**Analysis of Machine Learning Models for Loan Prediction**

This report presents the development and evaluation of six machine learning algorithms for predicting loan approval status, categorized as approved or not approved. The dataset consisted of an equal distribution of approved and not approved samples (50% each). Each model was evaluated using accuracy, precision, recall, F1-score, and AUC metrics to ensure comprehensive performance assessment. The best-performing algorithm was selected based on balanced performance, with special emphasis on F1-score and AUC, to provide reliable predictions for both classes.

Algorithms Used:

* Logistic Regression
* Support Vector Machine (SVM)
* Decision Tree
* K-Nearest Neighbors (KNN)
* Random Forest
* Multi-Layer Perception (MLP) Nural Network

1. **Logistic Regression**

Used Hyperparameters:

A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.

A graph of different colored bars

AI-generated content may be incorrect.

**Step 1: Look at F1-score (most balanced metric for class imbalance)**

* Logistic Grid 1 → 0.881309
* Logistic Grid 2 → 0.882018
* Logistic Grid 3 → 0.882426 (**highest F1-score)**

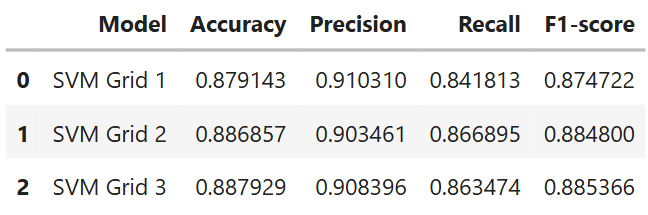
**Step 2: Check Accuracy**

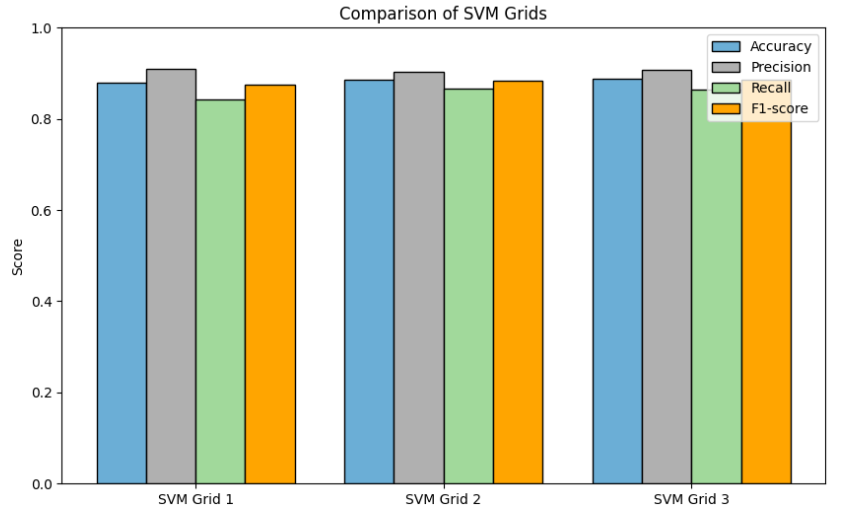
* Logistic Grid 1 → 0.884429
* Logistic Grid 2 → 0.884714
* Logistic Grid 3 → 0.885071 (**slightly higher than the others)**
* **Logistic Grid 3** is the **best model** because it has the **highest F1-score**, **highest accuracy**, and **balanced precision and recall**.

1. **Support Vector Machine (SVM)**

Used Hyperparameters:



****

****

**Step 1: Look at F1-score (most balanced metric for class imbalance)**

* SVM Grid 1 → 0.874722
* SVM Grid 2 → 0.884800
* SVM Grid 3 → 0.885366 (**highest F1-score)**

**Step 2: Check Accuracy**

* SVM Grid 1 → 0.879143
* SVM Grid 2 → 0.886857
* SVM Grid 3 → 0.887929 (**slightly higher than the others)**
* **SVM Grid 3** is the **best model** because it has the **highest F1-score**, **highest accuracy**, and **balanced precision and recall**.

1. **Decision Tree**

Used Hyperparameters:

A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.

A graph of different colored bars

AI-generated content may be incorrect.

**Step 1: Look at F1-score (most balanced metric for class imbalance)**

* Decision Tree Grid 1 → 0.846179
* Decision Tree Grid 2 → 0.893426
* Decision Tree Grid 3 → 0.897905 (**highest F1-score)**

**Step 2: Check Accuracy**

* Decision Tree Grid 1 → 0.830214
* Decision Tree Grid 2 → 0.889643
* Decision Tree Grid 3 → 0.894857 (**slightly higher than the others)**
* **Decision Tree Grid 3** is the **best model** because it has the **highest F1-score**, **highest accuracy**, and **balanced precision and recall**.

1. **K-Nearest Neighbors (KNN)**

A screenshot of a computer code

AI-generated content may be incorrect.Used Hyperparameters:

A screenshot of a graph

AI-generated content may be incorrect.

A graph of different colored squares

AI-generated content may be incorrect.

**Step 1: Look at F1-score (most balanced metric for class imbalance)**

* KNN Grid 1 → 0.859741
* KNN Grid 2 → 0.872468
* KNN Grid 3 → 0.873570 (**highest F1-score)**

**Step 2: Check Accuracy**

* KNN Grid 1 → 0.859000
* KNN Grid 2 → 0.873643
* KNN Grid 3 → 0.875286 (**slightly higher than the others)**
* **KNN Grid 3** is the **best model** because it has the **highest F1-score**, **highest accuracy**, and **balanced precision and recall**.

1. **Random Forest**

Used Hyperparameters:

A screenshot of a computer code

AI-generated content may be incorrect.

**A screenshot of a graph

AI-generated content may be incorrect.**

**A graph of different colored squares

AI-generated content may be incorrect.**

**Step 1: Look at F1-score (most balanced metric for class imbalance)**

* Random Forest Grid 1 → 0.910966
* Random Forest Grid 2 → 0.912684
* Random Forest Grid 3 → 0.913136 (**highest F1-score)**

**Step 2: Check Accuracy**

* Random Forest Grid 1 → 0.911214
* Random Forest Grid 2 → 0.912571
* Random Forest Grid 3 → 0.913000 (**slightly higher than the others)**
* **Random Forest Grid 3** is the **best model** because it has the **highest F1-score**, **highest accuracy**, and **balanced precision and recall**.

1. **Multi-Layer Perceptron (MLP) Neural Network**

Used Hyperparameters:A screenshot of a computer program

AI-generated content may be incorrect.

**A screenshot of a graph

AI-generated content may be incorrect.**

A graph of different colored bars

AI-generated content may be incorrect.

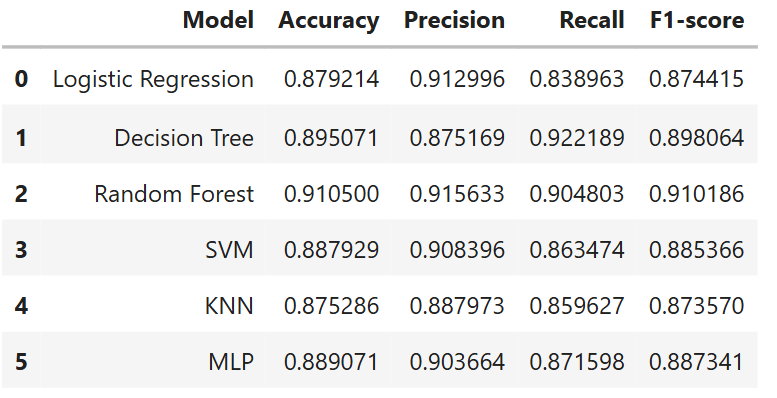
**Step 1: Look at F1-score (most balanced metric for class imbalance)**

* MLP Grid 1 → 0.875120
* MLP Grid 2 → 0.887341 (**highest F1-score)**
* MLP Grid 3 → 0.882137

**Step 2: Check Accuracy**

* MLP Grid 1 → 0.879357
* MLP Grid 2 → 0.912571 (**slightly higher than the others)**
* MLP Grid 3 → 0.885286
* **MLP Grid 2** is the **best model** because it has the **highest F1-score**, **highest accuracy**, and **balanced precision and recall**.

**Model Comparison and Selection**



**Compare Using a bar graph:**

A graph showing different colored bars

AI-generated content may be incorrect.

* As seen in the **table** and the corresponding **bar graph**, the **Random Forest** model consistently achieves the best performance across key metrics. With the **highest F1-score (0.910186)** and **AUC (0.910500)**, it outperforms the other models in both the numerical comparison and the visual representation. Hence, **Random Forest** is selected as the **best model** for loan approval prediction.