive model using deep learning algorithms to forecast the likelihood of missing appointments. By using the inherent capabilities of deep learning, we look to uncover hidden patterns and dependencies within appointment data that can enhance the accuracy and reliability of predictions. The use of deep learning algorithms offers several advantages in this context. First, deep learning models can automatically learn hierarchical representations of data, enabling them to capture intricate patterns and dependencies. This is particularly relevant in appointment prediction, where many factors such as time, demographics, and historical appointment records may influence attendance [2].

Second, deep learning models have the ability to process sequential and temporal data effectively. By using recurrent neural networks (RNNs) or other temporal modelling techniques, the predictive model can learn from the sequential nature of appointment data, incorporating the temporal dynamics that affect attendance patterns. Furthermore, deep learning models can manage diverse types of data, including structured, unstructured, or multimodal data [3]. This flexibility allows the integration of various features, such as demographic information, textual data from appointment notes, or even visual data from images or videos, to enrich the predictive model's inputs. The outcomes of this research hold enormous potential for improving appointment management systems. By accurately predicting the likelihood of missing appointments, healthcare providers, businesses, and service industries can proactively distribute resources, perfect scheduling, and implement targeted interventions to mitigate the negative impact of no-shows [4].

In this paper, we present the method, experimental results, and analysis of our predictive model. We evaluate the performance of different deep learning algorithms, considering various evaluation metrics, and discuss the insights gained from our experiments. Additionally, we discuss the implications and potential applications of our findings, along with areas for future research to further advance appointment prediction using deep learning algorithms [5].

1.1 Problem statement

The problem addressed in this research is the prediction of missing appointments in various domains, such as healthcare, business, and service industries. Missed appointments lead to inefficiencies, lost revenue, and reduced customer or patient satisfaction. Traditional statistical and machine learning approaches often struggle to capture the complex patterns and dependencies present in appointment data. Therefore, there is a need to develop a correct predictive model that can expect the likelihood of appointment no-shows [6].

1.2 Objective

The aim of this research is to develop a predictive model using deep learning algorithms to forecast the likelihood of missing appointments. By using the capabilities of deep learning, the model aims to uncover hidden patterns and dependencies within appointment data to enhance the accuracy and reliability of predictions. Specifically, the aims are as follows:

- Design and implement a deep learning-based predictive model for appointment prediction.
- Explore and use proper deep learning algorithms, such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), or their variants, to capture the complex relationships in appointment data.
- Preprocess and transform the appointment dataset, considering factors such as time, demographics, and historical appointment records, to extract meaningful features for the predictive model.
- Train and perfect the predictive model using proper techniques, such as gradient descent, regularization, and hyperparameter tuning, to achieve the best performance.
- Evaluate the performance of the predictive model using suitable evaluation metrics, such as accuracy, precision, recall, F1-score, or area under the receiver operating characteristic (ROC) curve.

By achieving these aims, this research aims to provide a valuable contribution to the development of correct and dependable predictive models for predicting missing appointments.

2 Literature review

2.1 Missed appointments in healthcare

Missed appointments, often referred to as "no-shows," present a substantial challenge in healthcare settings. They lead to inefficiencies, increased operational costs, and adverse impacts on patient care. Understanding the many factors contributing to missed appointments is essential for developing effective strategies to mitigate this issue.
Factors Contributing to Missed Appointments

2.1.1 Patient demographics and socioeconomic status

Age and Gender: Research shows that younger patients and certain gender demographics have higher rates of missed appointments. Younger individuals may have less predictable schedules or competing commitments such as education or early career responsibilities [7]. Gender-related responsibilities, such as caregiving roles more often associated with women, can also influence appointment adherence.
Income and Education: Lower-income individuals and those with limited educational backgrounds often miss appointments due to financial constraints, a lack of transportation, or insufficient health literacy. Financial difficulties can make it challenging to afford travel costs or take time off work, while lower health literacy may affect one's understanding of the importance of appointments and adherence. Ethnicity and Cultural Factors: Cultural attitudes towards healthcare and systemic barriers can also influence appointment adherence. For instance, some ethnic groups may face language barriers or distrust in the healthcare system, which affects their likelihood of attending scheduled appointments [8].

2.1.2 Health-related factors

- Chronic Illness: Patients with chronic conditions may experience higher no-show rates due to the physical and emotional burden of their illness. Factors such as fatigue, depression, and complex medication regimens can hinder their ability to attend appointments consistently.
- Mental Health: Individuals with mental health conditions, including anxiety and depression, are more prone to missing appointments. These conditions can lead to a lack of motivation, organizational challenges, and avoidance behaviors.
- Earlier Attendance History: Past behavior is a strong predictor of future behavior. Patients with a history of missed appointments are more likely to miss future ones, suggesting a pattern that can be found and addressed through predictive modelling. Appointment Scheduling Practices [9].
- Appointment Lead Time: The interval between scheduling an appointment and the appointment date can significantly affect attendance. Longer wait times can lead to higher no-show rates as patients may forget, lose interest, or find alternative solutions.
- Reminder Systems: Effective reminder systems, such as automated phone calls, SMS, and emails, have been shown to reduce no-show rates. However, the effectiveness can vary based on the mode of communication and patient preferences.
- Appointment Type and Time: The nature and timing of the appointment also affect attendance. Routine check-ups might have higher no-show rates compared to urgent care visits. Similarly, appointments scheduled during inconvenient times, such as early mornings or late afternoons, may see higher no-show rates.

2.1.3 Addressing the issue

Traditional methods to reduce missed appointments include reminder systems, patient education, and flexible scheduling practices. However, these strategies often fall short due to their reactive nature. The advent of predictive modelling, particularly through machine learning and deep learning, offers an initiative-taking approach by finding high-risk patients before they miss appointments. By understanding and addressing the factors contributing to missed appointments, healthcare providers can develop targeted interventions to improve patient adherence, reduce costs, and enhance overall healthcare efficiency [10].

Missed appointments are a multifaceted problem with significant implications for healthcare systems. By using data on patient demographics, health status, and scheduling practices, predictive models can offer a nuanced understanding of why appointments are missed. This knowledge can inform more effective strategies to improve patient attendance, ultimately leading to better health outcomes and more efficient use of healthcare resources [11].

2.2 Predictive modelling in healthcare

Predictive modelling in healthcare has gained prominence as a vital tool for addressing various challenges, including disease prediction, patient readmission, and appointment no-shows. These models use historical data to find patterns and make informed predictions about future outcomes, enabling healthcare providers to intervene proactively and perfect their operations. The evolution of predictive modelling has seen the application of various machine learning techniques, each contributing unique strengths and facing specific limitations [12].

2.2.1 Traditional machine learning techniques

Traditional machine learning techniques such as logistic regression, decision trees, and support vector machines (SVM) have been foundational in predictive modelling within healthcare. Logistic Regression: This statistical model is often used for binary classification problems, such as predicting the likelihood of a patient attending an appointment. It is valued for its simplicity and interpretability, enabling healthcare professionals to understand the impact of various factors on the predicted outcome. However, its linear nature can limit its ability to capture more complex, non-linear relationships in data [13].

Decision Trees: Decision trees offer a more intuitive approach by recursively splitting the data into subsets based on feature values. They provide clear decision rules that are easy to interpret and visualize. Despite their simplicity and ease of use, decision trees can be prone to overfitting, especially when dealing with noisy data, which can lead to reduced generalization performance. Support Vector Machines: SVMs are powerful for classification tasks, especially in high-dimensional spaces. They work by finding the best hyperplane that separates different classes with the largest margin. While SVMs can manage non-linear relationships through the use of kernel functions, they can be computationally intensive and less interpretable than simpler models like logistic regression.

2.2.2 Deep learning algorithms

Deep learning algorithms, a subset of machine learning, have revolutionized predictive modelling in healthcare by offering the ability to model intricate patterns and relationships within data. These algorithms, particularly neural networks, have several advantages over traditional methods:

Neural Networks: Neural networks consist of layers of interconnected nodes (neurons) that can learn complex representations of data. By adjusting the weights of these connections through backpropagation, neural networks can capture non-linear and high-dimensional patterns that traditional models might miss [14]. This makes them particularly effective for tasks like image and speech recognition, as well as predicting outcomes in healthcare, where the relationships between variables can be overly complex.

Convolutional Neural Networks (CNNs): Originally developed for image processing, CNNs have been adapted for various healthcare applications, such as analyzing medical images (e.g., MRI, CT scans) to detect anomalies or diseases. Their ability to automatically detect relevant features without extensive pre-processing has led to significant improvements in diagnostic accuracy. Recurrent Neural Networks (RNNs) and Long Short-Term Memory Networks (LSTMs): These architectures are designed for sequential data, making them ideal for time-series analysis and predicting patient outcomes based on historical health records [15]. LSTMs, in particular, address the vanishing gradient problem in standard RNNs, allowing them to capture long-term dependencies and trends in patient data over time.

2.3 Benefits and challenges

The integration of deep learning into predictive modelling in healthcare brings many benefits, including higher predictive accuracy and the ability to manage large and complex datasets. These models can uncover hidden patterns and insights that traditional techniques might overlook, leading to more effective interventions and improved patient outcomes. For instance, predictive models can find patients at elevated risk of missing appointments, enabling healthcare providers to implement targeted reminder systems or support services to reduce no-show rates. However, deep learning models also present challenges. They need substantial computational resources and copious amounts of labelled data for training, which can be a barrier in some healthcare settings. Moreover, the complexity of these models often comes at the expense of interpretability, making it difficult for healthcare professionals to understand the decision-making process and trust the model's predictions fully. Ensuring data privacy and security is also critical, given the sensitive nature of healthcare data [17]. Predictive modelling using both traditional machine learning and advanced deep learning algorithms has become a cornerstone in modern healthcare. While traditional techniques provide a solid foundation and are easier to interpret, deep learning models offer superior performance by capturing complex patterns and relationships in data. The ongoing challenge lies in balancing the need for accuracy with the requirement for interpretability, ensuring that these models can be effectively integrated into healthcare practices to enhance patient care and operational efficiency. As computational power and data availability continue to grow, the potential for predictive modelling in healthcare will only expand, promising even greater advancements in the future.

2.4 Deep learning algorithms

Deep learning, a subset of machine learning, has revolutionized the field by using neural networks with multiple layers to learn hierarchical representations of data. Unlike traditional machine learning algorithms that require manual feature extraction, deep learning models can automatically discover intricate patterns and relationships in large datasets. This ability has led to remarkable advancements across various domains, including image recognition, natural language processing, and time series prediction [16]

2.5 Feedforward neural networks

Feedforward neural networks (FNNs) are the simplest type of artificial neural network. They consist of an input layer, one or more hidden layers, and an output layer. Each neuron in a layer is connected to every neuron in the later layer, with data flowing in one direction—from input to output. FNNs are particularly effective for tasks where the relationship between input and output is relatively straightforward, such as basic classification and regression problems. Despite their simplicity, they can model complex functions by stacking multiple layers and employing non-linear activation functions [17]. However, FNNs can struggle with more complex data structures, such as sequential or spatial data, which has led to the development of more specialized architectures.

2.6 Convolutional Neural Networks (CNNs)

Convolutional neural networks (CNNs) are designed to manage spatial data, making them exceptionally powerful for image processing tasks. CNN typically consists of multiple convolutional layers, pooling layers, and fully connected layers. Convolutional layers apply a series of filters to the input data to detect various features, such as edges, textures, and shapes. Pooling layers reduce the spatial dimensions of the data, which helps to decrease computational load and mitigate overfitting [22]. CNNs can automatically learn hierarchical feature representations from raw image pixels, enabling them to excel in tasks like image classification, object detection, and segmentation. Their success in these areas has made CNNs a cornerstone in the development of advanced computer vision applications.

2.7 Recurrent Neural Networks (RNNs)

Recurrent neural networks (RNNs) are specifically designed for sequential data, making them suitable for tasks that involve time series, natural language processing, and any data where temporal or sequential patterns are important. Unlike FNNs, RNNs have connections that form directed cycles, allowing information to persist across different time steps. This architecture enables RNNs to keep a memory of earlier inputs, which is crucial for understanding context in sequences. However, traditional RNNs suffer from issues like vanishing and exploding gradients, which make training them on long sequences challenging. To address these issues, more advanced variants such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) were developed [21]

2.8 Long Short-Term Memory Networks (LSTMs) and Gated Recurrent Units (GRUs)

LSTMs and GRUs are enhancements to the standard RNN architecture that improve the network’s ability to capture long-term dependencies. LSTMs incorporate memory cells and gating mechanisms that regulate the flow of information, effectively mitigating the vanishing gradient problem. These mechanisms include input gates, forget gates, and output gates, which control the input, retention, and output of information, respectively. GRUs simplify the LSTM architecture by combining the input and forget gates into a single update gate, while keeping comparable performance. Both LSTMs and GRUs have been successfully applied to a range of sequential data tasks, such as language modelling, machine translation, and speech recognition [23].

2.9 Advantages and applications

The primary advantage of deep learning algorithms is their ability to automatically learn features from raw data, which reduces the need for manual feature engineering [20]. This capability is particularly beneficial when dealing with large, complex datasets where traditional methods might fall short. Deep learning models have proved superior performance in various applications:

- Image Recognition: CNNs have achieved ultramodern results in image classification tasks, enabling applications like autonomous driving, medical image analysis, and facial recognition.
- Natural Language Processing (NLP): RNNs, LSTMs, and GRUs have been instrumental in advancing NLP tasks, including language translation, sentiment analysis, and text generation.
- Time Series Prediction: RNNs and their variants excel in forecasting and predicting trends from temporal data, making them invaluable in fields such as finance, weather forecasting, and supply chain management.

2.10 Challenges

Despite their advantages, deep learning algorithms also present several challenges. They need copious amounts of labelled data for training, which can be resource-intensive to obtain. The models are computationally demanding, often needing specialized hardware such as GPUs for efficient training. Additionally, the interpretability of deep learning models stays a significant concern, as their complex architectures can make it difficult to understand how they arrive at specific predictions. Ensuring the ethical use of these models, especially in sensitive domains like healthcare, adds another layer of complexity [24]

Deep learning algorithms have transformed predictive modelling by enabling the automatic extraction of complex patterns from large datasets. Feedforward neural networks, convolutional neural networks, and recurrent neural networks, along with their advanced variants, have shown immense potential in various applications, from image recognition to natural language processing. While these models offer significant advantages, their implementation requires careful consideration of data requirements, computational resources, and ethical implications. As research and development in deep learning continue to progress, these models are poised to become even more integral to predictive modelling and beyond

3 Methodology

3.1 Import libraries

popular libraries such as NumPy, pandas, scikit-learn, and PyTorch for data processing, visualization, and building machine learning models. In addition, tools such as imbalanced learning and joblib are used to address specific challenges in machine learning projects, such as dealing with imbalanced datasets and implementing parallel computing tasks

3.2 Data description

The dataset provided has information on 110,527 medical appointments, each associated with various characteristics. The primary focus of this dataset is to decide whether a patient showed up for their scheduled appointment or was a no-show. By examining the different variables associated with each appointment, patterns and relationships can be found to predict the likelihood of a patient not attending their appointment

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, StandardScaler, RobustScaler, MinMaxScaler
from sklearn.preprocessing import QuantileTransformer, PowerTransformer
from imblearn.pipeline import Pipeline
from imblearn.over_sampling import SMOTE
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import RandomizedSearchCV
from sklearn.metrics import make_scorer, f1_score, classification_report
import torch
import torch.nn as nn
from torch.utils.data import Dataset, DataLoader
from tqdm import tqdm
from itertools import product
from joblib import Parallel, delayed
from sklearn.metrics import classification_report
```

Figure 1: libraries function from models

The dataset includes essential variables such as Patient and AppointmentID, which serve as unique identification numbers for each patient and appointment, respectively. Gender is another variable that shows whether the patient is male or female. The dates of DataMarcacaoConsulta and Data Amendment represent the actual appointment date and the date when the appointment was scheduled or registered, respectively.

The patient's age is included as a variable to analyze if age influences the likelihood of being a no-show. The neighborhood variable provides information about the neighborhood where the appointment takes place, which can help find any geographic patterns or influences on attendance. One crucial variable in the dataset is Scholarship, which shows whether the patient is enrolled

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 110527 entries, 0 to 110526
Data columns (total 14 columns):
 #   Column          Non-Null Count   Dtype  
---  ------          --------------   -----  
 0   PatientId       110527 non-null  float64
 1   AppointmentID   110527 non-null  int64  
 2   Gender          110527 non-null  object 
 3   ScheduledDay    110527 non-null  object 
 4   AppointmentDay  110527 non-null  object 
 5   Age             110527 non-null  int64  
 6   Neighbourhood   110527 non-null  object 
 7   Scholarship     110527 non-null  int64  
 8   Hipertension    110527 non-null  int64  
 9   Diabetes        110527 non-null  int64  
 10  Alcoholism      110527 non-null  int64  
 11  Handcap         110527 non-null  int64  
 12  SMS_received    110527 non-null  int64  
 13  No-show         110527 non-null  object 
dtypes: float64(1), int64(8), object(5)
memory usage: 11.8+ MB
```

Figure 2: Data set information

in the Bolsa Familia program. Bolsa Familia is a government social welfare program in Brazil aimed at providing financial aid to poor families. This variable allows for an investigation into whether participation in this program affects appointment attendance. The presence of hypertension, diabetes, alcoholism, and disability are represented by the variables Hypertension, Diabetes, Alcoholism, and Handcap, respectively. These variables help assess if any of these conditions influence the likelihood of a patient being a no-show. The SMS_received variable shows whether one or more text messages were sent to remind the patient about their appointment. This variable allows for an examination of the effectiveness of SMS reminders in improving attendance rates. Finally, the No-show variable serves as the target variable, showing whether the patient was a no-show (True) or attended the appointment (False). By analyzing the relationships between this variable and the other characteristics, predictive models can be built to decide the factors that contribute to appointment no-show.

3.3 Model training and evaluation

A machine learning pipeline is defined using scikit-learn's Pipeline module. This pipeline consists of two main steps: oversampling using the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance, followed by training a Random Forest classifier. The hyperparameters of the Random Forest classifier are specified using a parameter grid, which includes various options such as the number of estimators, maximum depth of trees, maximum features to consider, minimum number of samples required to split a node, minimum number of samples required at each leaf node, and whether bootstrap samples are used when building trees. To find the best combination of hyperparameters, a

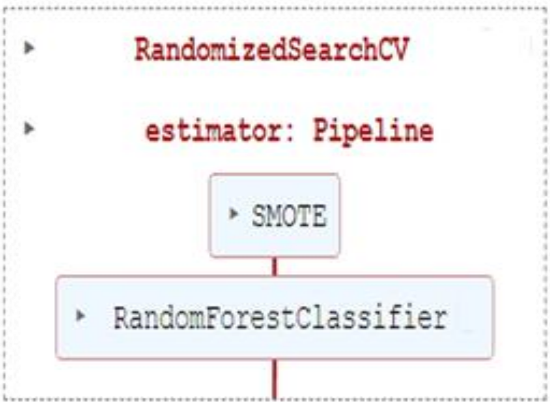

Figure 3: Random Forest classifier algorithms

Randomized Search object is instantiated. This object performs a randomized search over the parameter grid, evaluating each combination of hyperparameters using cross-validation. The scoring metric used for evaluation is the F1-score, with a positive label of 1 showing the minority class (which may be the class of interest in imbalanced dataset.

4 Results

This process allows for the identification of the best model among the trained models based on the specified evaluation metric (in this case, the F1-score), which can then be further analyzed or used for inference on new data. The classification report provides a comprehensive evaluation of the performance of the machine learning model on the test dataset. It includes metrics such as precision, recall, and F1-score for each class, as well as overall accuracy. In this case, the model achieved an accuracy of 63%, showing that it correctly classified 63% of the instances in the test set. For class 0, which likely is the negative class, the precision is 90%, showing that when the model predicts a negative outcome, it is correct 90% of the time. However, the recall is 60%, suggesting that the model misses finding approximately 40% of the actual negative instances in table 1 The F1-score, which is the harmonic mean of precision and recall, is 72%, providing a balanced measure of the model's performance for class 0. For class 1, which likely is the positive class, the precision is 32%, showing that when the model predicts a positive outcome, it is correct 32% of the time. The recall is higher at 73%, suggesting that the model is better at finding actual positive instances compared to negatives. However, the low precision shows a high rate of false positives. The F1-score for class 1 is 44%, showing a moderate balance between precision and recall. Overall, while the model proves decent performance for class 0 (negative), it struggles with class 1 (positive), particularly in terms of precision. This imbalance in performance between the two classes could show issues with class imbalance, model complexity, or feature representation. Further analysis and refinement of the model may be necessary to improve its overall performance and address the class imbalance issue

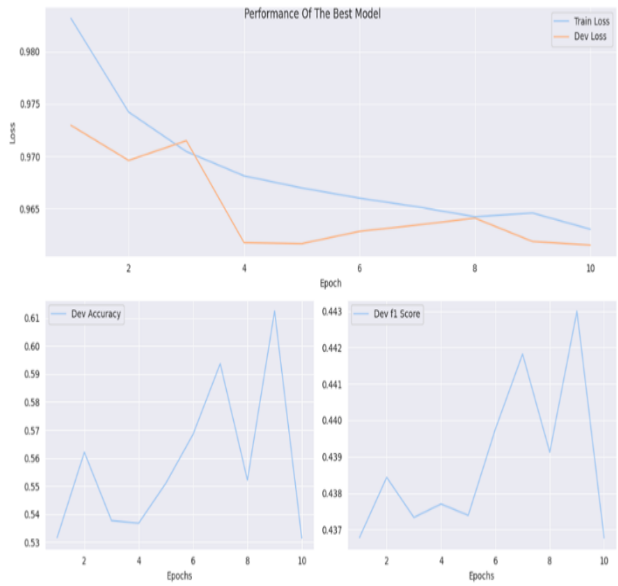

Figure 4: The classification report provides score

The classification report provided evaluates the performance of a machine learning model on a test dataset, encompassing metrics such as precision, recall, and F1-score for each class, along with overall accuracy. In this case, the model achieved an accuracy of 53%, showing its correct classification of 53% of the instances in the test set. For class 0 (likely being the negative class), the precision is 95%, showing that when the model predicts a negative outcome, it is correct 95% of the time. However, the recall is relatively low at 43%, suggesting that the model misses finding approximately 57% of the actual negative instances.

Table 1: Confusion matrix of models

	precision	**recall**	**f1-score**	**support**
0	0.9	0.61	0.73	17642
1	0.32	0.72	0.44	4464
accuracy			0.63	22106
macro avg	0.61	0.67	0.59	22106
weighted avg	0.78	0.63	0.67	22106

The F1-score, which is the harmonic mean of precision and recall, is 59%, offering a balanced measure of the model's performance for class 0. For class 1 (likely being the positive class), the precision is lower at 29%, showing that when the model predicts a positive outcome, it is correct only 29% of the time. However, the recall is much higher at 91%, suggesting that the model is better at finding actual positive instances compared to negatives. The F1-score for class 1 is 44%, showing a moderate balance between precision and recall. The model demonstrates significant class imbalance, with much lower precision for class 1 compared to class 0. This imbalance could show issues with the model's ability to correctly find positive instances. The reported metrics offer insights into the strengths and weaknesses of the model and can guide further analysis and improvement efforts to enhance its performance.

Table 2: Confusion matrix of models

	precision	**recall**	**f1-score**	**support**
0	0.95	0.43	0.59	17642
1	0.29	0.91	0.44	4464
accuracy			0.53	22106
macro avg	0.62	0.67	0.51	22106
weighted avg	0.82	0.53	0.56	22106

5 Discussion

The comparison between the two classification reports offers valuable insights into the performance of machine learning models on different datasets. The first report shows a model with an accuracy of 63%, achieving a relatively balanced F1-score of 72% for the negative class and 44% for the positive class. However, the second report shows a model with a lower accuracy of 53% but with higher precision and recall for the positive class (29% precision and 91% recall). The discrepancies in performance between the two models highlight the importance of considering various evaluation metrics, especially in the context of imbalanced datasets. While the first model proves better overall accuracy and balance between precision and recall for both classes, it struggles with finding positive instances accurately. On the other hand, the second model achieves higher recall for the positive class, showing its ability to capture more true positive instances, albeit with lower precision. The class imbalance issue is clear in both reports, with the second model particularly affected by the imbalance, leading to a significantly lower precision for the positive class. This underscores the need for proper handling of class imbalance during model training and evaluation, such as using techniques like oversampling (e.g., SMOTE) or adjusting class weights to mitigate the impact of class imbalance on model performance. the discussion emphasizes the importance of interpreting classification reports comprehensively, considering various metrics, and addressing class imbalance effectively to develop robust machine learning models for real-world applications. Further exploration and refinement of the models based on these insights can lead to improved performance and better decision-making capabilities in practical scenarios.

6 Conclusion

This paper proposed a predictive model to predict missing appointment using recurrent neural networks. The comparison of two classification reports highlights the nuanced performance of machine learning models on different datasets. While one model achieves higher overall accuracy and balanced performance across classes, it struggles with finding positive instances accurately. Conversely, the other model shows higher recall for the positive class but at the cost of lower precision, showing challenges with class imbalance. These findings underscore the importance of considering multiple evaluation metrics, especially in the context of imbalanced datasets. Addressing class imbalance through techniques like oversampling or adjusting class weights is crucial for improving model performance and ensuring reliable predictions, particularly for minority classes. In practical applications, interpreting classification reports comprehensively and understanding the trade-offs between precision and recall is essential for informed decision-making. Further refinement and optimization of machine learning models based on these insights can lead to more effective solutions for real-world problems.

References

[1] Alshammari, A., Almalki, R., & Alshammari, R. (2021). Developing a predictive model of predicting appointment no-show by using machine learning algorithms. Journal of Advances in Information Technology Vol, 12(3). https://doi.org/10.12720/jait.12.3.234-239

[2] Dashtban, M., & Li, W. (2022). Predicting non-attendance in hospital outpatient appointments using a deep learning approach. Health Systems, 11(3), 189-210. https://doi.org/10.1080/20476965.2021.1924085

[3] Salazar, L. H., Fernandes, A., Dazzi, R., Garcia, N., & Leithardt, V. R. (2020). Using different models of machine learning to predict attendance at medical appointments. Journal of Information Systems Engineering and Management, 5(4), em0122. https://doi.org/10.29333/jisem/8430

[4] Chong, L. R., Tsai, K. T., Lee, L. L., Foo, S. G., & Chang, P. C. (2020). Artificial intelligence predictive analytics in the management of outpatient MRI appointment no-shows. American Journal of Roentgenology, 215(5), 1155-1162. https://doi.org/10.2214/ajr.19.22594

[5] Alshammari, R., Daghistani, T., & Alshammari, A. (2020). The prediction of outpatient no-show visits by using deep neural network from large data. International Journal of Advanced Computer Science and Applications, 11(10). https://doi.org/10.14569/ijacsa.2020.0111066

[6] Salah, H., & Srinivas, S. (2022). Predict, then schedule: Prescriptive analytics approach for machine learning-enabled sequential clinical scheduling. Computers & Industrial Engineering, 169, 108270. https://doi.org/10.1016/j.cie.2022.108270

[7] Dunstan, J., Villena, F., Hoyos, J. P., Riquelme, V., Royer, M., Ramírez, H., & Peypouquet, J. (2023). Predicting no-show appointments in a pediatric hospital in Chile using machine learning. Health Care Management Science, 26(2), 313-329. https://doi.org/10.1007/s10729-022-09626-z

[8] Joseph, J., Senith, S., Alfred Kirubaraj, A., & Ramson, S. J. (2023, August). Feature Selection Techniques to Enhance Prediction of Clinical Appointment No-Shows Using Neural Network. In International Conference on Artificial Intelligence on Textile and Apparel (pp. 275-285). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-99-8479-4_20

[9] Almutairi, T. H., & Olatunji, S. O. (2024). The use of AI in healthcare to predict no-shows for dental appointments: A case study conducted in Saudi Arabia. Informatics in Medicine Unlocked, 46, 101472. https://doi.org/10.1016/j.imu.2024.101472

[10] Alaidah, A., Alamoudi, E., Shalabi, D., AlQahtani, M., Alnamshan, H., & Abubacker, N. F. (2021). Mining and predicting no-show medical appointments: Using hybrid sampling technique. In Communication and Intelligent Systems: Proceedings of ICCIS 2020 (pp. 315-333). Springer Singapore. https://doi.org/10.1007/978-981-16-1089-9_27

[11] Joseph, J., Senith, S., Kirubaraj, A. A., & Ramson, S. J. (2024, June). Predicting missed appointments in dental clinic: Enhanced risk assessment with multiple learning techniques. In Advances in AI for Biomedical Instrumentation, Electronics and Computing: Proceedings of the 5th International Conference on Advances in AI for Biomedical Instrumentation, Electronics and Computing (ICABEC-2023), 22–23 December 2023, India (p. 13). CRC Press. https://doi.org/10.1201/9781032644752-3

[12] Aladeemy, M., Adwan, L., Booth, A., Khasawneh, M. T., & Poranki, S. (2020). New feature selection methods based on opposition-based learning and self-adaptive cohort intelligence for predicting patient no-shows. Applied Soft Computing, 86, 105866. https://doi.org/10.1016/j.asoc.2019.105866

[13] Giunta, D. H., Huespe, I. A., Alonso Serena, M., Luna, D., & Gonzalez Bernaldo de Quirós, F. (2023). Development and validation of nonattendance predictive models for scheduled adult outpatient appointments in different medical specialties. The International Journal of Health Planning and Management, 38(2), 377-397. https://doi.org/10.1002/hpm.3590

[14] Babayoff, O., Shehory, O., Geller, S., Shitrit-Niselbaum, C., Weiss-Meilik, A., & Sprecher, E. (2022). Improving Hospital Outpatient Clinics Appointment Schedules by Prediction Models. Journal of Medical Systems, 47(1), 5. https://doi.org/10.1007/s10916-022-01902-3

[15] Ahmadi, E., Garcia-Arce, A., Masel, D. T., Reich, E., Puckey, J., & Maff, R. (2019). A metaheuristic-based stacking model for predicting the risk of patient no-show and late cancellation for neurology appointments. IISE Transactions on Healthcare Systems Engineering, 9(3), 272-291. https://doi.org/10.1080/24725579.2019.1649764

[16] Nguyen, P. T., Dang, D. T., & Nguyen, V. D. (2023, September). A Robust Deep Learning Techniques for No-Show Prediction in Hospital Appointments. In International Conference on Advanced Intelligent Systems and Informatics (pp. 3-18). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43247-7_1

[17] Maskew, M., Sharpey-Schafer, K., De Voux, L., Crompton, T., Bor, J., Rennick, M., ... & Pisa, P. (2022). Applying machine learning and predictive modeling to retention and viral suppression in South African HIV treatment cohorts. Scientific reports, 12(1), 12715. https://doi.org/10.1038/s41598-022-16062-0

[18] Raja, R., Mukherjee, I., & Sarkar, B. K. (2021). A machine learning-based prediction model for preterm birth in rural India. Journal of Healthcare

Engineering, 2021. https://doi.org/10.1155/2021/6665573

[19] Taylor, R. A., Pare, J. R., Venkatesh, A. K., Mowafi, H., Melnick, E. R., Fleischman, W., & Hall, M. K. (2016). Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data–driven, machine learning approach. Academic Emergency Medicine, 23(3), 269-278. https://doi.org/10.1111/acem.12876

[20] Kunze, K. N., Karhade, A. V., Sadauskas, A. J., Schwab, J. H., & Levine, B. R. (2020). Development of machine learning algorithms to predict clinically meaningful improvement for the patient-reported health state after total hip arthroplasty. The Journal of Arthroplasty, 35(8), 2119-2123. https://doi.org/10.1016/j.arth.2020.03.019

[21] Shour, A. R., Jones, G. L., Anguzu, R., Doi, S. A., & Onitilo, A. A. (2023). Development of an evidence-based model for predicting patient, provider, and appointment factors that influence no-shows in a rural healthcare system. BMC Health Services Research, 23(1), 989. https://doi.org/10.1186/s12913-023-09969-5

[22] Farran, B., Channanath, A. M., Behbehani, K., & Thanaraj, T. A. (2013). Predictive models to assess risk of type 2 diabetes, hypertension and comorbidity: machine-learning algorithms and validation using national health data from Kuwait a cohort study. BMJ open, 3(5), e002457. https://doi.org/10.1136/bmjopen-2012-002457

[23] Pianykh, O. S., Guitron, S., Parke, D., Zhang, C., Pandharipande, P., Brink, J., & Rosenthal, D. (2020). Improving healthcare operations management with machine learning. Nature Machine Intelligence, 2(5), 266-273. https://doi.org/10.1038/s42256-020-0176-3

[24] Kotsiantis, S. B. (2012). Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades. Artificial Intelligence Review, 37, 331-344. https://doi.org/10.1007/s10462-011-9234-x