Case Study Solution: Customer Service Vertical

Understanding of the Problem

The case study focuses on enhancing customer service efficiency by analyzing historical data. The objective is to predict certain customer behaviors based on features such as age, location, portfolio type, and interaction history. This prediction helps in personalizing customer interactions, leading to improved customer satisfaction and retention.

Approach

1. **EDA:** Explore data distribution, identify missing values, and detect outliers.

2. **Data Imputation:** Handle missing data using SimpleImputer with appropriate strategies.

3. **Feature Engineering:** Encode categorical variables and scale numerical features.

4. **Class Imbalance:** Apply SMOTE to balance the dataset.

5. **Model Training:** Train and evaluate the model, using cross-validation and metrics like accuracy and F1-score.

6. **Hyperparameter Tuning:** Optimize model performance with GridSearchCV or RandomizedSearchCV.

7. **Deployment:** Deploy the model and monitor its performance in production.

Expected Benefit (Value)

The solution enables accurate predictions of customer behaviors, allowing the customer service team to proactively address potential issues. This approach improves customer satisfaction and retention by providing personalized and efficient service.

Algorithm Used

1. **Imputation:** Apply SimpleImputer with the most frequent strategy to handle missing values.

2. **Encoding:** Use OneHotEncoder to convert categorical features into a numerical format.

3. **Scaling:** Standardize numerical features using StandardScaler.

4. **Balancing:** Address class imbalance with SMOTE to create synthetic samples for the minority class.

5. **Modeling:** Train a machine learning model to predict customer behavior.

Final Evaluation Metric

The model's performance was evaluated using metrics such as accuracy and F1 score, ensuring a balanced measure of precision and recall, particularly in the context of class imbalance.

Implementation

Missing Values Heatmap

Correlation Matrix