**Loan Default Prediction – Model Performance Report**

**->Objective**

The goal of this project is to predict whether a loan application will be approved (`Loan\_Status`) based on applicant details such as income, education, credit history, and more. We trained and evaluated multiple machine learning models to identify the best-performing one.

**->Dataset Summary**

- Dataset: Loan Prediction Dataset (Kaggle)

- Target Variable: `Loan\_Status` (`Y` = Approved, `N` = Not Approved)

- Total Records: 614

- Features: 13 (categorical + numerical)

- Preprocessing: Missing value imputation, label encoding, scaling

**->Models Trained**

1. Decision Tree Classifier

2. Random Forest Classifier

3. Gradient Boosting Classifier

Each model was evaluated using the following metrics:

- Accuracy

- Precision

- Recall

- Confusion Matrix

**-> Hyperparameter Tuning:**

I used GridSearchCV on the Gradient Boosting Classifier to optimize parameters such as:

- `n\_estimators`

- `max\_depth`

- `learning\_rate`

**->Best Parameters Found:**

```python

{'n\_estimators': 100, 'max\_depth': 3, 'learning\_rate': 0.1}

After tuning, the Gradient Boosting model slightly improved in performance with better generalization and reduced overfitting.

**->** **Best Model Choice: Gradient Boosting Classifier**

* Highest accuracy and recall among all models
* Performs well with imbalanced datasets (by focusing on difficult samples)
* Improved further with GridSearchCV
* Offers feature importance, aiding interpretability

**->Top Influential Features:**

* Credit\_History
* ApplicantIncome
* LoanAmount
* Education
* Property\_Area

Gradient Boosting stands out as the most effective model for predicting loan approvals in this dataset. With hyperparameter tuning, its predictive performance becomes even more reliable, making it suitable for real-world deployment scenarios like banking loan systems.