Model Metrics – Sentiment Classification

1. Overall Model Performance (Test Set)

Model	Test Samples	Accuracy	Macro Precision	Macro Recall	Macro F1	Weighted Precision	Weighted Recall	Weighted F1
Logistic Regression (TF-IDF)	158175	0.88	0.75	0.67	0.69	0.86	0.88	0.87
Linear SVM (TF-IDF)	158175	0.88	0.76	0.66	0.67	0.86	0.88	0.86
Naive Bayes (TF-IDF)	158175	0.85	0.72	0.6	0.62	0.83	0.85	0.83

Interpretation: Macro-averaged metrics are emphasized due to class imbalance. Logistic Regression and Linear SVM perform similarly (accuracy ≈ 0.88) with comparable macro F1.

2. Per-Class Performance Metrics

Model	Class	Precision	Recall	F1	Support
Logistic Regression (TF-IDF)	negative	0.79	0.79	0.79	23392
Logistic Regression (TF-IDF)	neutral	0.54	0.26	0.35	14686
Logistic Regression (TF-IDF)	positive	0.92	0.97	0.94	120097
Linear SVM (TF-IDF)	negative	0.77	0.81	0.79	23392
Linear SVM (TF-IDF)	neutral	0.59	0.19	0.29	14686
Linear SVM (TF-IDF)	positive	0.91	0.98	0.94	120097
Naive Bayes (TF-IDF)	negative	0.79	0.68	0.73	23392
Naive Bayes (TF-IDF)	neutral	0.51	0.13	0.2	14686
Naive Bayes (TF-IDF)	positive	0.87	0.98	0.92	120097