**Predicting Employee Attrition (Source – Kaggle )**

Uncovering the factors that cause employees to leave the organization and explore important questions such as ‘compare average monthly income by education and attrition’ or ‘is distance from home a case for attrition’?

Build classification models to predict which employee is likely to churn and help the business to devise policies and attract back the right talent.

Evaluation is done as below .

**Feature Selection (variance and entropy based)**

Select the right features based on importance and significance.

**Feature Engineering**

Apply feature engineering techniques to see how new features can be created to improve the model. Check for Interaction.

**Model Comparison**

Apply multiple classification algorithms and compare results.

**Model Selection**

Select the best model. Model selection to be based on Accuracy, Sensitivity & Specificity and Kappa value.

**DATA : IBM HR Analytics Employee Attrition & Performance**

Data contains differnet attributes of an employee and the target variable Atrition.  EmployeeNumber is the primary key. We use this dataset to predict employee churn.

Data definitions for categorical variables:

Education 1 'Below College' 2 'College' 3 'Bachelor' 4 'Master' 5 'Doctor'

EnvironmentSatisfaction 1 'Low' 2 'Medium' 3 'High' 4 'Very High'

JobInvolvement 1 'Low' 2 'Medium' 3 'High' 4 'Very High'

JobSatisfaction 1 'Low' 2 'Medium' 3 'High' 4 'Very High'

PerformanceRating 1 'Low' 2 'Good' 3 'Excellent' 4 'Outstanding'

RelationshipSatisfaction 1 'Low' 2 'Medium' 3 'High' 4 'Very High'

WorkLifeBalance 1 'Bad' 2 'Good' 3 'Better' 4 'Best'