A Data-Driven Approach to Patient Satisfaction: Unveiling Key Influencers Through Machine Learning

Md. Mahafuzur Rahman1\*, Charls Darwin2, Md. Mushaddiqul Islam Amin1

1 Department of Statistics, Begum Rokeya University, Rangpur-5404, Bangladesh

2 Associate Professor, Department of Statistics, Begum Rokeya University, Rangpur-5404, Bangladesh

**Corresponding author:** Md. Mahafuzur Rahman

Gmail: [mahafuzur.brur@gmail.com](file:///C:\Users\User\Downloads\11111)

Abstract

**Background:** Patient satisfaction is a crucial measure of healthcare quality, influencing both health outcomes and care experiences. This study aims to identify the factors influencing patient satisfaction in healthcare facilities using machine learning algorithms due to their strong predictive capabilities.

**Methods:** A cross-sectional survey was conducted with 312 patients from two private hospitals in Rangpur, Bangladesh. Machine learning models, including LightGBM, Random Forest, XGBoost, and CatBoost, were used to predict patient satisfaction, and SHAP value analysis was employed for interpretation.

**Results:** The LightGBM classifier (before SMOTE) outperforms other models across metrics, with the highest accuracy, MCC, and ROC-AUC scores, establishing it as the optimal predictive model for this study. SHAP analysis further reveals that factors such as treatment plan, age, appointment ease, waiting time, and medication details significantly influence satisfaction levels. Overall, results indicate that structured interactions, shorter waiting times, and clear communication are associated with higher satisfaction, while extended waiting times and lack of decision involvement negatively impact patient experiences.

**Conclusion:** To enhance patient satisfaction, healthcare providers should prioritize improving communication, reducing wait times, and offering clear treatment plans. Future research should explore additional factors to refine predictive models. The LightGBM classifier provides valuable insights into the key determinants of patient satisfaction, enabling healthcare practitioners and researchers to utilize it for targeted prediction and analysis in their specific contexts.

Keywords: Patient satisfaction, machine learning, Bangladesh, LightGBM, SHAP

# Introduction

Patients' views and needs on the use of health services are reflected in their level of satisfaction. Evaluating patient satisfaction is crucial because it frequently aids in determining the standard of healthcare delivery and the responsiveness of the healthcare system in the absence of indicators of healthcare service quality. Better health outcomes are the consequence of higher patient satisfaction levels, which also reflect higher levels of patient empowerment, dedication to care, and adherence to advised management. Assessing patient satisfaction also aids in investigating care delivery, prioritizing resource allocation and capacity building requirements. Prior research examined this subject from the viewpoints of the healthcare system or the quality of care (Adhikary et al., 2018). The healthcare industry is essential to the growth of a nation because it provides a vital service for the development and upkeep of healthy human capital, which is necessary to achieve national goals. Globally, the healthcare industry is expanding quickly and becoming a very competitive sector of the service economy (Irfan & Farooq, 2012). Better compliance, continuity of care, and, eventually, improved health outcomes are all fostered by a patient's ability to form a deep and lasting relationship with their healthcare provider (Ferdousi, 2015). The sense of how someone feels about their experience in relation to their expectations is called satisfaction. For patients, it has to do with how well their demands for both general healthcare and care particular to their conditions are satisfied. Medical consultation is an integral aspect of any patient's sickness management plan. Clinically significant is the assessment of patients' satisfaction levels with health care, as contented patients are more likely to adhere to their treatment plans (Norhayati et al., 2017).

Despite having few resources, Bangladesh, a country in South Asia, tries to provide healthcare that meets international standards. Four main sectors make up the nation's healthcare delivery system: the public sector, the for-profit commercial sector, the not-for-profit private sector (which is mainly made up of NGOs), and several international development organizations (Begum et al., 2021). In the global healthcare scene, patient happiness is becoming more and more crucial. It is acknowledged that patient-perceived outcomes are reliable, significant, and conventional measures of the caliber of care. Patients keep complaints and concerns from doctors when they are unhappy with the way the doctor is treating them. Despite their opinions, patients often behave in a submissive manner during medical consultations, hence doctors are often uninformed of the patient satisfaction levels throughout consultations (Kuteyi et al., 2010) . In contrast to wealthy nations, medical professionals in Bangladesh are not taught the value of ethical behavior and effective communication during their training. Medical professionals employed in public hospitals typically treat individuals from lower socioeconomic backgrounds who have poor cleanliness and little awareness of health issues. In public outpatient clinics in Bangladesh, one of the biggest challenges faced by doctors is understanding their patients and helping them know (Jalil et al., 2017).

In our study, we used a machine learning approach to investigate patient satisfaction, where we applied an advanced methodology. Several biases (e.g., common method bias, non-response bias, and sample selection bias) may exist in survey-based research and influence the validity and accuracy in determining patient satisfaction (Ahmed, 2005). The prediction results generated by machine learning models demonstrate strong robustness and consistency across different samples within the same population (Shrestha et al., 2021). These attributes help to mitigate potential prediction errors caused by sampling variability. This study leverages machine learning to examine the factors influencing patient satisfaction from three perspectives: patient characteristics, interaction dynamics, and consultation volume. Specifically, we address the following research questions:

1. Patient characteristics
2. Doctor-patient interaction characteristics

However, in the Rangpur region, limited research has focused solely on detecting service quality, primarily using descriptive methods. In contrast, our study applies a machine learning approach, including SHAP analysis, to analyze patient satisfaction and incorporate doctor-patient interactions for deeper insights. Despite the lack of empirical research utilizing machine learning models to evaluate patient satisfaction with the quality of care in healthcare facilities across Rangpur, this study aims to fill that gap. By leveraging machine learning and SHAP analysis, we seek to identify the factors influencing patient satisfaction, providing a more comprehensive understanding. The results will not only contribute to the existing literature but also offer actionable insights for enhancing the quality of care in the region.

The research is divided into six sections. The first outlines the context and rationale, followed by a review of relevant literature in the second. The third details data collection and statistical methods, including machine learning and SHAP analysis. Empirical findings are presented in the fourth, with discussions in the fifth. The final section provides policy recommendations for improving healthcare quality and addresses the study's limitations.

# Literature Review

Patient satisfaction, as demonstrated by (Kuteyi et al., 2010) hinges on effective communication, trust, and information from physicians, with minimal influence from demographic factors. (Xu et al., 2022) explored how machine learning models can be used to predict patient satisfaction with doctors in online medical communities. The study uses data from the online platform Good Doctor and applies the XGBoost algorithm combined with the SMOTE technique to handle imbalanced data. The study identifies key factors influencing satisfaction, including doctors’ competence, online efforts, service quality, and patient treatment processes. (Liu et al., 2023) investigated the key factors influencing service satisfaction on online healthcare platforms. Using LightGBM highlights key determinants of patient satisfaction, including consultation volume, physician interactions (thank-you letters, gifts), and attributes like gender. It found that free consultations and long wait times lower satisfaction, while gifts and patient votes boost it.

(Al & Thesis, 2022) discussed patient satisfaction in telemedicine as a growing area of interest due to its potential to improve healthcare access and efficiency. The review also highlighted the importance of advanced techniques like machine learning, particularly tree-based algorithms such as Random Forest and XGBoost, in effectively identifying and ranking the determinants of patient satisfaction in telemedicine​. Patient satisfaction, a crucial health outcome influenced by provider communication and empathy, can be effectively identified and predicted using machine learning algorithms like Random Forest and XGBoost, ultimately enhancing healthcare service quality (Zhang et al., 2020). (Batbaatar et al., 2017) emphasized that patient satisfaction is a key indicator of healthcare quality, with factors like interpersonal care, communication, and empathy of healthcare providers being crucial determinants. Despite many studies, there is no globally accepted framework for measuring patient satisfaction, and results vary widely across studies due to differences in methods and contexts. (Aktar, 2021) highlighted that service quality significantly impacts patient satisfaction in healthcare settings. Various studies demonstrate that factors like reliability, responsiveness, empathy, assurance, and cost affect patient perceptions of service quality. Research shows that patient satisfaction can lead to improved health outcomes, higher retention rates, and better adherence to treatment plans. These determinants of service quality are crucial for enhancing healthcare systems, particularly in private hospital settings​.

(Clever et al., 2008) explored the relationship between physicians' communication behaviors and patient satisfaction with hospital care. Through an analysis using an instrumental variable (IV) approach, the study found a significant positive association between effective communication from physicians and overall patient satisfaction. This relationship holds even after adjusting for potential confounding patient factors, suggesting that improvements in physician communication could meaningfully enhance patients' perceptions of care quality. (Li et al., 2021) used the random forest algorithm to identify key factors impacting patient satisfaction, such as patients' right to know, timely nursing response, and staff service. The random forest model showed higher prediction accuracy compared to logistic regression and naive Bayes models. (Shin et al., 2024) highlighted how long waiting times are a key contributor to patient dissatisfaction in outpatient settings. Existing studies have employed linear models to identify features affecting waiting times, but these approaches often lack accuracy and interpretability. Recent advances in machine learning (ML), particularly interpretable machine learning (IML), have been applied to address these challenges, emphasizing the need for models that balance predictive accuracy with operational insights for service improvements. Using a random forest algorithm, the study highlighted that patient satisfaction is driven by demographics like age during registration, while physician attentiveness and behavior are key factors during consultations (Simsekler et al., 2021).

(Wu et al., 2023) emphasized that ICU healthcare professionals face high turnover intentions due to factors like long hours, burnout, and low salary satisfaction. Previous research has focused on these aspects using traditional methods, but this study uses XGBoost to reveal key predictors like income satisfaction, years of service, and night shift frequency, offering a more detailed analysis of turnover risks. (Alhashem et al., 2011) emphasized that patient satisfaction is crucial for assessing healthcare quality, influenced by factors like communication, accessibility, and care quality. Studies in the Gulf region highlight the impact of socio-demographic variables, such as age and nationality, on satisfaction, underscoring the need for targeted healthcare improvements​.

# Data and Methodology

## Data collection

A cross-sectional study was conducted on patients treated from October 22, 2023, to January 1, 2024, in the Rangpur region of Bangladesh, assessing satisfaction levels and identifying key determinants. Data was collected through a structured 11-question survey from two private hospitals in Rangpur, with a total of 312 responses gathered. An 8-member team of final-year B.Sc. honors statistics students from BRUR was recruited for data collection. Participants were thoroughly briefed on the study’s objectives and questionnaire details. They also underwent intensive training to ensure unbiased and meaningful data collection. Authorization from the diagnostic center manager ensured full cooperation, with research teams overseeing data collection across both hospitals. The study's results will provide valuable insights into patient satisfaction and help shape future healthcare service improvements in the region. Additionally, this research can serve as a foundation for more comprehensive studies on healthcare quality and patient care in Bangladesh.

## Methodology

The methodology involved a comprehensive machine learning approach to uncover key factors influencing patient satisfaction (Figure 1). After collecting and preprocessing the data, a train-test split (75% vs. 25%) was applied to ensure model validation. To address potential class imbalance, SMOTE oversampling was used, improving model robustness. Several advanced ensemble classifiers, including Random Forest, XGBoost, LightGBM, and CatBoost, were employed to predict patient satisfaction. Performance was measured using metrics like accuracy, recall, and AUC, and the best-performing model was selected after comparison. To enhance interpretability, SHAP value analysis was conducted, which identified the most impactful features influencing patient satisfaction. This comprehensive workflow (as illustrated in Figure 1) not only provided accurate predictions but also shed light on actionable factors that healthcare providers can target to improve patient experiences.

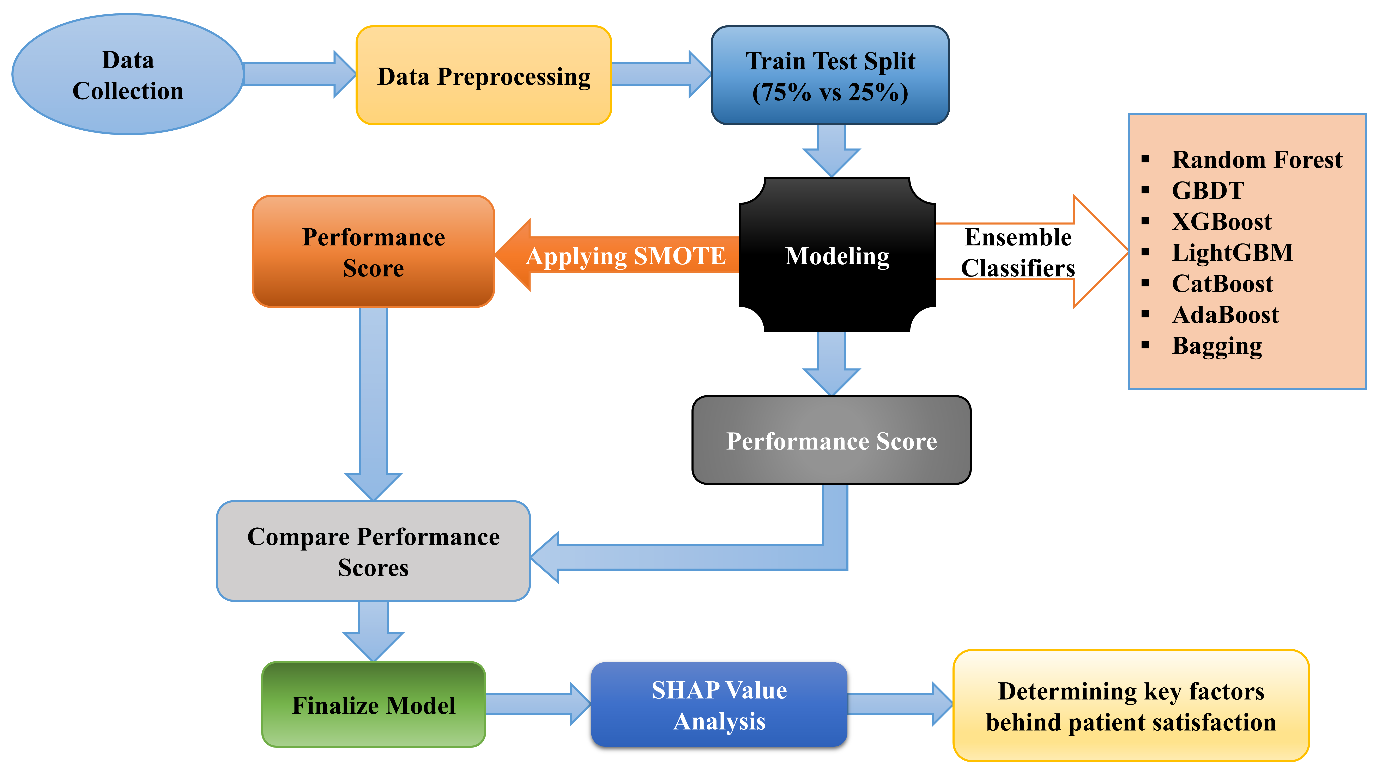


Figure : Analytical framework

## Baseline Approaches

To assess the performance of our proposed LightGBM algorithm, we conducted a comparative analysis with several other ensemble methods rooted in tree-based learning algorithms, including XGBoost, CatBoost, AdaBoost, Bagging, Random Forest, and GBDT. Decision trees, introduced by (Quinlan, 1986), structure data hierarchically, where each path from the root to a leaf represents a classification rule. However, a single decision tree can be prone to bias and instability. Ensemble methods, which combine multiple decision trees, offer a powerful solution by turning weak learners into a robust, high-performing model. Bagging and Boosting are two popular ensemble techniques. Bagging, as described by Fan and Zhang (2009), works by creating different subsets of data, training individual models on each subset, and then aggregating their predictions for the final output. Random Forest (Breiman, 2001) builds on this by introducing randomness in the selection of data and features to further enhance model accuracy. Boosting, on the other hand, incrementally builds a strong model by combining weak models through iterative refinement. Some of the most widely used Boosting algorithms include LightGBM (Ke et al., 2017), CatBoost (Prokhorenkova et al., 2018), AdaBoost (Freund & Schapire, 1997), GBDT (Friedman, 2000), and XGBoost (Chen & Guestrin, 2016). These approaches were chosen as baselines due to their shared foundation in decision trees, while their distinct mechanisms for generating final predictions allow for a diverse and meaningful comparison.

## Evaluation Metrics

In this study, we prioritize the Matthews Correlation Coefficient (MCC) as the primary evaluation metric for model selection, with AUC and accuracy scores as secondary criteria. The choice of MCC is driven by its robustness, particularly in imbalanced datasets, where traditional metrics like accuracy or F1 score can be misleading. MCC evaluates the full confusion matrix, incorporating true positives, true negatives, false positives, and false negatives, thus providing a balanced and holistic assessment of model performance. As highlighted by (Chicco & Jurman, 2020), accuracy and F1 score often give inflated results when faced with imbalanced data because they fail to account for misclassification rates in both classes, which can skew the model's perceived effectiveness. In contrast, MCC avoids this pitfall by ensuring a more nuanced evaluation, balancing the contributions of both classes, which is especially crucial in healthcare-related data analysis where minority class predictions may have significant consequences (Chicco & Jurman, 2020). Moreover, MCC offers advantages over commonly used metrics like the AUC, which, while popular, does not consider precision or negative predictive value—two critical factors in determining a model's real-world utility. A high AUC can mask poor precision, leading to overconfidence in a model's performance. MCC, however, generates high scores only when a classifier performs well across all critical metrics: sensitivity, specificity, precision, and negative predictive value. This makes MCC not only more reliable but also more indicative of a model's practical effectiveness (Chicco & Jurman, 2023). Given the complexity and stakes of healthcare data, MCC emerges as the ideal primary metric for our study, ensuring a comprehensive, fair, and insightful evaluation of machine learning models, while AUC and accuracy offer supplementary performance insights.

# Results

## Frequency Analysis

The frequency analysis in Table 1 provides a detailed view of patient demographics, doctor-patient interaction characteristics, and satisfaction levels. Age distribution is varied, with the largest group between 36-45 years (23.72%) and the smallest aged 46-55 (15.38%), and a near-equal gender balance of males (48.72%) and females (51.28%). Educational attainment shows that 37.18% of patients have higher education, while 16.67% are illiterate. Interaction characteristics reveal that 62.82% of patients find scheduling easy, though waiting time varies, with 70.19% waiting under an hour and 10.90% waiting over two hours, potentially impacting satisfaction. Treatment plans are widely provided (78.85%), and appointment lengths mostly span 10 minutes (38.46%), with some variation. Decision-making involvement is reported by 51.92% of patients, and while 60.26% receive detailed medication information, 19.23% report feeling neglected during visits. Patient satisfaction is generally high, with 63.46% satisfied; however, factors such as waiting times, decision-making involvement, and communication about medications emerge as potential areas for improvement to enhance overall satisfaction.

Table : An overview of frequency analysis of the variables

|  |  |  |  |
| --- | --- | --- | --- |
| **Categories** | **Main categories** | **Sub-categories** | **Frequency (percent)** |
| Patient characteristics | Age | 16-25 | 70 (22.44%) |
| 26-35 | 63 (20.19%) |
| 36-45 | 74 (23.72%) |
| 46-55 | 48 (15.38%) |
| 55+ | 57 (18.27%) |
| Gender | male | 152 (48.72%) |
| female | 160 (51.28%) |
| Education level | illiterate | 52 (16.67%) |
| primary | 49 (15.71%) |
| secondary | 95 (30.45%) |
| higher education | 116 (37.18%) |
| Interaction Characteristics | Appointment ease | yes | 196 (62.82%) |
| no | 116 (37.18%) |
| Waiting time (minutes) | 0-60 | 219 (70.19%) |
| 61-120 | 59 (18.91%) |
| 120+ | 34 (10.90%) |
| Treatment plan | yes | 246 (78.85%) |
| no | 66 (21.15%) |
| Visiting time (minutes) | 5 | 84 (26.92%) |
| 10 | 120 (38.46%) |
| 15 | 81 (25.96%) |
| 15+ | 27 (8.65%) |
| Decision involves | yes | 162 (51.92%) |
| no | 150 (48.08%) |
| Medicine details | yes | 188 (60.26%) |
| no | 124 (39.74%) |
| Ignore patient | yes | 60 (19.23%) |
| no | 252 (80.77%) |
| Target | Patient satisfaction | satisfied | 198 (63.46%) |
| dissatisfied | 114 (36.54%) |

## Performance Comparison

As mentioned before, the original data is imbalanced in that a large difference exists between the number of satisfied and dissatisfied samples. To address the issue of imbalanced samples, SMOTE (Synthetic Minority Over-sampling Technique) was used to create synthetic samples so that the categories in the original data (Satisfied and Dissatisfied) were not seriously imbalanced. Since parameters can influence the performance of prediction approaches, we adjusted the parameters to achieve the best results. For every approach, their performance was evaluated in terms of three metrics: MCC, ROC AUC score and testing data accuracy score. For every approach, we compare their performance in different contexts: before SMOTE and after SMOTE, using training or test set.

The evaluation results of four approaches are shown in Table 2. From Table 2, we can see that LightGBM (before SMOTE) can achieve the best prediction performance in all three contexts, especially in MCC score. As shown in the second column, CatBoost (before SMOTE) can achieve the best performance in that the accuracy score (0.86) of this approach is jointly highest along with LightGBM (before SMOTE) approach. However, the other two metrics of CatBoost (before SMOTE) is not as good as LightGBM (before SMOTE). As shown in the third column, AUC score of XGBoost (after SMOTE) is 0.835 which is slightly higher than the AUC score (0.833) of LightGBM classifier (before SMOTE). As the MCC score of LightGBM classifier (before SMOTE) is the highest among all the classifiers and other two performance are relatively good, we can consider the LightGBM classifier as our optimum model for this study.

Table : Performance scores before and after over sampling

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Accuracy** | **ROC-AUC** | **MCC** |
| Adaptive Boosting Classifier | 0.833333 | 0.746071 | 0.630346 |
| Adaptive Boosting Classifier (SMOTE) | 0.820513 | 0.723214 | 0.601338 |
| Bagging Classifier | 0.846154 | 0.791786 | 0.668492 |
| Bagging Classifier (SMOTE) | 0.833333 | 0.8175 | 0.631432 |
| Categorical Boosting Classifier | 0.858974 | 0.828929 | 0.6906 |
| Categorical Boosting Classifier (SMOTE) | 0.833333 | 0.828929 | 0.63524 |
| Extreme Gradient Boosting | 0.846154 | 0.816429 | 0.662292 |
| Extreme Gradient Boosting (SMOTE) | 0.846154 | 0.835357 | 0.671898 |
| Gradient Boosting Classifier | 0.846154 | 0.805357 | 0.662292 |
| Gradient Boosting Classifier (SMOTE) | 0.807692 | 0.803214 | 0.579062 |
| Light Gradient Boosting Machine | 0.858974 | 0.832857 | 0.696065 |
| Light Gradient Boosting Machine (SMOTE) | 0.846154 | 0.811786 | 0.659399 |
| Random Forest Classifier | 0.833333 | 0.806071 | 0.640714 |
| Random Forest Classifier (SMOTE) | 0.820513 | 0.8025 | 0.61 |

[**Note:** SMOTE is only applied in training set, the test data remains unchanged so that it correctly represents the original data.]

Figure 2 presents an insightful evaluation of the LightGBM classifier's predictive capability through a confusion matrix and ROC curve, showcasing the model’s robust performance in classifying patient satisfaction. The confusion matrix reveals a high accuracy of 85.9%, with 49 true positives and 18 true negatives, alongside relatively low false positive (10) and false negative (1) counts. This indicates that the model is adept at correctly identifying both satisfied and dissatisfied patients, while maintaining a low error rate. Notably, the minimal false negatives suggest the model’s strong reliability in not overlooking dissatisfied patients—a crucial aspect in healthcare settings where misclassification can obscure critical insights into patient experiences. Complementing the confusion matrix, the ROC curve provides further validation of the model's discriminative ability. The area under the curve (AUC) is 0.833, signifying a high degree of sensitivity and specificity across various classification thresholds. This AUC score underscores the model’s overall efficacy in distinguishing between satisfied and dissatisfied patients, reflecting its strong balance between minimizing false positives and maximizing true positives.

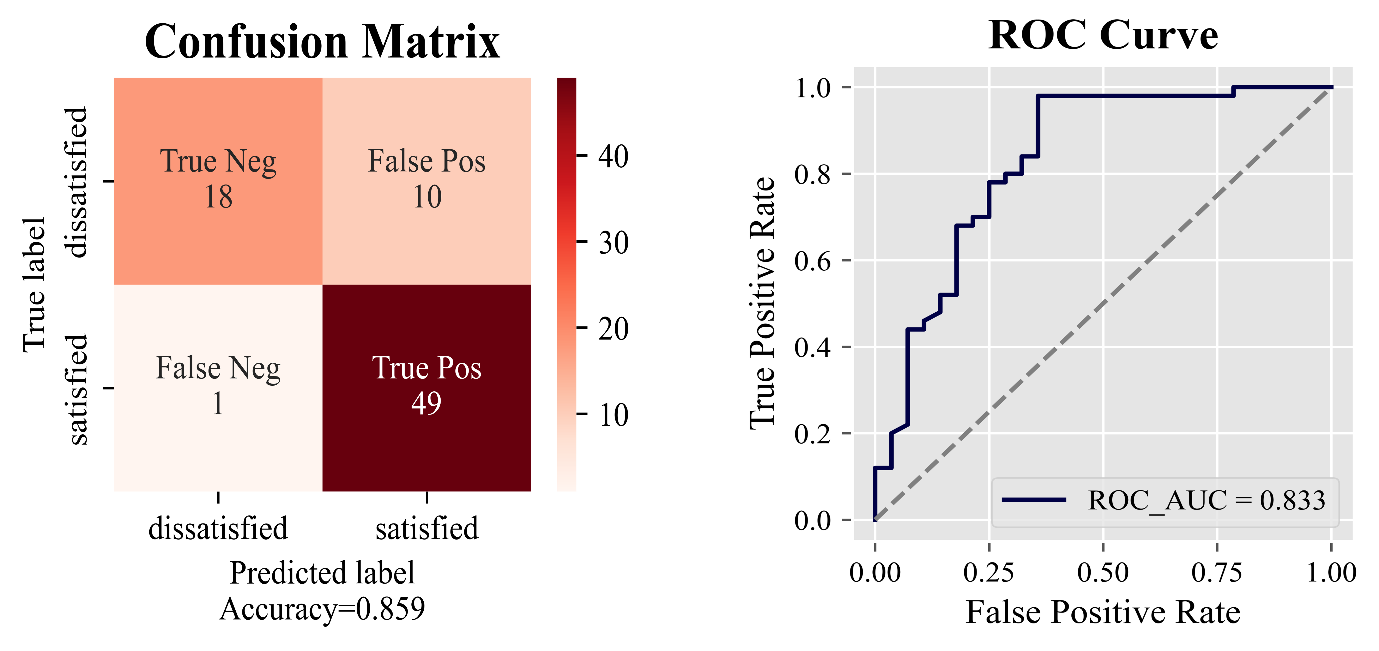


Figure : Confusion matrix and ROC curve for LightGBM classifier

We observe that, in most cases, approaches applied before the use of SMOTE outperform those that incorporate SMOTE. This can be attributed to SMOTE artificially increasing the proportion of positive (minority) samples, allowing the model to learn features related to these underrepresented instances more effectively. However, this synthetic augmentation often introduces redundant or less meaningful information, which may impair the model's ability to generalize effectively. As a result, we see a decline in the Matthews Correlation Coefficient (MCC) scores after applying SMOTE, as demonstrated in the last column of Table 2. Interestingly, both the AUC and accuracy scores also tend to decrease following SMOTE application, suggesting that while the model becomes more attuned to the minority class, it may overfit on the synthetic samples and lose its overall discriminatory power on unseen data. This emphasizes the complexity of using SMOTE in predictive modeling: although it improves class balance, it can negatively impact key performance metrics, including AUC and MCC. Consequently, careful consideration is required to weigh the advantages of minority class representation against the risks of overfitting, and to decide whether SMOTE should be used alongside other techniques to optimize performance. In conclusion, LightGBM classifier (before SMOTE) is the optimal model for our dataset, as it has the highest MCC, accuracy score and the AUC score is also relatively good. Therefore, we have reason to conclude that the SHAP interpretation for LightGBM (before SMOTE) is the most accurate for our dataset.

## Model Interpretability

From the experimental results, we know that LightGBM classifier can achieve the best performance. Next, we will analyze the roles of features played in LightGBM. We used TreeSHAP (Lundberg et al., 2020) to interpret the relationship between features and patient satisfaction because LightGBM classifier, based on the tree model, was the best model in our experiments. We use a global feature importance plot and a SHAP summary plot based on a LightGBM classifier to validate our model and identify influential factors. SHAP, a unified interpretability framework, aids researchers in explaining complex model predictions (Lundberg & Lee, 2017). Its theoretical robustness makes SHAP particularly useful in regulated environments. SHAP values provide quantitative insights into each feature's contribution, verifying its relationship with the target variable (Liu et al., 2023). Figure 3 shows the global importance of each feature that affected the prediction of patient satisfaction based on the SHAP values. The influence of treatment plan was the most significant, followed by age. However, simple quantitative values could not determine whether these features have a positive or negative effect on patient satisfaction. To identify the overall local importance according to the feature value, a SHAP summary plot sorted by global importance is also shown in Figure 3. Comparing SHAP values across features allows us to explore each feature's contribution, thereby estimating its impact on the dependent variable. In Figure 3, SHAP values for each feature are displayed on individual rows, with color gradients indicating whether an observation is lower (blue) or higher (red) for that feature. The features are ranked in descending order of importance: treatment plan, age, appointment ease, waiting time, medicine details, gender, visiting time, education level, patient neglect, and involvement in decision-making. To explore each feature’s impact through their SHAP values, the SHAP summary plot (Figure 3) shows a straightforward SHAP plot, where the SHAP value greater than 0 indicates a positive effect (satisfied) and less than 0 indicates a negative effect (dissatisfied) on the target variable named patient satisfaction. The figure also demonstrates that greater value of waiting time, ignore patient and interestingly decision involve negatively impact patient satisfaction. By contrast, treatment plan, age, appointment ease, medicine details, gender and visiting time positively affect patient satisfaction. Education level shows a complex relationship where most of the time higher values positively affect patient satisfaction and sometimes it also negatively impacts on patient satisfaction.

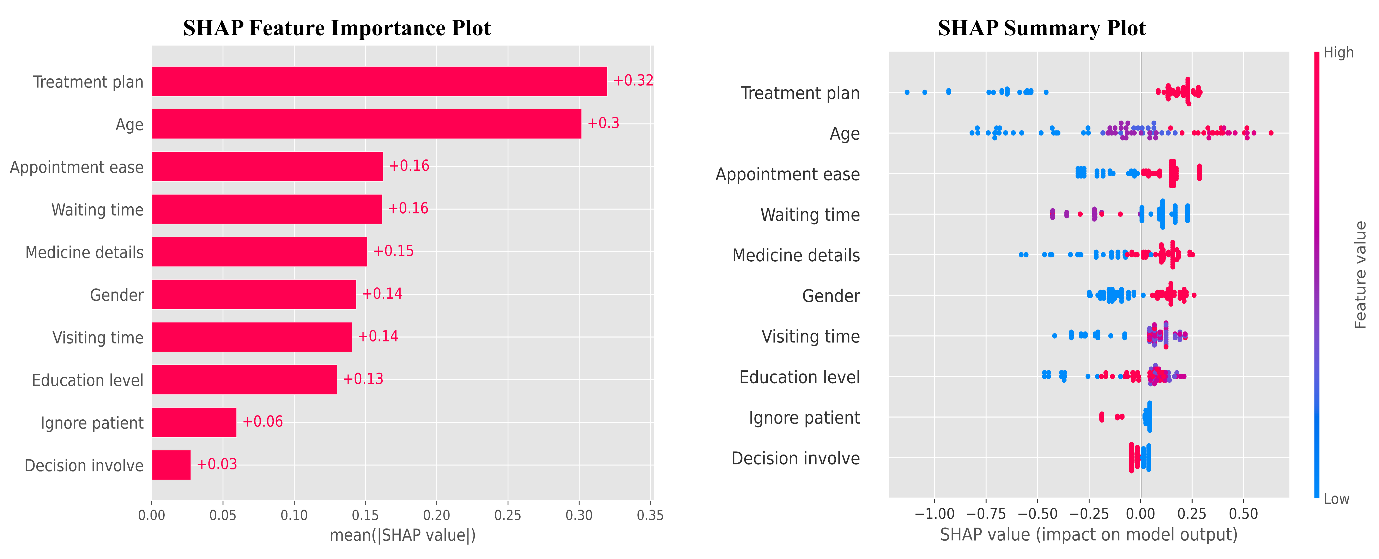


Figure : SHAP global feature importance and summary plot of LightGBM classifier

Figure 4 shows dependence plots for patient characteristics. When the value of attribute “Age” is 3 and 4 (i.e., a patient’s age is 45+), the SHAP value presented by the plot is significantly higher than the SHAP value when the value of age is 0 (i.e., a patient’s age is 16-25), and the SHAP value is around 0 when the value of age is 1 and 2 (i.e., a patient’s age is 26-45). Thus, we can infer that younger patients generally obtain less service satisfaction than older patients. The SHAP values for the "Gender" attribute show a minimal influence on service satisfaction, with values consistently close to zero across all gender categories. This suggests that gender may not be a strong predictor of patient satisfaction within this model, indicating that the satisfaction levels are not significantly biased by gender. For “Education level”, patients with higher education levels (indicated by values 2 and 3) show slightly positive SHAP values, suggesting a modest association with higher satisfaction. In contrast, those with lower education levels (values 0 and 1) generally have lower or neutral SHAP values, implying a lesser or neutral impact on satisfaction. This could mean that patients with higher education may have better communication skills or greater health literacy, leading to higher satisfaction with the services received.

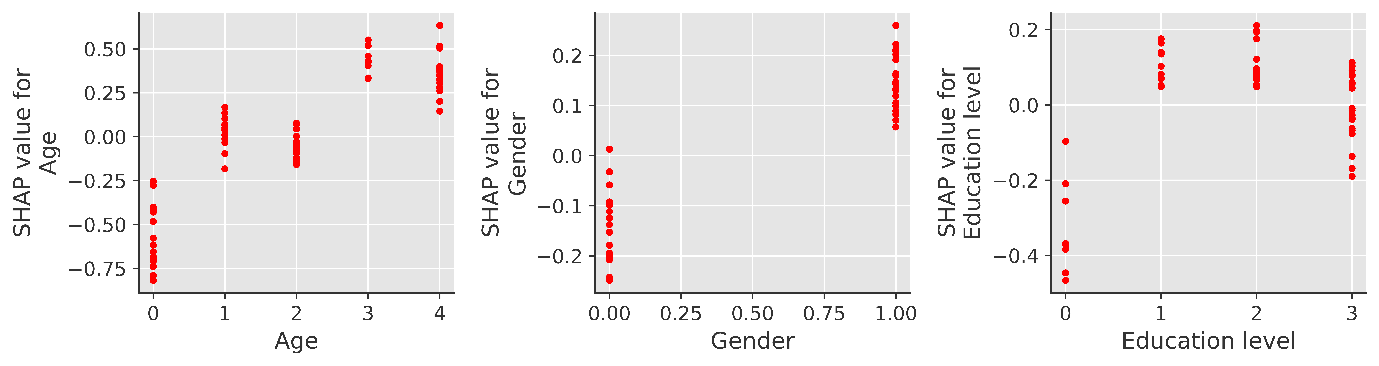


Figure 4: Dependence plot for patient characteristics

Figure 5 presents SHAP dependence plots highlighting the influence of various doctor-patient interaction characteristics on patient satisfaction. The SHAP values indicate a slight positive impact on satisfaction when a treatment plan is provided (value of 1), while no treatment plan (value of 0) appears to have a neutral impact. This suggests that patients may feel more satisfied when they receive a clear treatment plan, possibly due to greater clarity and direction in their care. The SHAP values associated with appointment cases are relatively stable and close to zero across categories, implying a negligible impact on satisfaction. This suggests that, within this model, the procedural aspects of scheduling and managing appointments do not significantly affect patient satisfaction levels. Extended waiting times (values 1 and 2) are associated with negative SHAP values, indicating a reduction in patient satisfaction, while shorter waiting times (value 0) yield SHAP values near zero. This underscores the expectation that shorter wait times contribute positively to the patient experience, likely by reducing feelings of frustration and enhancing perceived efficiency in care delivery. For “Medicine details” Patients who received detailed explanations about their medications (value 1) show slightly positive SHAP values, suggesting a modest improvement in satisfaction. This finding highlights the role of clear communication regarding treatment details in enhancing patients’ trust and adherence, thereby contributing to higher satisfaction levels. SHAP values indicate a slightly negative impact on satisfaction for extended visiting times (value 2), while shorter visits (values 0 and 1) are associated with neutral or mildly positive effects. This pattern suggests that longer visits do not necessarily lead to increased satisfaction, possibly due to perceptions of inefficiency or diminished attention to patient needs during prolonged consultations. Interestingly patients not involved in decision-making (value 0) show slightly positive SHAP values, suggesting that leaving decisions to the healthcare provider may enhance satisfaction by reinforcing trust and confidence in professional expertise.

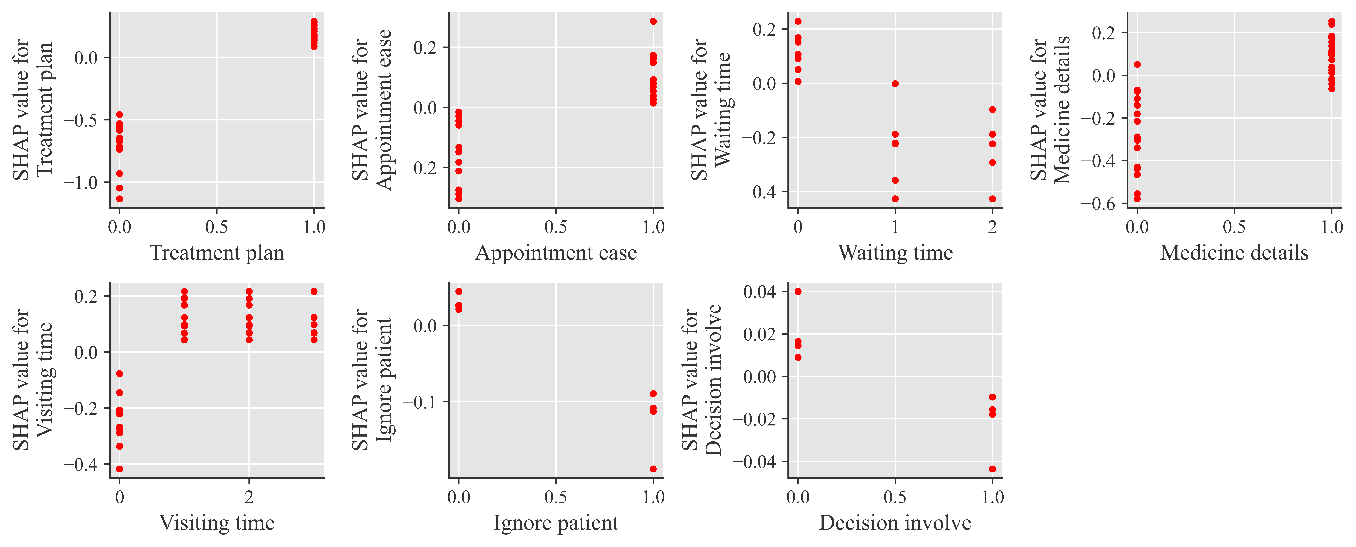


Figure : Dependence plot for interaction characteristics

# Discussions

In this study, we used a cross-section study and assumed the association between 11 factors using a survey questionnaire and machine learning analysis. A total of 312 patients were taken part in the survey and 198 (63.46%) of them were satisfied. The Light Gradient Boosting Machine (LightGBM) algorithm was utilized to pinpoint the primary drivers of patient satisfaction, effectively identifying the critical factors that contribute most significantly to patient satisfaction outcomes. This approach allowed for a nuanced analysis, enabling the identification of key determinants with precision, thereby offering valuable insights into enhancing patient care experiences while Random Forest algorithm was used in (Simsekler et al., 2021). The LightGBM classifier was employed to generate a confusion matrix, demonstrating a notable accuracy of 85.9%. Additionally, the ROC curve provided further validation of the model's discriminative strength, with an Area Under the Curve (AUC) of 0.833. Furthermore, the Matthews Correlation Coefficient (MCC) score surpassed those of all other models tested, highlighting LightGBM superior predictive performance. Based on these results, we conclude that the LightGBM classifier is the optimal model for our dataset, offering robust insights and reliable classification for our analysis which is consistent with the findings of (Xu et al., 2022). Using SHAP analysis with the LightGBM model, we identified the key features influencing patient satisfaction. The features with the highest importance scores included treatment plan, age, and ease of appointment (0.32, 0.30 & 0.16), indicating their strong impact on satisfaction levels. In contrast, decision involvement, ignoring patient concerns, and education level (0.03, 0.06 & 0.13) showed the lowest feature importance scores, suggesting a lesser influence on overall patient satisfaction in our model which aligns with the findings of (Al & Thesis, 2022) where we found surprisingly little evidence for the influence of culturally acceptable, comfortable communication, doctors’ explanations, attitudes and caring.

The implications of this study extend beyond mere predictive accuracy, as the LightGBM model provides a framework for healthcare practitioners to address specific areas affecting satisfaction proactively. By adopting a data-driven approach to patient care, healthcare providers can identify improvement opportunities that resonate directly with patient priorities. This approach not only enhances patient satisfaction but may also lead to improved health outcomes, as satisfied patients are generally more likely to engage positively with their treatment plans. Finally, it is important to note that while our study provides valuable insights, limitations such as the cross-sectional design and reliance on self-reported data should be considered. Future research may benefit from longitudinal studies to explore the causal relationships between patient satisfaction factors and health outcomes over time. Additionally, expanding the sample size and incorporating diverse healthcare settings may yield a broader understanding of patient satisfaction dynamics.

# Conclusion and Recommendations

This study, Unveiling the Determinants of Patient Satisfaction Using Machine Learning Algorithms, offers critical insights into how advanced machine learning techniques can be harnessed to analyze and predict patient satisfaction in healthcare settings. By employing several state-of-the-art classification algorithms and rigorous evaluation metrics, we identified that the LightGBM classifier, especially when applied prior to SMOTE, outperformed other models across key metrics, including the Matthews Correlation Coefficient (MCC), ROC-AUC score, and accuracy. This demonstrates LightGBM’ s robustness in effectively distinguishing between satisfied and dissatisfied patients, even when faced with the inherent challenge of imbalanced data. The superior performance of LightGBM underscores the importance of selecting algorithms that can maintain accuracy and precision while balancing the trade-off between minority and majority classes.

In addition, our SHAP analysis provided interpretability to the model, offering a clear understanding of the underlying factors driving patient satisfaction. The results indicated that treatment plans, age, and appointment ease were the most influential predictors, with these features playing critical roles in shaping patient experiences. Furthermore, factors such as waiting time, involvement in decision-making, and clear communication regarding medication significantly impacted patient satisfaction levels. Interestingly, the analysis revealed that younger patients with lower education levels generally reported lower satisfaction, pointing toward potential areas where healthcare providers could enhance their engagement strategies, particularly by improving communication and tailoring services to better meet the needs of these groups.

However, it is important to acknowledge the limitations of this study. Firstly, the focus was limited to a specific context—intensive care medical personnel in Bangladesh—which may restrict the generalizability of the findings to healthcare professionals in other countries or different medical settings. Moreover, the use of a cross-sectional survey design limits our ability to establish causal relationships between the identified factors and patient satisfaction. Future longitudinal studies would be valuable in examining how these factors and satisfaction levels evolve over time, providing a more dynamic view of patient experiences.

In conclusion, this research illustrates the immense potential of machine learning not only to predict patient satisfaction but also to provide deep, actionable insights into the factors that influence it. The findings suggest that healthcare providers should prioritize reducing waiting times, providing clear and structured treatment plans, and enhancing communication to improve patient satisfaction. As the healthcare industry increasingly adopts data-driven approaches, this study contributes to the growing body of research focused on leveraging machine learning to improve patient care and satisfaction outcomes. Future research could extend this work by incorporating additional social, psychological, and behavioral variables, and by exploring the integration of these models across diverse healthcare environments to enhance their adaptability and predictive power. The continued refinement of these models will help ensure that healthcare providers can better understand and address the evolving needs of patients in a rapidly changing medical landscape.

Authors’ contributions

MMR, CD & MMIA have contributed equally in all the sections of this research paper. The authors read and approved the final manuscript.

Funding

The author did not receive any specific funding for this research work.

**Data Availability**

The datasets are available upon reasonable request from the corresponding author, Md. Mahafuzur Rahman, at [mahafuzur.brur@gmail.com](file:///C:\Users\User\Downloads\mahafuzur.brur@gmail.com). Access may be subject to ethical and institutional restrictions.

Declarations

### Ethics Approval

Ethical approval was granted by the Ethics Committee of the Department of Statistics, Begum Rokeya University, Rangpur.

### Accordance with Guidelines

This study adhered to all relevant guidelines and regulations in conducting research involving human participants.

### Informed Consent

Informed consent was obtained from all participants prior to their involvement in the study

References

Adhikary, G., Shajedur Rahman Shawon, M., Wazed Ali, M., Shamsuzzaman, M., Ahmed, S., Shackelford, K. A., Woldeab, A., Alam, N., Lim, S. S., Levine, A., Gakidou, E., & Jasim Uddin, M. (2018). Factors influencing patients’ satisfaction at different levels of health facilities in Bangladesh: Results from patient exit interviews. *PLoS ONE*, *13*(5). https://doi.org/10.1371/journal.pone.0196643

Ahmed, S. M. (2005). *Exploring health-seeking behaviour of disadvantaged populations in rural Bangladesh*. Institutionen för folkhälsovetenskap / Department of Public Health Sciences. http://openarchive.ki.se/xmlui/handle/10616/39135

Aktar, M. (2021). Determinates Service Quality and Its Effect on Patients’ Satisfaction of Private Medical College Hospitals, Rangpur, Bangladesh. *International Journal of Management and Accounting*. https://doi.org/10.34104/ijma.021.0910105

Al, F., & Thesis, H. M. (2022). *Evaluation of Patient Satisfaction in Telemedicine Using Machine Learning Algorithms*.

Alhashem, A. M., Alquraini, H., & Chowdhury, R. I. (2011). Factors influencing patient satisfaction in primary healthcare clinics in Kuwait. *International Journal of Health Care Quality Assurance*, *24*(3), 249–262. https://doi.org/10.1108/09526861111116688

Batbaatar, E., Dorjdagva, J., Luvsannyam, A., Savino, M. M., & Amenta, P. (2017). Determinants of patient satisfaction: A systematic review. In *Perspectives in Public Health* (Vol. 137, Issue 2, pp. 89–101). SAGE Publications Ltd. https://doi.org/10.1177/1757913916634136

Begum, F., Said, J., Hossain, S. Z., Hasan, M. J., & Binti, N. (2021). *Healthcare Cost and Patient Satisfaction: A Comparative Analysis Between Public and Private Hospitals in Bangladesh*. In Review. https://www.researchsquare.com/article/rs-325300/v1

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13-17-August-2016*, 785–794. https://doi.org/10.1145/2939672.2939785

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1), 6. https://doi.org/10.1186/s12864-019-6413-7

Chicco, D., & Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining*, *16*(1), 4. https://doi.org/10.1186/s13040-023-00322-4

Clever, S. L., Jin, L., Levinson, W., & Meltzer, D. O. (2008). Does doctor-patient communication affect patient satisfaction with hospital care? Results of an analysis with a novel instrumental variable. *Health Services Research*, *43*(5 P1), 1505–1519. https://doi.org/10.1111/j.1475-6773.2008.00849.x

Ferdousi, M. J. (2015). Patient satisfaction with community clinic care: facility and household based survey in a sub-district in Bangladesh. *Mediscope*, *1*(1), 23–28. https://doi.org/10.3329/mediscope.v1i1.21633

Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, *55*(1), 119–139. https://doi.org/10.1006/JCSS.1997.1504

Friedman, J. (2000). Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, *29*(5), 1189–1232.

Irfan, S. M., & Farooq, M. M. (2012). Patient Satisfaction and Service Quality of Public Hospitals in Pakistan: An Empirical Assessment. *Middle-East Journal of Scientific Research*, *12*(6), 870–877. https://doi.org/10.5829/idosi.mejsr.2012.12.6.2743

Jalil, A., Zakar, R., Zakar, M. Z., & Fischer, F. (2017). Patient satisfaction with doctor-patient interactions: a mixed methods study among diabetes mellitus patients in Pakistan. *BMC Health Services Research*, *17*(1). https://doi.org/10.1186/s12913-017-2094-6

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: a highly efficient gradient boosting decision tree. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 3149–3157.

Kuteyi, E. A., Bello, I., Olaleye, T., Ayeni, I., & Amedi, M. (2010). Determinants of patient satisfaction with physician interaction: a cross-sectional survey at the Obafemi Awolowo University Health Centre, Ile-Ife, Nigeria. *South African Family Practice*, *52*(6). https://www.ajol.info/index.php/safp/article/view/63273

Li, C., Liao, C., Meng, X., Chen, H., Chen, W., Wei, B., & Zhu, P. (2021). Effective analysis of inpatient satisfaction: The random forest algorithm. *Patient Preference and Adherence*, *15*, 691–703. https://doi.org/10.2147/PPA.S294402

Liu, C., Li, Y., Fang, M., & Liu, F. (2023). Using machine learning to explore the determinants of service satisfaction with online healthcare platforms during the COVID-19 pandemic. *Service Business*, *17*(2), 449–476. https://doi.org/10.1007/s11628-023-00535-x

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From local explanations to global understanding with explainable ai for trees. *Nature Machine Intelligence*, *2*(1), 2522–5839. https://doi.org/10.1038/s42256-019-0138-9

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Neural Information Processing Systems*. https://api.semanticscholar.org/CorpusID:21889700

Norhayati, M. N., Masseni, A. A., & Azlina, I. (2017). Patient satisfaction with doctor-patient interaction and its association with modifiable cardiovascular risk factors among moderately-high risk patients in primary healthcare. *PeerJ*, *2017*(2), 1–12. https://doi.org/10.7717/peerj.2983

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, 6639–6649.

Quinlan, R. J. (1986). Induction of Decision Trees. *Machine Learning*, *1*(1), 81–106. https://doi.org/10.1007/bf00116251

Shin, J., Lee, D. A., Kim, J., Lim, C., & Choi, B. K. (2024). Dissatisfaction-considered waiting time prediction for outpatients with interpretable machine learning. *Health Care Management Science*. https://doi.org/10.1007/s10729-024-09676-5

Shrestha, Y. R., He, V. F., Puranam, P., & von Krogh, G. (2021). Algorithm supported induction for building theory: How can we use prediction models to theorize? *Organization Science*, *32*(3), 856–880. https://doi.org/10.1287/orsc.2020.1382

Simsekler, M. C. E., Alhashmi, N. H., Azar, E., King, N., Luqman, R. A. M. A., & Al Mulla, A. (2021). Exploring drivers of patient satisfaction using a random forest algorithm. *BMC Medical Informatics and Decision Making*, *21*(1). https://doi.org/10.1186/s12911-021-01519-5

Wu, J., Zhang, C., He, F., Wang, Y., Zeng, L., Liu, W., Zhao, D., Mao, J., & Gao, F. (2023). Factors Affecting Intention to Leave Among ICU Healthcare Professionals in China: Insights from a Cross-Sectional Survey and XGBoost Analysis. *Risk Management and Healthcare Policy*, *16*, 2543–2553. https://doi.org/10.2147/RMHP.S432847

Xu, Y., Wu, G., & Chen, Y. (2022). Predicting patients’ satisfaction with doctors in online medical communities: An approach based on XGBoost Algorithm. *Journal of Organizational and End User Computing*, *34*(4). https://doi.org/10.4018/JOEUC.287571

Zhang, H., Wang, W., Haggerty, J., & Schuster, T. (2020). Predictors of patient satisfaction and outpatient health services in China: Evidence from the WHO SAGE survey. In *Family Practice* (Vol. 37, Issue 4, pp. 465–472). Oxford University Press. https://doi.org/10.1093/FAMPRA/CMAA011