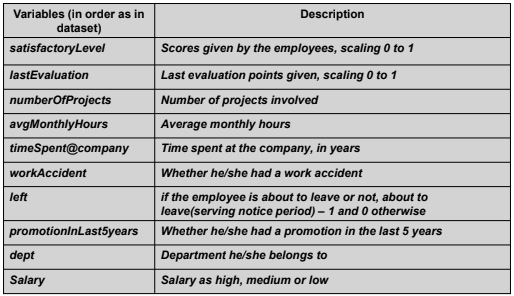
‘People Charm’, a growing company is facing a high attrition rate among their employees which in turn affects their business due to lack of expertise and experience. Their HR department is assigned the task to reduce the attrition rate by retaining employees who are about to churn out. They need to recommend special plans or strategies which will help them to retain their employees which in turn will help them to grow bigger as a company.

The file ‘People Charm Case.csv’ has several attributes as described in the table below.



Perform data analysis like missing value treatment, encoding, normalization etc, wherever necessary.

Apply following classification algorithms on the dataset:

* KNN
* Logistic Regression
* Naïve Bayes
* Decision Tree
* Random Forest
* Adaboost
* Gradient Boosting
* XG Boost

Display confusion matrix, calculate accuracy, plot ROC in all the cases.

Provide your views on the following issues:

* Which model gives best accuracy and why?
* Do you feel accuracy is a good performance evaluation metric for the given data? If yes, justify your answer. If no, justify your answer and suggest alternative metric/s.