Predicting Customer Satisfaction Scores from Customer Support Data Using Random Forest and Gradient Boosting

Matthew Owen Gunawan

Department of Computer Science. Bina Nusantara University, Jakarta, Indonesia

Email: matthew.gunawan002@binus.ac.id

This project was completed in January 2025 while taking the Machine Learning course as part of the Data Science program at Binus University. The author, Matthew Owenn Gunawan, wishes to express gratitude to Mr. Jaffarus Sodiq for his assistance and guidance in completing this project.

Abstract: This study focuses on predicting customer satisfaction using a customer support dataset. Data is processed through several common stages such as handling missing values, removing outliers, and creating feature engineering to capture patterns used for later analysis. Exploratory Data Analysis is also conducted to identify key factors that influence customer satisfaction. Two classifiers are used, namely Random Forest and Gradient Boosting, which are implemented and optimized using hyperparameter tuning. Model performance results are evaluated using accuracy, precision, recall, and F1-score, with special attention to addressing class imbalance. The results show that ensemble-based methods are effective in modeling customer satisfaction, with Gradient Boosting achieving the best overall performance. The results indicate that ensemble-based methods are effective in modeling customer satisfaction, with Gradient Boosting achieving the best overall performance.

Index Terms: Customer Satisfaction Prediction, Customer Support Data, Random Forest, Gradient Boosting, Machine Learning, Classification

1. Introduction

Customer Satisfaction is defined as the pleasure or disappointment customers feel after comparing what they believed would happen with what actually happened. Customer Satisfaction plays a critical role in driving business success especially in highly competitive environments where Customer Satisfaction will directly impact Consumer Loyalty, Consumer Retention and long-term profitability. Customer Satisfaction not only provides an indication of the Quality of Customer Support Services but Customer

Satisfaction will also affect how Customers View a Brand and Customer's Future Behaviour. Customer Satisfaction also provides organisations with a way to improve the quality of Customer Support Services [2][3].

Customer Satisfaction is particularly important to Customer Support because it represents the overall Quality and Effectiveness of how Customers Interact with the Organisation. A high level of Customer Support, which includes Timely Response, Accurate Problem Resolution, and Good Communication, can increase the likelihood that a customer will have a positive experience and therefore be satisfied. Prior research suggests that service quality dimensions such as Responding Quickly, being Reliable, and showing Empathy towards the Customer significantly affect Customer Satisfaction and in turn directly influence Customer Loyalty, Consumer Retention and long-term Profitability [4][5]. Therefore, utilising data analysis and predictive modelling to measure and forecast Customer Satisfaction in the context of Customer Support services will be a critical component for organisations to improve the overall Quality of the Customer Support Experience [6].

Despite its importance, accurately measuring and predicting customer satisfaction in customer support environments remains a crucial and challenging task. Customer satisfaction is influenced by numerous factors, such as response time, handling duration, communication channels, and agent characteristics, which interact in complex and non-linear ways [6]. Previous analysis has shown that service quality dimensions significantly influence customer satisfaction. Furthermore, customer satisfaction data often suffers from class imbalance, where high satisfaction scores dominate the dataset, making it difficult for models to accurately predict lower satisfaction levels, a challenge known as imbalanced classification problems [7]. This challenge highlights the need for robust machine learning approaches that can effectively capture patterns in customer support data and address imbalanced class distributions using two ensemble modeling methods, Random Forest and Gradient Boosting [8][9].

The main objective of this research is to develop and evaluate a machine learning approach to build a customer satisfaction prediction model using a customer support dataset. Specifically, this study aims to:

- Identify the key factors influencing customer satisfaction based on customer support dataset.
- Build a customer satisfaction prediction model using two ensemble learning modeling methods, Random Forest and Gradient Boosting method.
- Compare the performance of both models using relevant evaluation metrics like accuracy, precision, f1-score, and support.
- Analyze the impact of class imbalance on model prediction result.

2. Methodology

2.1 Description of techniques implemented

2.1.1 Ensemble Learning

Ensemble learning is a technique in machine learning that combines multiple base models to produce better predictive performance than a single model. The main idea of ensemble learning is to combine predictions from multiple models, where the overall model can reduce variance, minimize bias, and improve generalization. This is especially true when dealing with complex and non-linear data patterns. Ensemble methods are very effective in real-world applications where data characteristics are heterogeneous and noisy, such as customer support and customer satisfaction data, which involve many interacting factors and uncertain patterns [2]. By leveraging the strengths of multiple models, ensemble learning can achieve more robust and stable predictions compared to individual model classifiers.

2.1.2 Random Forest

Random Forest is an ensemble learning method that builds multiple decision trees and combines prediction results to produce a more stable and accurate final decision. This method uses bootstrap aggregation techniques such as bagging and random feature selection to reduce correlation between trees and reduce model variance. Random Forest is effective in handling high-dimensional data, non-linear relationships, and data containing noise, making it more resistant to overfitting than a single model [11]. In this study, Random Forest was used to predict customer satisfaction due to its ability to handle imbalanced data and identify important features that influence customer satisfaction [8] [11].

2.1.3 Gradient Boosting

Gradient Boosting is an ensemble learning algorithm that combines several simple models such as weak learners gradually to form a more accurate model by correcting previous prediction errors through an iterative approach. At each iteration, a new model is built to learn the residual error from the previous model, so that the final model can capture complex patterns in the data and improve prediction accuracy [12]. In addition, Gradient Boosting excels in handling regression and classification problems, especially when the dataset has non-linear relationships and complex feature interaction patterns. However, this algorithm requires hyperparameter tuning to avoid overfitting and achieve the best performance in prediction [13].

2.2 Dataset Detail

This dataset consists of detailed records of customer support interactions collected from an e-commerce service platform. It includes a combination of categorical, numerical, temporal, and textual features that describe customer issues, service processes, and satisfaction outcomes. The variables provide comprehensive information ranging from communication channels and issue categories to agent-related attributes and service response timelines. Such diversity of features enables an in-depth analysis of factors influencing customer satisfaction and supports the development of predictive models, with the Customer Satisfaction (CSAT) score serving as the target variable [1].

2.2.1 Unique ID

- Represents a unique identifier for each customer support interaction.
- Used to uniquely distinguish each record in the dataset.
- Treated as a categorical identifier and not used directly as a predictive feature.

2.2.2 Channel_name

- Indicates the communication channel through which the customer contacted customer support (e.g., Inbound, Outcall).
- This feature helps analyze differences in customer satisfaction across communication channels.
- Treated as a categorical variable.

2.2.3 category

- Describes the high-level classification of the customer issue (e.g., Product Queries, Returns, Order Related).
- Provides an overview of the nature of customer concerns.
- Treated as a categorical feature.

2.2.4 Sub Category

- Represents a more detailed classification of the customer issue within each main category.
- Captures specific problem types encountered by customers.
- Treated as a categorical variable.

2.2.5 Customer Remarks

- Contains textual feedback or comments provided by customers regarding their issues.

- This feature may include unstructured text data reflecting customer sentiment.

2.2.6 Order ID

- Represents the unique identifier of the customer's order related to the reported issue.
- Enables linkage between customer support interactions and transaction records.
- Treated as a categorical identifier.

2.2.7 order_date_time

- Records the date and time when the related order was placed.
- This temporal feature can provide context regarding the timing of issues relative to order placement.
- Stored as a datetime attribute, with many missing values.

2.2.8 Issue_reported at

- Indicates the timestamp when the customer issue was initially reported.
- Serves as the starting point of the customer support process.
- Used in calculating response time metrics.

2.2.9 issue_responded

- Records the timestamp when the customer support team first responded to the reported issue.
- Reflects the responsiveness of the support team.
- Combined with Issue_reported at to measure service response delay.

2.2.10 Survey_response_Date

- Represents the date when the customer completed the satisfaction survey.
- Captures the timing of customer feedback after issue handling.
- Used to align satisfaction scores with service interactions.

2.2.11 Customer_City

- Indicates the city from which the customer contacted customer support.
- Provides geographical context for customer interactions.
- Treated as a categorical feature with many missing values.

2.2.12 Product_category

- Specifies the category of the product associated with the customer issue.
- Helps identify product-level patterns in customer satisfaction.
- Treated as a categorical variable.

2.2.13 Item_price

- Represents the price of the product associated with the reported issue.
- This numerical feature may influence customer expectations and satisfaction.
- Contains missing values and is treated as a continuous variable.

2.2.14 connected_handling_time

- Indicates the duration (in time units) that the agent spent handling the customer issue.
- Reflects service effort and interaction length.
- Due to extensive missing values, its use is carefully considered during preprocessing.

2.2.15 Agent_name

- Identifies the customer support agent who handled the issue.

- Used for internal analysis of agent performance.
- Treated as a categorical feature and not directly used for prediction.

2.2.16 Supervisor

- Represents the supervisor overseeing the assigned customer support agent.
- Provides hierarchical information within the support organization.
- Treated as a categorical attribute.

2.2.17 Manager

- Indicates the manager responsible for the customer support team.
- Adds organizational context to each interaction.
- Treated as a categorical feature.

2.2.18 Tenure Bucket

- Categorizes agents based on their length of service (e.g., On Job Training, >90 days).
- Reflects agent experience level.
- Used to analyze the impact of experience on customer satisfaction.

2.2.19 Agent Shift

- Indicates the work shift during which the issue was handled (e.g., Morning, Evening).
- Useful for analyzing service performance across different time periods.
- Treated as a categorical variable.

2.2.20 CSAT Score

- Represents the Customer Satisfaction (CSAT) score provided by the customer.
- This score serves as the target variable in the predictive modeling process.
- Stored as an integer numerical variable.

2.3 preprocessing steps

The preprocessing part of this study was conducted to ensure the data was ready for use by machine learning algorithms. The preprocessing steps included handling missing values, removing duplicate data, transforming the data in time, handling outliers, encoding categorical data, normalizing numerical features, and handling class calculations.

2.3.1 Handling missing value

Missing values are handled to maintain data quality and prevent bias in the model training process. For categorical features, blank values are filled with a special label, such as "unknown," to preserve the information contained in the data. Meanwhile, for numeric features, the median value is used to minimize the influence of outliers on the data distribution. This approach is commonly used in preprocessing due to its stability and effectiveness in dealing with data containing missing values [14].

2.3.2 Outlier Detection and removal

Outliers in numerical features are identified using an interquartile range (IQR)-based statistical approach. Values outside the lower and upper limits are removed to prevent distortion of the data distribution [14]. The presence of outliers has a negative impact on model performance, especially in tree-based and ensemble algorithms. Therefore, outlier removal is a crucial step in the preprocessing stage to achieve good results][15].

2.3.3 Encoding for categorial features

This dataset contains categorical features that cannot be directly processed by machine learning algorithms, so they need to be converted into numerical form [15]. In this study, the Label Encoding technique is needed to convert categories such as channel, agent shift, and tenure bucket into numerical representations. This technique was chosen because it is suitable for categorical features with a limited number of classes. The encoding process allows the model to effectively learn patterns from categorical data [16].

2.3.4 Feature Scaling and Normalization

Feature scaling is performed to equalize the scale between numeric features so that no single variable dominates the model learning process. In this study, standardization is used to change feature values to have a mean of zero and a standard deviation of one. This step is crucial because the dataset contains features with varying value ranges. By performing feature scaling, the model training process becomes more stable and prediction results can be consistently improved [17].

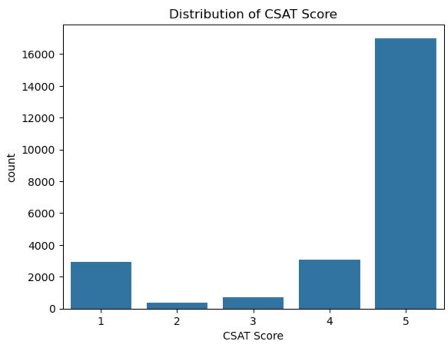
2.3.5 Handling Class Imbalance

In the customer satisfaction dataset, the data is unevenly distributed across classes. The majority of data fall within high satisfaction classes, while the number of low satisfaction classes is significantly smaller. This can lead to the model tending to overestimate the majority class and underestimating the minority class. Therefore, this study applies the SMOTE method to increase the number of data points in the minority class to improve the data distribution and allow the model to better learn all classes [18].

2.4 Exploratory Data Analysis

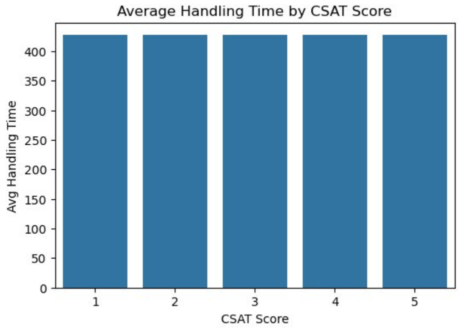
Exploratory Data Analysis (EDA) was conducted to examine the distribution of variables, identify patterns, and highlight features most relevant for predicting customer satisfaction.

2.4.1 Distribution of CSAT Score



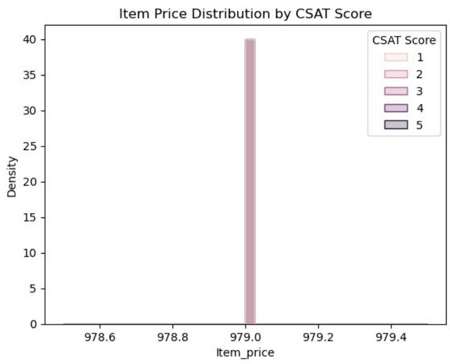
Among all responding customers, CSAT ratings were predominately rated as 5. This response indicates overall high levels of customer satisfaction. Moderate numbers of ratings of 4 and 1 suggest that a moderate percent of surveyed customers are dissatisfied or only somewhat satisfied with the brand. Ratings of 2 and 3 were less frequent with the majority of customer sentiment clustering toward either end of the CSAT scoring system. While service quality, as a whole, received high ratings by customers and was perceived positively, it will be important to continue to support those with lower CSAT satisfaction scores.

2.4.2 Relationship between CSAT Score and Handling Time



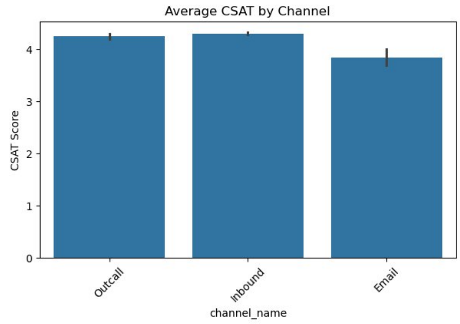
The average time required to handle a case appears relatively stable regardless of customer satisfaction (CSAT) score and shows no statistically significant difference for low scores (1-2) compared to high scores (4-5). Time appears to have little impact on overall customer satisfaction. In other words, even if handling times are fairly comparable across multiple cases, the customer's experience with different agents. For example, agent attitudes, or responses to cases may have a much more meaningful impact on CSAT scores than how long it took to handle their case. This provides an opportunity to focus on enhancing interaction quality rather than simply improving time efficiency.

2.4.3 Item Price Distribution by CSAT Score



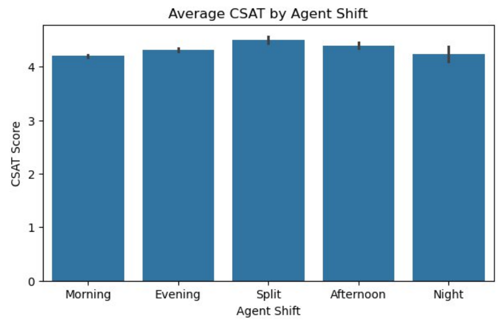
The price of items shared by customers who rated their experience as having a Customer Satisfaction (CSAT) level between extremely low to extremely high is very closely grouped together with a mean price of roughly 979.0. Therefore, the price of the item has nearly no effect on CSAT. Even though the prices were very close, we see that there is a very good representation of both the low and high scores falling within the price range of the item. It indicates that something besides the price of an item will have a greater influence on a customer's CSAT score.

2.4.4 Average CSAT Score by Communication Channel



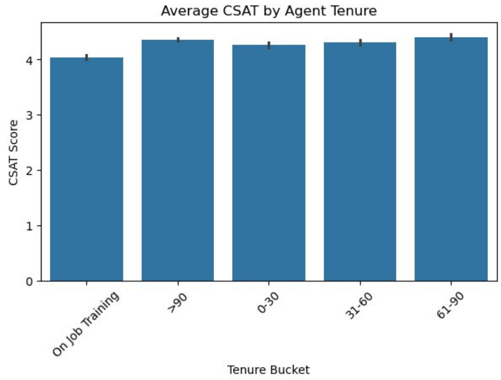
The highest average CSAT scores were obtained from the outcall and inbound channels, both slightly above 4. The email channel had a lower score, closer to 4, indicating a relatively lower level of satisfaction. This difference indicates that in-person interactions of phone calls tend to result in a more satisfying customer experience than email communications. Error bars indicate some variation, but the general pattern remains consistent across channels.

2.4.5 Average CSAT Score by Agent Shift



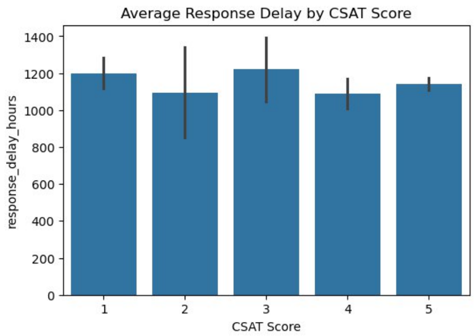
On CSTA score, the average for all shifts was slightly higher than 4, suggesting the level of satisfaction was consistent across all shifts. Assistant agent hours do not seem to have a significant impact on the customer experience although some slight differences exist between morning, afternoon, evening, and split shifts. Other factors should be targeted for improvement, including interaction quality and the ability to resolve problems.

2.4.6 Average CSAT Score by Agent Tenure



Customer Satisfaction Scores (CSAT) generally increased as the amount of time as an agent increased. The CSAT scores were highest for agents who had been in their positions for more than 90 days but lower for agents still in training. This indicates that the agent's experience is a positive influence on Customer Satisfaction Scores. Therefore, to enhance service quality, it may be beneficial to continue to train and retain agents who have had greater amounts of experience working directly with customers.

2.4.7 Relationship between Response Delay and CSAT Score



The highest average response time occurred with a CSAT score of 3, while a score of 4 had the fastest response time. This pattern suggests that response delays don't always correlate with customer dissatisfaction. Lower scores like 1 and 2 actually had faster response times than a score of 3, indicating that other factors, such as the quality of the response or the agent's attitude, may be more influential. This insight encourages a deeper evaluation of interaction quality, not just response speed.

Overall, the exploratory analysis indicates that customer satisfaction scores are highly imbalanced, with most customers reporting very high satisfaction levels. Several operational factors such as handling time, item price, and agent shift show limited influence on CSAT scores. In contrast, interaction-related factors, including communication channel and agent experience, demonstrate a clearer association with customer satisfaction. These findings suggest that customer satisfaction in this dataset is driven more by service quality and agent-related factors rather than operational efficiency alone. Based on these observations, the modelling stage focuses on handling class imbalance and capturing patterns related to customer–agent interactions.

3. Result

In this section, the performance of two models: Random Forest and Gradient Boosting Classifier are evaluated:

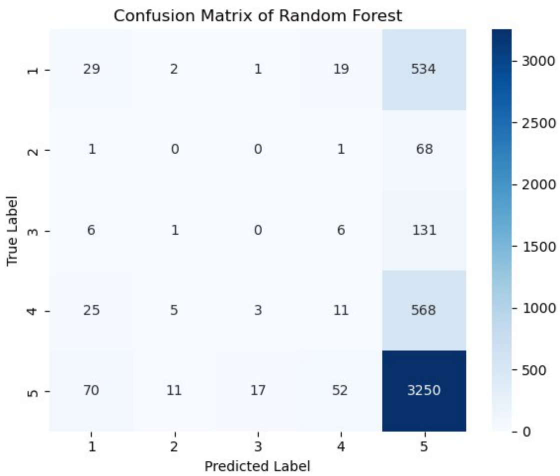
3.1 Random Forest

```python
smote = SMOTE(
    random_state=42,
    sampling_strategy={
        1: 3000,
        2: 800,
        3: 1000,
        4: 3000
    }
)

X_train_res, y_train_res = smote.fit_resample(X_train, y_train)

rf = RandomForestClassifier(
    random_state=42,
    n_estimators=400,
    max_depth=30,
    min_samples_split=5,
    min_samples_leaf=2,
    max_features='sqrt'
)

rf.fit(X_train_res, y_train_res)

y_pred = rf.predict(X_test)

print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nClassification Report:\n", classification_report(y_test, y_pred))

pred_counts = pd.Series(y_pred).value_counts().sort_index()
print("\nPredicted class counts:\n", pred_counts)
```

The CSAT dataset exhibited class imbalance, which was addressed using SMOTE to increase the number of samples in the minority classes according to a custom sampling strategy. A Random Forest model with 400 decision trees and a maximum depth of 30 was then trained on the balanced data. Model performance was evaluated using accuracy, precision, recall, F1-score, and the number of predictions per class. The use of SMOTE helped improve prediction balance across all classes and reduced the model’s tendency to favor the majority class.

```
Accuracy: 0.6838495115360632

Classification Report:
              precision    recall  f1-score   support

           1       0.22      0.05      0.08       585
           2       0.00      0.00      0.00        70
           3       0.00      0.00      0.00       144
           4       0.12      0.02      0.03       612
           5       0.71      0.96      0.82      3400

    accuracy                           0.68      4811
   macro avg       0.21      0.20      0.19      4811
weighted avg       0.55      0.68      0.59      4811


Predicted class counts:
 1     131
2      19
3      21
4      89
5    4551
Name: count, dtype: int64
```

The Random Forest model achieved an accuracy of 0.6838495115360632 on the test data. It performs very well in predicting the majority class (CSAT 5), with recall 0.96 and F1-score 0.82, making it reliable for identifying highly satisfied customers. For the minority classes (CSAT 1–4), performance is still low, although some predictions for these classes do appear. Overall, even with SMOTE applied, the model's predictions are still dominated by the majority class, which reflects the severe imbalance in the dataset rather than a problem with the Random Forest model itself.



The confusion matrix visualization illustrates that most samples across different true classes are predicted as CSAT Score 5. While correct predictions are concentrated along the diagonal for class 5, misclassifications are common for lower CSAT scores, which are frequently confused with the highest score. This highlights the challenge of distinguishing dissatisfied customers in an imbalanced multi-class classification problem.

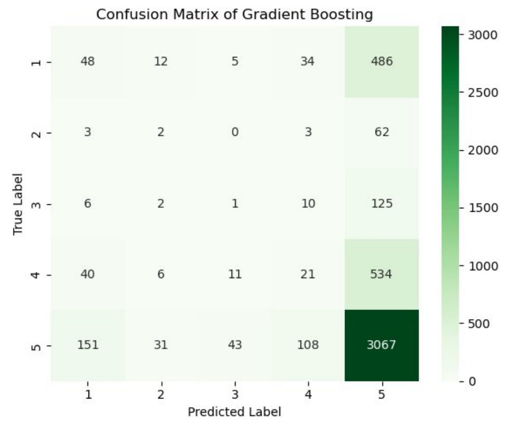
3.2 Gradient Boosting Classifier

```python
param_dist_gb = {
    'n_estimators': [100, 200, 300],
    'learning_rate': [0.01, 0.05, 0.1],
    'max_depth': [3, 5, 7],
    'subsample': [0.8, 1.0]
}

gb = GradientBoostingClassifier(random_state=42)

gb_random = RandomizedSearchCV(
    gb,
    param_distributions=param_dist_gb,
    n_iter=15,
    cv=3,
    scoring='f1_macro',
    random_state=42,
    n_jobs=-1,
    verbose=2
)

gb_random.fit(X_train, y_train)

best_gb = gb_random.best_estimator_
y_pred_gb = best_gb.predict(X_test)

print("Best Params GB:", gb_random.best_params_)
print("GB Accuracy:", accuracy_score(y_test, y_pred_gb))
print(classification_report(y_test, y_pred_gb))
```

Since the CSAT dataset has a class imbalance, SMOTE was used to balance the number of samples in the minority class before training. The Gradient Boosting Classifier was then optimized using RandomizedSearchCV with 3-fold cross-validation, testing several important parameters such as the number of estimators, learning rate, maximum tree depth, and subsampling ratio, with the F1-macro evaluation metric. The best model obtained was used to predict the test data, and its performance was evaluated using accuracy, classification report, and the number of predictions per class. The results showed strong performance on the majority class (CSAT 5) while being able to capture some predictions for the minority class.

```
Fitting 3 folds for each of 10 candidates, totalling 30 fits
Best Params GB: {'subsample': 0.8, 'n_estimators': 400, 'max_depth': 10, 'learning_rate': 0.05}
GB Accuracy: 0.6524631053834962

Classification Report:
              precision    recall  f1-score   support

           1       0.19      0.08      0.12       585
           2       0.04      0.03      0.03        70
           3       0.02      0.01      0.01       144
           4       0.12      0.03      0.05       612
           5       0.72      0.90      0.80      3400

    accuracy                           0.65      4811
   macro avg       0.22      0.21      0.20      4811
weighted avg       0.55      0.65      0.59      4811


Predicted class counts (GB):
 1     248
2      53
3      60
4     176
5    4274
Name: count, dtype: int64
```

Gradient Boosting achieved an accuracy of approximately 0.6524631053834962 on the test data. The model performed quite well in predicting highly satisfied customers (CSAT Score 5), with a recall of 0.90 and an F1-score of 0.80. For minority classes (CSAT 1–4), the predictions were lower, but the model still generated some predictions, for example, 248 for CSAT 1 and 176 for CSAT 4. These results indicate that Gradient Boosting is able to handle some data imbalance, especially in the dominant class, while still providing information for less frequent classes.



The confusion matrix visualization shows that most of the data from various actual classes are predicted as CSAT Score 5. Correct predictions are most common for this class, while classes with low satisfaction scores are often misclassified as high scores. Although Gradient Boosting provides improved accuracy compared to Random Forest, the problem of class imbalance remains a major challenge in this modeling.

4. Discussion

4.1 Analysis of the result

Evaluation results show that both models perform quite well in identifying highly satisfied customers (CSAT score 5). However, they still struggle to predict customers with low scores due to data imbalance.

4.1.1 Random Forest

Random Forest is able to recognize the majority class (CSAT Score 5) well, with high recall and F1-score. For the minority classes (CSAT 1–4), performance is still low, although some predictions for these classes appear. The confusion matrix shows that most errors occur because low-score customers are often predicted as class 5. Overall, Random Forest is reliable for identifying highly satisfied customers, but less effective at distinguishing dissatisfied customers.

4.1.2 Gradient Boosting Classifier

Gradient Boosting showed slightly lower accuracy than Random Forest, at 0.652, but was able to predict the minority class better. Predictions for

CSAT 5 remained strong, with a recall of 0.90 and an F1-score of 0.80, while for CSATs 1–4, the model remained low but produced noticeable predictions, for example: 248 for CSAT 1, 53 for CSAT 2, 60 for CSAT 3, and 176 for CSAT 4. This indicates that Gradient Boosting was able to capture some minority class patterns that did not appear in Random Forest. The confusion matrix shows that majority errors still occur when low scores are predicted as 5, but the distribution of predictions is more balanced than in Random Forest.

4.2 Insights gained

Evaluation results show that both models can reliably identify highly satisfied customers (CSAT 5). For dissatisfied customers (CSAT 1–4), prediction remains limited, although Gradient Boosting is able to capture some of the minority classes—producing measurable predictions for CSAT 1–4 while Random Forest tends to focus more heavily on the majority class. This suggests that Gradient Boosting provides a slightly more balanced distribution of predictions across classes, even though challenges from data imbalance remain.

- Both models consistently detected highly satisfied customers (CSAT 5), making them useful for analyzing the majority of customers.
- Prediction for low-scoring customers (CSAT 1–4) was limited, highlighting the challenges in detecting dissatisfaction.
- Using SMOTE helped improve the prediction distribution for minority classes, but extreme imbalance still impacted performance.
- Gradient Boosting showed a slightly more balanced prediction distribution than Random Forest, although the improvement for minority classes was still limited.
- Further improvements in minority class prediction may require a more aggressive oversampling strategy, a minority focus algorithm, or additional data for underrepresented classes.

4.3 Challenges faced

The main challenge is the strong class imbalance in the dataset, with most observations concentrated on CSAT 5. This affects the models' learning process, causing them to focus primarily on the majority class. Even with SMOTE and hyperparameter tuning, both Random Forest and Gradient Boosting still have difficulty predicting the minority classes accurately.

Another challenge comes from the limited features available to differentiate low scores. While the models can easily learn patterns of highly satisfied customers, the characteristics of less satisfied customers are less distinct. Efforts to improve recall for these minority classes sometimes reduce overall accuracy, showing that resampling and tuning alone cannot fully solve the problem. This indicates that the limitations are not only due to the choice of model but also related to the structure and information content of the dataset.

5. Conclusion

5.1 Random Forest

Random Forest successfully identified highly satisfied customers (CSAT 5), as indicated by high recall and F1-score. However, for the minority class (CSAT 1–4), the model's accuracy remained limited, although some predictions were still made. This indicates that Random Forest is reliable for detecting high satisfaction but less effective in distinguishing less satisfied customers. Using SMOTE helped improve the prediction distribution, although the majority class still dominated.

5.2 Gradient Boosting Classifier

Gradient Boosting performed well, with an accuracy of around 0.652. Predictions for CSAT 5 remained strong, with a recall of 0.90 and an F1-score of 0.80, while for CSATs 1–4, the model produced predictions that looked like 248 for CSAT 1 and 176 for CSAT 4. This demonstrates Gradient Boosting's ability to capture minority class patterns better than Random Forest. While prediction difficulties for minority classes persist, the distribution of predictions is more even, making the model slightly more balanced overall.

6. References

[1] Kaggle, "Ecommerce Customer Service Satisfaction Dataset," 2023. [Online]. Available: https://www.kaggle.com/datasets/ddosad/ecommerce-customer-service-satisfaction

[2] D. Januaji, "Customer satisfaction adalah: ini pengertian dan cara meningkatkannya," Accurate.id, Jun. 04, 2024. [Online]. Available: https://accurate.id/bisnis-ukm/customer-satisfaction-adalah/

[3] Sprint Asia Technology, “5 Strategies to Increase Customer Satisfaction Effectively.” [Online]. Available: https://sprintasia.co.id/5-strategies-to-increase-customer-satisfaction-effectively/
[4] KPSG, “Layanan Customer Service untuk Customer Satisfaction.” [Online]. Available: https://kpsg.com/layanan-customer-service-untuk-customer-satisfaction/
[5] A. F. Fauzi and Y. J. Purnomo, “The influence of service quality and consumer value on customer satisfaction,” *Research Horizon*, vol. 3, no. 4, pp. 421–432, Aug. 2023. [Online]. Available:
https://journal.lifescifi.com/index.php/RH/article/view/151/122

[6] K. Viadi, “Analysis the impact of service quality, customer satisfaction, customer commitment on word of mouth,” *Social Economics and Ecology International Journal*, vol. 5, no. 2, May 2023, doi: 10.21512/seeij.v5i2.10031.
[Online]. Available:
https://journal.binus.ac.id/index.php/SEEIJ/article/view/10031/4772
[7] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer,
“SMOTE: Synthetic minority over-sampling technique,” *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002.
[8] L. Breiman, “Random forests,”
Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[9] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
[10] IBM, “What is ensemble learning?” IBM Think. [Online]. Available: https://www.ibm.com/think/topics/ensemble-learning
[11] IBM, “What is gradient boosting?” IBM Think, 2024. [Online]. Available: https://www.ibm.com/think/topics/gradient-boosting
[12] DataCamp, “Random forest classifier in Python,” DataCamp Tutorial. [Online]. Available: https://www.datacamp.com/tutorial/random-forests-classifier-python
[13] Trivusi, “Algoritma gradient boosting,” Trivusi, Mar. 2023. [Online]. Available: https://www.trivusi.web.id/2023/03/algoritma-gradient-boosting.html
[14] M. M. Rahman *et al.*, “Data cleaning and transformation techniques for machine learning,” *Data*, vol. 6, no. 10, Art. no. 257, Oct. 2021. [Online]. Available: https://www.mdpi.com/2673-2688/6/10/257

[15] N. Cheke, “Data preprocessing in machine learning: A detailed guide,” *Medium*, 2023. [Online]. Available: https://nicks-cheke44.medium.com/data-preprocessing-in-machine-learning-a-detailed-guide-c710df69073f
[16] Aiskunks, “Categorical data encoding techniques,” *Medium*, 2022. [Online]. Available: https://medium.com/aiskunks/categorical-data-encoding-techniques-d6296697a40f
[17] J. Brownlee, “Feature scaling in practice: What works and what doesn’t,” *Machine Learning Mastery*, 2016. [Online]. Available: https://machinelearningmastery.com/feature-scaling-in-practice-what-works-and-what-doesnt/
[18] C. Maklin, “Synthetic minority over-sampling technique (SMOTE),” *Medium*, May 14, 2022. [Online]. Available: https://medium.com/@corymaklin/synthetic-minority-over-sampling-technique-smote-7d419696b88c