GUIDING THE WAY: NAVIGATING THE SOLUTION MAP FOR OPTIMIZED OUTCOME

Data Preprocessing

Data Cleaning
Data Visualization
Oversampling

Model Development

Feature selection
Predictive Modeling
Model Tuning

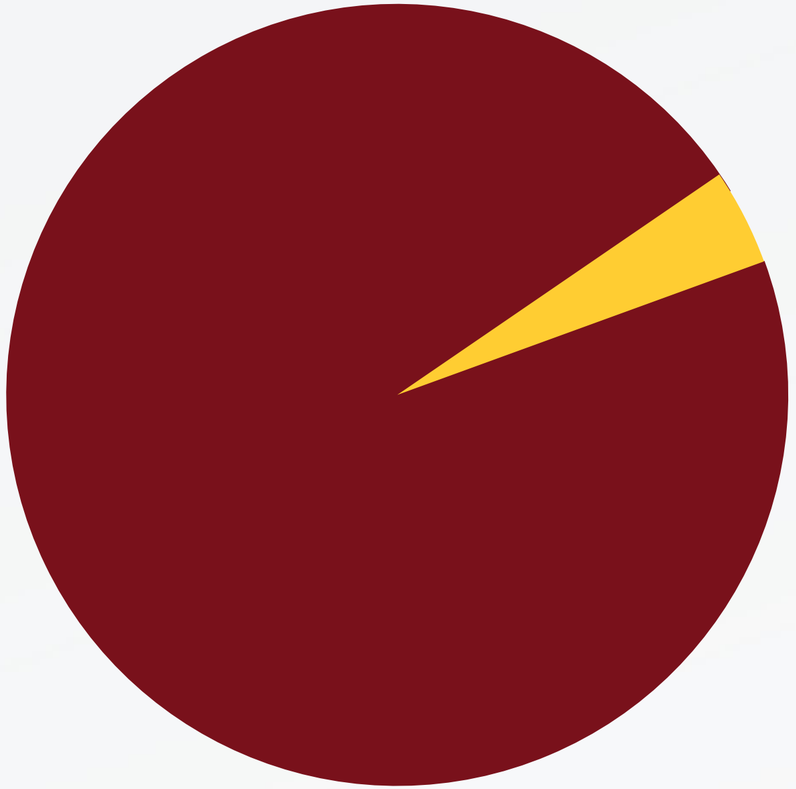
Model Performance

Precision
Recall
F-measure

THE DATA IS IMBALANCED, AND HAS MORE NON-ADOPTERS THAN ADPOTER

NON-ADOPTER

96%

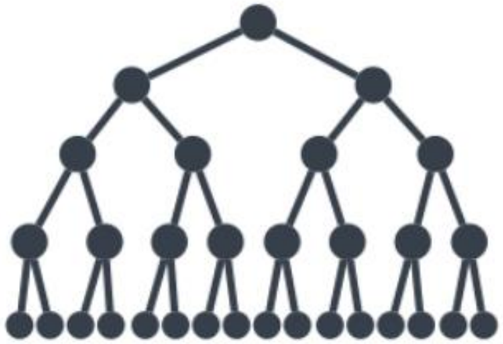


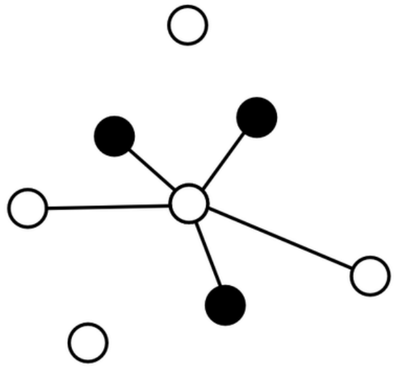
ADOPTER

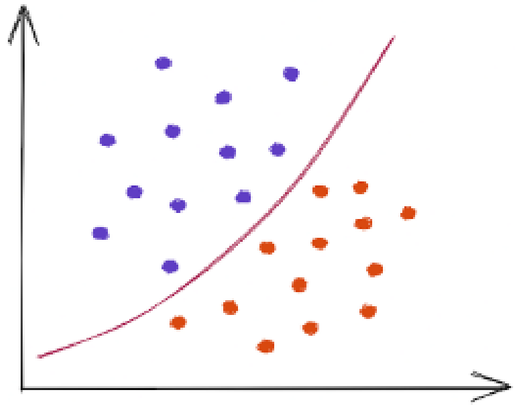
4%

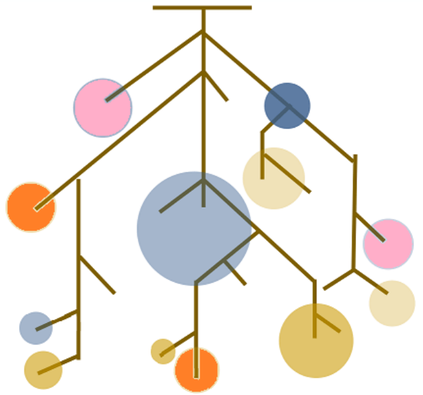
EXPLORING THREE PREDICTIVE MODELING TECHNIQUES

Decision Tree	K-NN	Logistic Regression	Random Forest







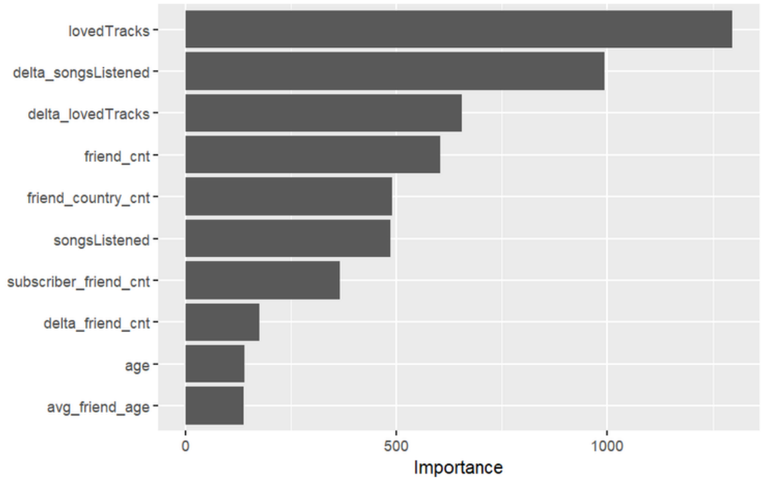


AMONG THE THREE MODELS, K-NN PERFORMED THE BEST PERFORMANCE

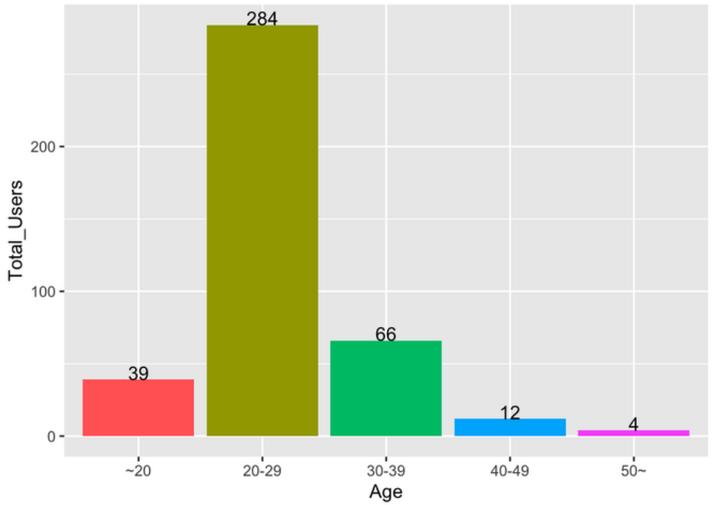
Model Evaluation Index	Decision Tree	K-NN	Logistic Regression
Precision	0.55	**0.71**	0.58
Recall	0.28	**0.91**	0.07
F-measure	0.37	**0.80**	0.14

WHAT DOES THE K-NN MODEL TELL US?

Important Features Effecting User Preferences



Most Potential Adopters are Around age 20-29



Average Number of loved tracks among the most potential adopters is 185

185