KNN

* Leave One out CV had best performance on unbalanced data set
* Using SMOTE to balance the dataset improved performance, with stratified k-fold cross-validation showed the best results.

Evaluation metrics using SMOTE and stratified CV ( Without Normalization)

A screenshot of a computer

Description automatically generated

Evaluation metrics using SMOTE and stratified CV ( With Normalization)

A screenshot of a computer

Description automatically generated

Feature Importance score with KNN

A graph of blue bars

Description automatically generated

SVM

* SVM performed poorly after balancing and normalization. Performance decreased from 89 to 86
* Leave one out CV was used on unbalanced data set.
* Stratified CV was used for balanced data set
* A screenshot of a graph

  Description automatically generated

After Balancing and normalization

A screenshot of a computer screen

Description automatically generated

Important Features

A graph with blue bars

Description automatically generated

Decision Tree

* K fold CV performed best after balancing and normalization.
* A screenshot of a computer

  Description automatically generated

Important Features For DT

A graph of a number of blue squares

Description automatically generated

Naïve bayes algorithm performed poorly. Kfold CV was used after balancing and normalization

A screenshot of a computer

Description automatically generated

A graph of a number of blue bars

Description automatically generated

Gradient Boosting performed well with k fold CV and after applying balancing and normalization

A screenshot of a computer

Description automatically generated

A graph of a number of blue bars

Description automatically generated

LGBM

Best performed with kfold CV after applying balancing and normalization

A screenshot of a computer

Description automatically generated

A graph with different colored bars

Description automatically generated with medium confidence