Model Evaluation Summary

Overview

We evaluated three different machine learning classifiers for a multiclass sentiment classification task involving the labels: negative, neutral, and positive. The models trained and tested include:

1. **Logistic Regression**

2. **Support Vector Machine (SVM)** with LinearSVC + CalibratedClassifierCV

3. **Random Forest Classifier**

Each model was trained on a preprocessed dataset and evaluated using the same test set. Key performance metrics such as Accuracy, F1 Score, Precision, Recall, and Confusion Matrix were collected.

Model Configurations

| Model	Configuration Details
Logistic Regression	solver='lbfgs', default settings
SVM (Calibrated)	LinearSVC with CalibratedClassifierCV
Random Forest	n_estimators=100, max_depth=None, random_state=0

Performance Comparison Table

Metric	Logistic Regression	SVM (Calibrated)	Random Forest
Accuracy	0.7202	0.7196	0.7100
F1 Score	0.6451	0.6343	0.6033
Negative F1	0.30	0.26	0.11
Neutral F1	0.05	0.01	0.02
Positive F1	0.83	0.83	0.83

Confusion Matrix Highlights

- Positive sentiment is predicted with high recall in all models (~97–100%).

- Negative and neutral classes are often misclassified as positive.

- SVM performs slightly better than Random Forest for negative and neutral precision but not recall.

Visual Comparison
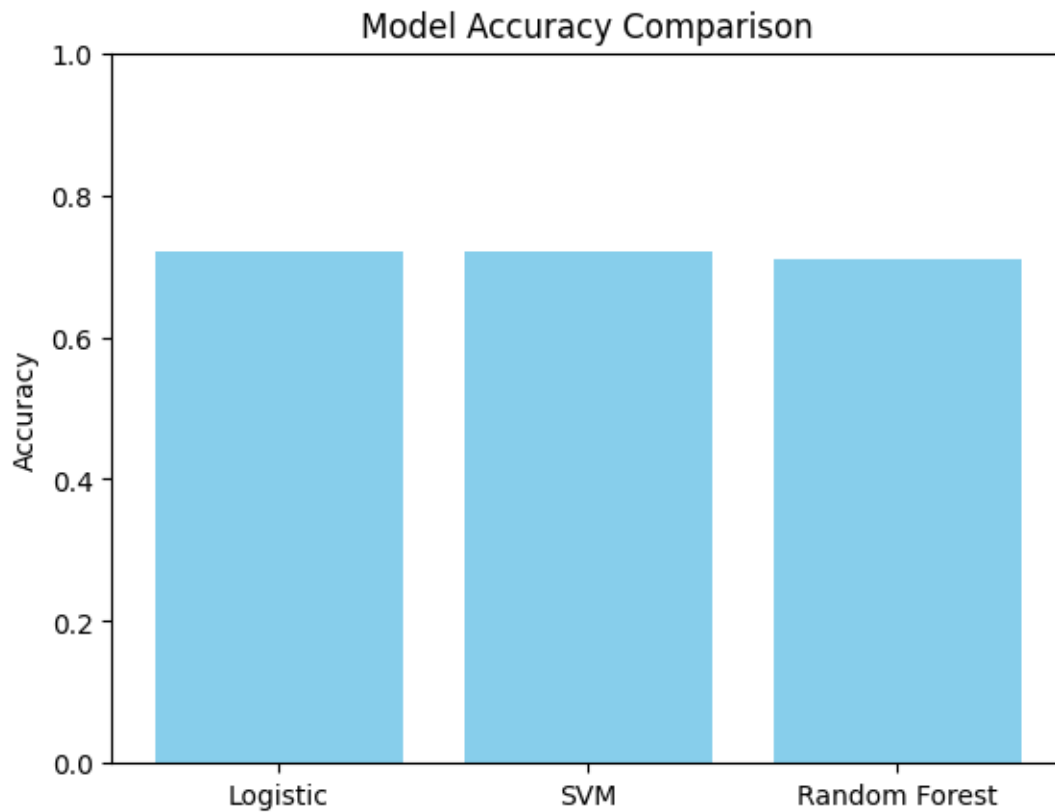
Recommendation

Logistic Regression and SVM (Calibrated) show similar overall accuracy and macro F1 scores, with Logistic Regression slightly outperforming in terms of handling negative and neutral classes.

Most Promising Model: Logistic Regression, due to its simplicity and relatively better balance across all classes.