

SongNet

Karthik Prabhu
J Guruprasad





Problem Statement

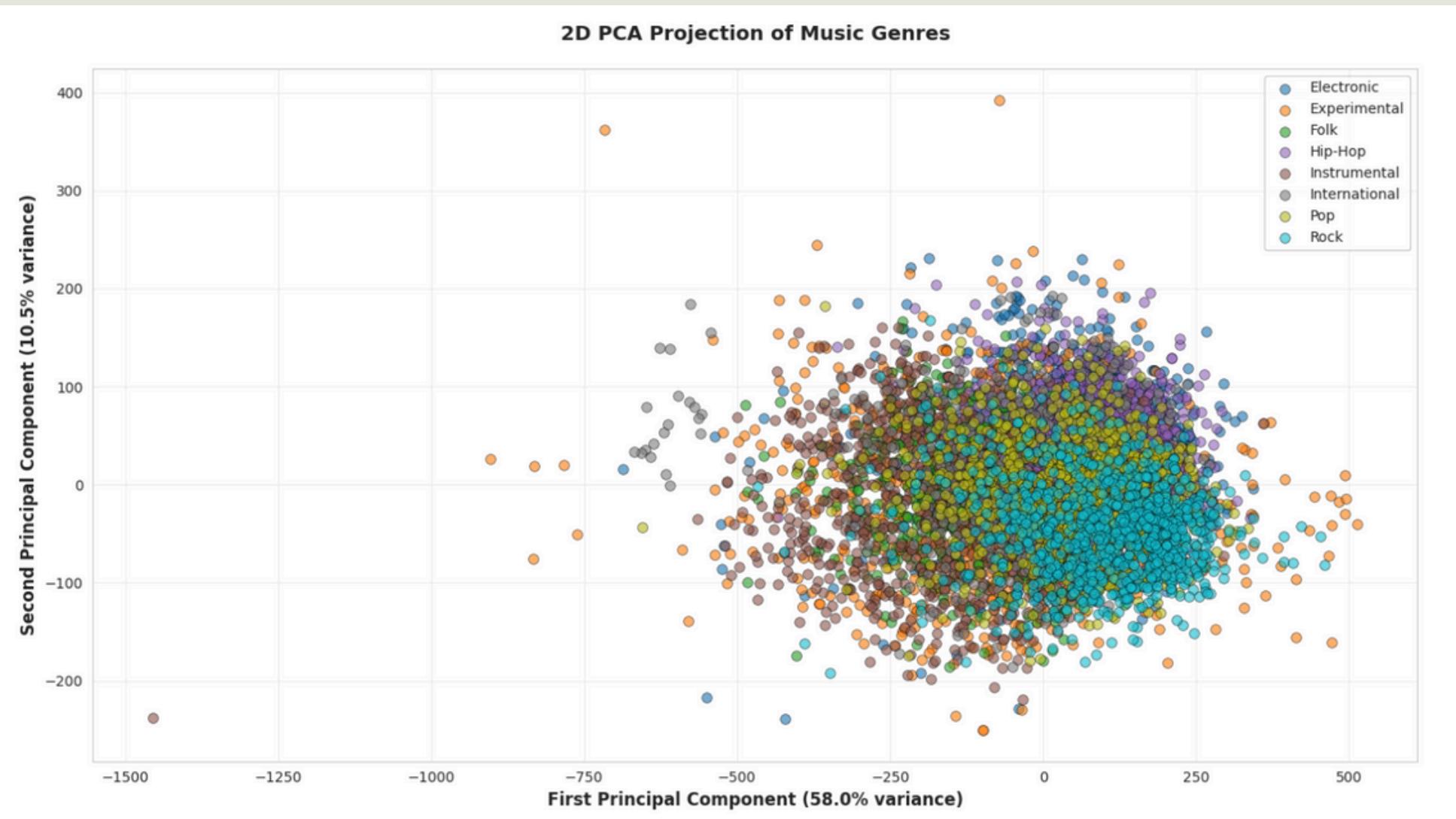
- **Goal:** Build a machine learning model to classify songs into 8 genres
- Hip-Hop, Pop, Folk, Experimental, Rock, International, Electronic, Instrumental
- **Input:** 140 MFCC features from audio signals
- **Output:** Predicted genre label

Dataset Overview

- **Dataset:** FMA-Small (8,000 tracks; 1,000 per genre)
- **Balanced dataset ensures fair model evaluation**
- **Visualization:** PCA revealed overlapping clusters

Data Preprocessing

- **Split:**
 - **Train:** 70% (5600 samples)
 - **Validation:** 20% (1600 samples)
 - **Test:** 10% (800 samples)
- **Scaling:** **StandardScaler** for normalization
→ Improves performance of distance-based algorithms (SVM, KNN)



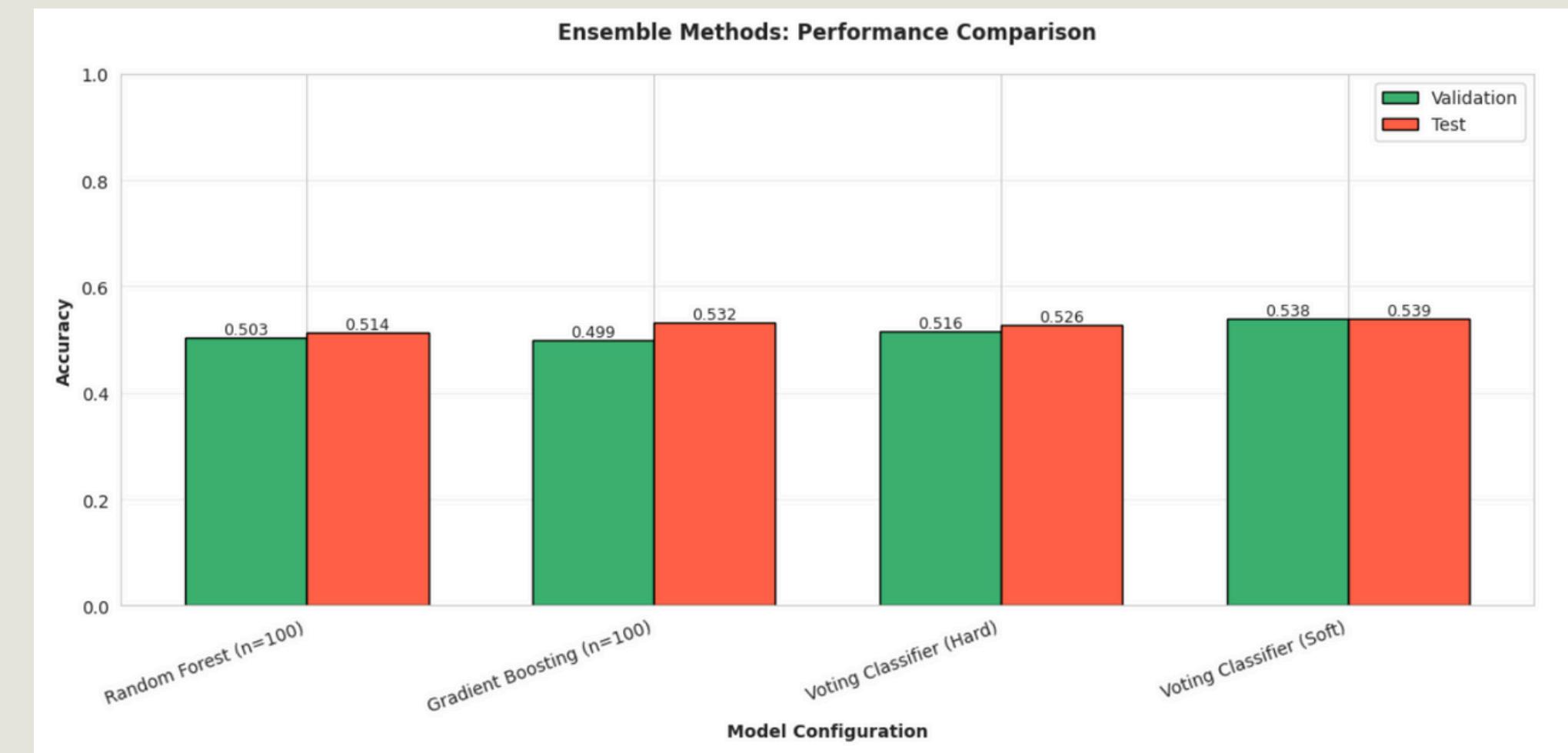
Models Implemented

Baseline Models

- K-Nearest Neighbors (KNN)
- Logistic Regression
- Multilayer Perceptron (MLP)
- Support Vector Machine (SVM)

Ensemble Learning

- Random Forest
- Gradient Boosting
- Voting Classifiers (Hard, Soft)



=====

STEP 4: OVERALL RESULTS COMPARISON

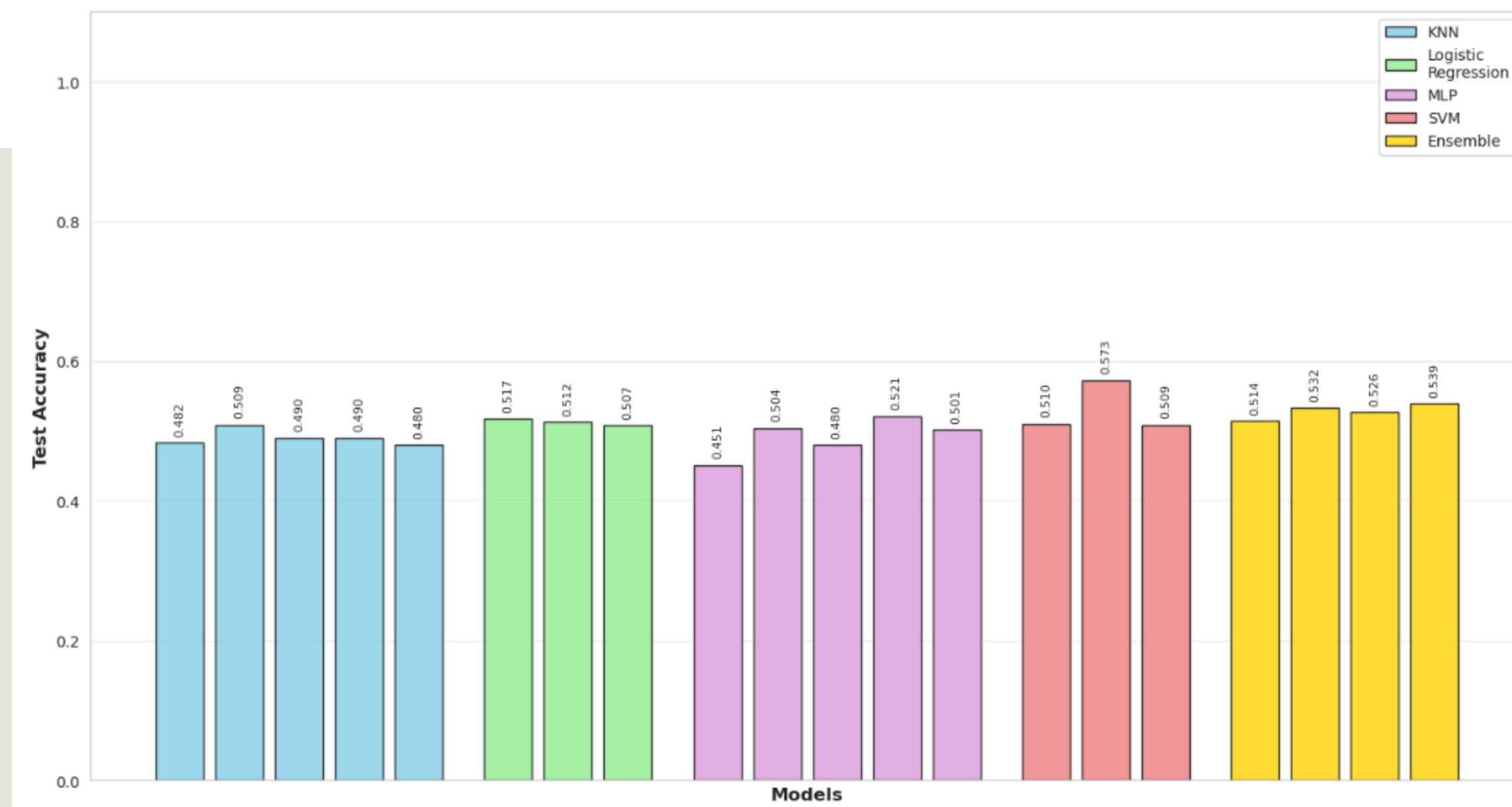
=====

Performance Ranking (by Test Accuracy):

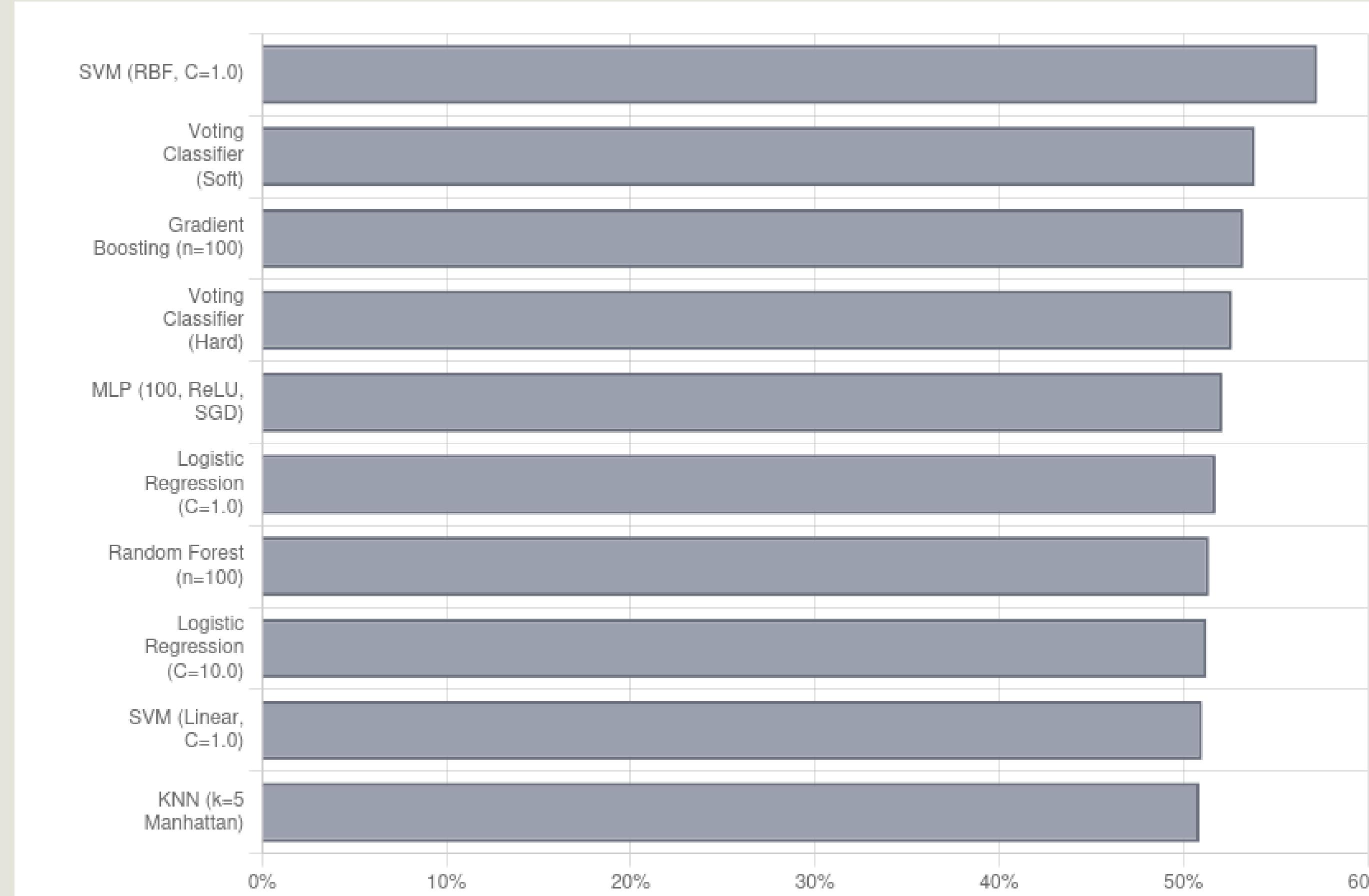
Rank	Model	Test Acc	Val Acc	F1-Score
1	SVM (RBF, C=1.0)	0.5725	0.5381	0.5691
2	Voting Classifier (Soft)	0.5387	0.5381	0.5356
3	Gradient Boosting (n=100)	0.5325	0.4994	0.5307
4	Voting Classifier (Hard)	0.5262	0.5156	0.5189
5	MLP (100, ReLU, SGD)	0.5212	0.5269	0.5185
6	Logistic Regression (C=1.0)	0.5175	0.4894	0.5108
7	Random Forest (n=100)	0.5138	0.5031	0.5042
8	Logistic Regression (C=10.0)	0.5125	0.4850	0.5059
9	SVM (Linear, C=1.0)	0.5100	0.4825	0.5058
10	KNN (k=5, Manhattan)	0.5088	0.4731	0.5005
11	SVM (Linear, C=10.0)	0.5088	0.4763	0.5045
12	Logistic Regression (C=100.0)	0.5075	0.4869	0.4997
13	MLP (100,50, ReLU, Adam)	0.5038	0.4819	0.5024
14	MLP (100,50,50, ReLU)	0.5012	0.4756	0.4982
15	KNN (k=7, Euclidean)	0.4900	0.4819	0.4830
16	KNN (k=7, minkowski)	0.4900	0.4819	0.4830
17	KNN (k=5, Euclidean)	0.4825	0.4744	0.4734
18	KNN (k=9, minkowski)	0.4800	0.4788	0.4733
19	MLP (100, Tanh, Adam)	0.4800	0.4394	0.4789
20	MLP (100, ReLU, Adam)	0.4512	0.4550	0.4506

Model configurations

Overall Model Performance Comparison



Performance Ranking



The Champion: SVM with RBF Kernel

Key Performance Metrics

The Support Vector Machine (SVM) using a Radial Basis Function (RBF) kernel proved to be the most effective model. Its test accuracy of 57.25% demonstrates its superior ability to generalize to new, unseen music data compared to other models.

Test Accuracy: 57.25%



Correctly Classified



Incorrectly Classified

53.81%

Validation Accuracy

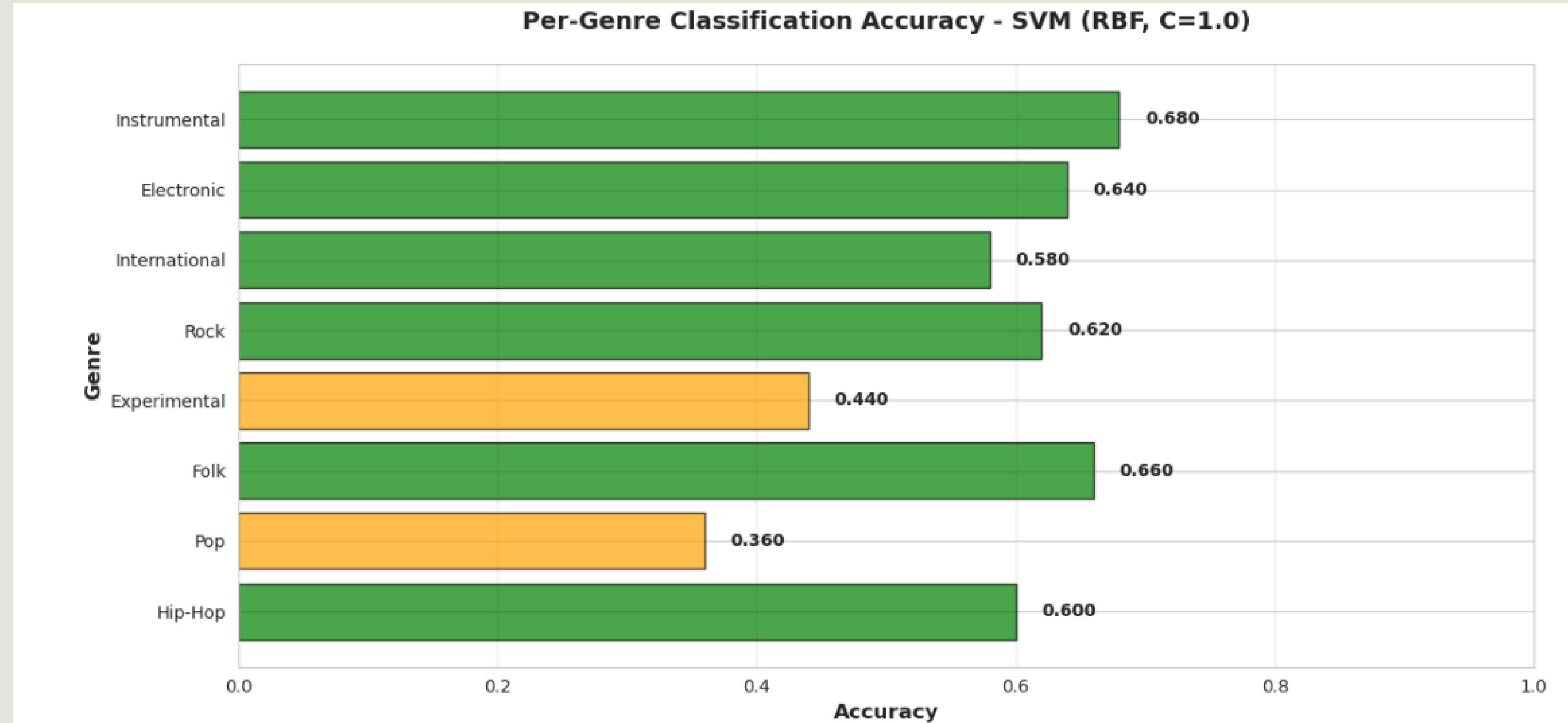
75.93%

Training Accuracy

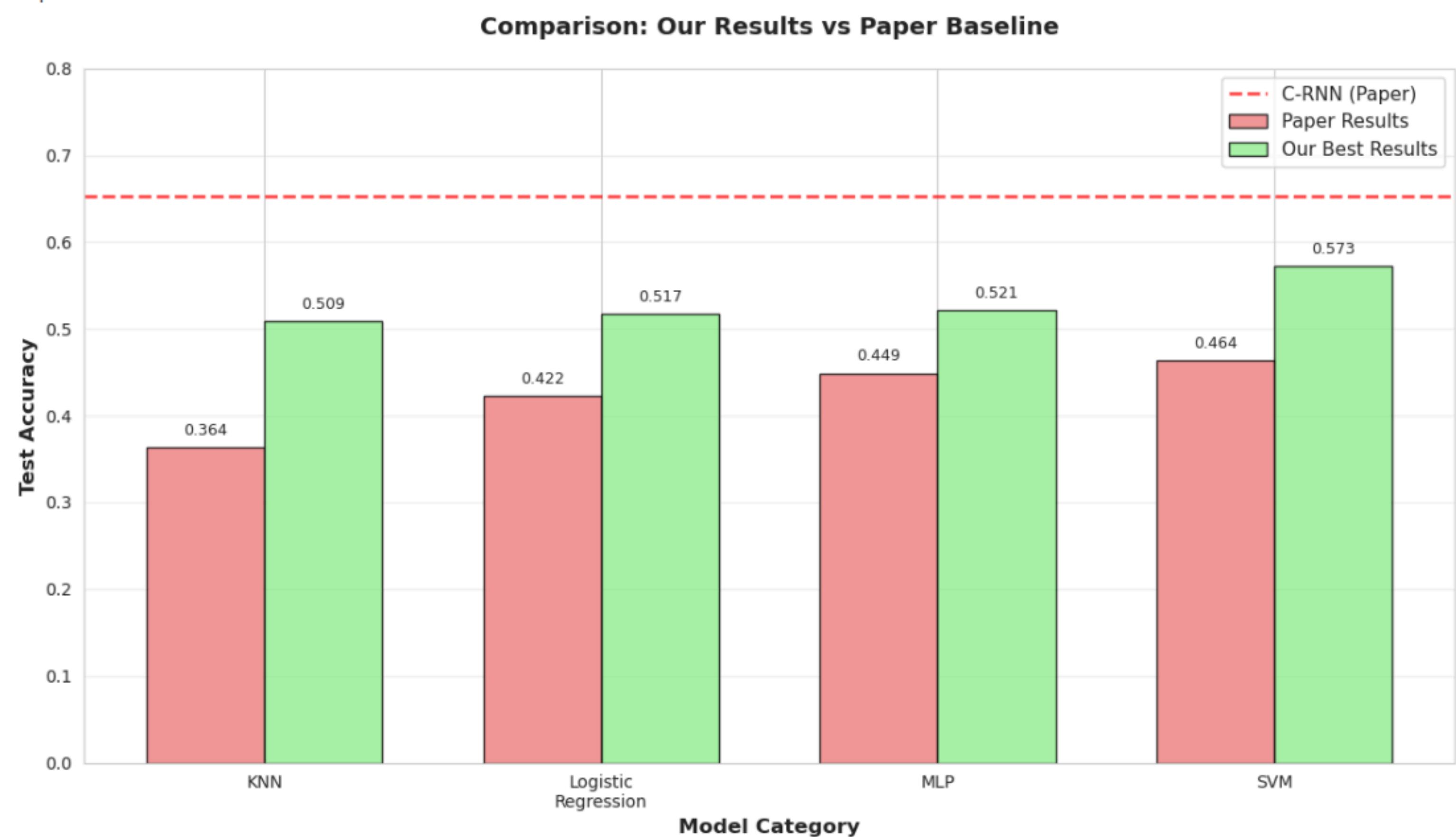
56.91%

Test F1-Score (Weighted)

More details on the SVM model



Paper vs Our results



Thank You

