**Comparison of Machine Learning Models for Loan Default Prediction**

**Results and Analysis**

**Experimental Setup**

In this study, three machine learning models were implemented and evaluated for predicting loan defaults:

1. Logistic Regression
2. Naive Bayes (Gaussian)
3. Neural Network (Multi-layer Perceptron)

These models were trained using scikit-learn implementations with the following configurations:

models = {

"Logistic Regression": LogisticRegression(max\_iter=1000),

"Naive Bayes": GaussianNB(),

"Neural Network": MLPClassifier(max\_iter=300, random\_state=42)

}

The dataset was split into training and testing sets, and each model was evaluated using standard classification metrics.

**Performance Metrics**

Table 1 presents the performance metrics for all three models:

| **Model** | **Accuracy** | **Precision** | **Recall** | **F1 Score** | **ROC AUC** |
| --- | --- | --- | --- | --- | --- |
| Logistic Regression | 0.8901 | 0.7663 | 0.7308 | 0.7482 | 0.9483 |
| Naive Bayes | 0.7369 | 0.4590 | 0.9965 | 0.6285 | 0.9360 |
| Neural Network | 0.9130 | 0.8460 | 0.7463 | 0.7930 | 0.9645 |

**Confusion Matrices**

The confusion matrices for each model provide deeper insights into their classification performance:

**Logistic Regression**

* True Negatives: 6542
* False Positives: 448
* False Negatives: 541
* True Positives: 1469

**Naive Bayes**

* True Negatives: 4629
* False Positives: 2361
* False Negatives: 7
* True Positives: 2003

**Neural Network**

* True Negatives: 6717
* False Positives: 273
* False Negatives: 510
* True Positives: 1500

**Analysis of Results**

**Neural Network Performance**

The Neural Network model demonstrated the best overall performance with the highest accuracy (91.30%), precision (84.60%), F1 score (0.7930), and ROC AUC (0.9645). This indicates that the model effectively balances identifying defaults while minimizing false predictions. The confusion matrix shows that this model had the lowest number of false positives (273), suggesting it is most conservative in predicting defaults.

**Naive Bayes Analysis**

The Naive Bayes classifier exhibited remarkable recall performance (99.65%), meaning it successfully identified almost all actual default cases (with only 7 false negatives). However, this came at the cost of precision (45.90%), as evidenced by the high number of false positives (2361). This model would be appropriate in scenarios where missing a potential default carries extremely high cost, and the organization can manage a higher number of false alarms.

**Logistic Regression Analysis**

Logistic Regression showed balanced performance with good accuracy (89.01%) and reasonable precision (76.63%) and recall (73.08%). It represents a middle ground between the more aggressive Naive Bayes approach and the more precise Neural Network model. Its F1 score of 0.7482 indicates good overall balance between precision and recall.

**ROC Analysis**

All three models demonstrated excellent discriminative ability with ROC AUC scores above 0.93. The Neural Network achieved the highest ROC AUC (0.9645), closely followed by Logistic Regression (0.9483) and Naive Bayes (0.9360). These high values indicate that all models can effectively separate default and non-default classes across various threshold settings.

**Discussion**

The results clearly demonstrate the trade-offs inherent in model selection for loan default prediction. The Naive Bayes model maximizes recall at the expense of precision, making it suitable for risk-averse scenarios. The Neural Network excels in overall accuracy and precision, making it appropriate when balanced performance is desired. The Logistic Regression model offers good all-around performance with the added benefit of interpretability.

For financial institutions concerned with managing loan default risk, the choice of model depends on their specific business objectives:

1. If the primary concern is to avoid missing any potential defaults (even at the cost of falsely flagging good loans), the Naive Bayes approach would be preferred.
2. If a balance between correctly identifying defaults and minimizing false alarms is desired, the Neural Network provides optimal performance.
3. For scenarios where model interpretability is critical for regulatory or business reasons, Logistic Regression offers the best compromise between performance and transparency.

The high ROC AUC values across all models suggest that the feature set used in this study contains strong predictive signals for loan default behavior. This indicates that the preprocessing steps and feature engineering were effective.

**Conclusion**

This study compared three machine learning approaches for loan default prediction. The Neural Network model demonstrated superior overall performance with the highest accuracy, precision, F1 score, and ROC AUC metrics. However, each model exhibited distinct strengths that may be valuable in different business contexts. The Naive Bayes model's exceptional recall makes it suitable for highly risk-averse scenarios, while Logistic Regression offers a good balance of performance and interpretability.

These findings suggest that financial institutions should consider deploying multiple models in parallel, potentially using the Naive Bayes model as an initial screening tool to catch almost all potential defaults, followed by more precise models like the Neural Network to reduce false positives among the flagged cases. Future work could explore ensemble methods that leverage the complementary strengths of these different approaches.