1.Models Developed

1. Logistic regression
2. Random forest
3. Decision tree
4. SVM
5. K-Nearest Neighbors
6. XGBoost

Baseline models

Model Name	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.8752	0.85	0.80	0.82
Decision Tree Classifier	0.8367	0.75	0.83	0.79
Random Forest Classifier	0.8829	0.86	0.81	0.84
SVM (Support Vector)	0.8582	0.84	0.76	0.80
KNN Classifier	0.8428	0.78	0.80	0.79
XGBoost	0.8737	0.82	0.83	0.83

Tuned models

Model Name	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.8767	0.85	0.81	0.83
Decision Tree Classifier	0.8336	0.77	0.77	0.77
Random Forest Classifier	0.8783	0.85	0.81	0.83
SVM (Support Vector)	0.8629	0.84	0.77	0.80
KNN Classifier	0.8459	0.78	0.81	0.79
XGBoost	0.8675	0.83	0.80	0.82

2.Final Selected Model

These results demonstrate that the XG boost model delivers the highest CV mean accuracy and balanced performance among the tested algorithms, making it the most suitable model for dropout prediction in this study.

Model	Test Accuracy	Precision	Recall	F1-Score
XGBoost(Baseline)	0.8737	0.82	0.83	0.83
XGBoost(Tuned)	0.8675	0.83	0.80	0.82

XG boost CV mean accuracy: 0.877204