Personal Loan Acquisition using Naive Bayes Classifier based Learning Model

Group Number: 14

Members:

21CS10042 - Meduri Harshith Chowdary

23CS60R30 - Dangodra Gautam Bharat Geeta

Project Code: PLNB

Project Title: Personal Loan Acquisition using Naive Bayes Classifier based Learning Model

1. Objective:

The bank aims to expand its customer base, primarily comprising liability customers with diverse deposit amounts. With a modest number of borrowers, the bank seeks to increase this base to enhance loan business and interest revenue. Following a successful campaign yielding over a 9% conversion rate among liability customers, the retail marketing department aims to refine targeting strategies to improve success rates within budget constraints. To achieve this, the department plans to develop a model to identify potential clients inclined to purchase loans, thus increasing success ratios and reducing campaign costs.

2. Dataset Description:

personal loan offered during the bank's previous campaign. It includes unique customer IDs, age, professional experience, annual income, home address ZIP code, family size, average monthly credit card spending, education level, house mortgage value, and binary indicators for personal loan acceptance, securities account, CD account, online banking usage, and credit card ownership. These attributes offer insights into customer characteristics and behaviors, aiding in analyzing patterns and predicting loan acceptance likelihood.

3. Implementation Details:

3.1 Data Preprocessing:

- Loading Dataset: The dataset was loaded using the Pandas library's read_csv function, converting it into a DataFrame.
- Preprocessing: String columns were converted to numeric values. For the 'Personal Loan' column, values were converted to integers.

3.2 Dataset Splitting:

- The dataset was randomly split into training and testing sets using a custom implementation of the train_test_split function.

3.3 Gaussian Naive Bayes Classifier:

- Custom Implementation: We developed a custom Gaussian Naive Bayes classifier. In the ‘CustomGaussianNaiveBayes’ class, methods were implemented for fitting the model (fit) and making predictions (predict). Additionally, a method for calculating the probability density function (pdf) was defined.
- Training: During training, the class calculates class priors, mean, and standard deviation for each feature based on the training data.
- Prediction: For prediction, the classifier computes posterior probabilities for each class using the Gaussian distribution and selects the class with the highest probability.
- Evaluation Metrics: We implemented functions for evaluating accuracy and calculating precision, recall, and F1-score metrics. Additionally, we processed the classification report to extract relevant metrics.

3.4 Comparison with scikit-learn:

- We compared the results of our custom implementation with scikit-learn's Gaussian Naive Bayes classifier.
- Training and Testing: Both models were trained and tested on the same dataset split.
- Evaluation: Evaluation metrics such as accuracy, precision, recall, and F1-score were calculated for both models. The results were compared to ensure consistency.

3.5 Results for 80/20 Split:

- Both our custom implementation and scikit-learn's Gaussian Naive Bayes classifier yielded similar results:
- Accuracy: 0.73
- Precision: 0.73 (class 0), 0.80 (class 1)
- Recall: 0.98 (class 0), 0.19 (class 1)
- F1-score: 0.83 (class 0), 0.31 (class 1)

4. Conclusion:

In conclusion, our project successfully addresses the objective of predicting potential loan acceptance among bank customers. The Gaussian Naive Bayes classifier, implemented from scratch, provides reliable predictions, enabling targeted marketing efforts and optimizing campaign success rates. Future work may involve further refining the model through continued hyperparameter tuning and exploring additional feature engineering techniques.