ML Algorithm	Algorithm Comparison
k-NN	**Type**: Supervised Classification, Non-Parametric, lazy learner, discriminative. **Feature Data:** For majority of the scenario, Numeric/Quantitative sometimes categorical. **Target Variable:** Categorical **Noisy Data:** Is sensitive to noisy data and outliers are liable to get misclassified. **Missing Data:** k-NN cannot train on missing values therefore some reasonable imputation has to be performed or NA records must be deleted. **Correlated Data:** Does not affect the classification. **Parameters and Hyperparameters**: K: number of neighbors, D: distance function like Euclidean, Cosine, Manhattan for numeric feature space and Hamming for categorical feature space. **Evaluation Measures:** Prediction Accuracy**,** ROC Curve**,** Confusion Matrix, Specificity, Sensitivity etc. **Package Implementation:** class, caret.
Naïve Bayes'	**Type:** Supervised Classification, Parametric, eager learner, generative. **Feature Data:** For majority of the scenario Categorical **Target Variable:** Categorical **Noisy Data:** Averages out the noisy therefore not sensitive to noisy data. Its performance degrades more slowly as compared to other algorithms. **Missing Data :** The specific record is ignored for the frequency count. The R package implementation can also determine whether to ignore the example/record or the feature. **Correlated Data:** Does not work optimally due to class conditional independence assumption. **Parameters and Hyperparameters** :Maximum Likelihood Estimation, Bayesian Estimation, optimization of loss criterion. **Evaluation Measures**: Prediction Accuracy**,** ROC Curve, Confusion Matrix ,Sensitivity, Specificity. **Package Implementation**: e1071,naïve bayes, caret
Classification Trees	**Type:** Supervised Classification and Prediction(Dual), Non-Parametric, eager learner, discriminative, greedy learner, top down divide and conquer, bottom up pruning, recursive partitioning **Feature Data:** Numeric/Quantitative or Categorical **Target Variable:** Categorical or Numeric. **Noisy Data:** Decision trees are very sensitive to the noise in the data and even a small change in the data can throw the classification/prediction off therefore an ensemble of tree is better than just one tree. **Missing Data :** Is handled as per the package implementation for instance CART provides the functionality of assigning the missing data examples to the partition which has maximum examples already assigned to it.

	Correlated Data: Typically, it will not be affected by correlated variables and the split will occur as per the best feature in terms of info gain. **Parameters and Hyperparameters** Gini index, information gain, chi square ,pre pruning ,post pruning, assigning levels of trees **Evaluation Measures:** Prediction Accuracy**,** ROC Curve**,** Confusion Matrix ,Sensitivity, Specificity, R squared, MSE. **Package Implementation:** CART,C5.0,ID 3,rpart,caret.
Classification Rules	**Type:** Supervised Classification , Non-Parametric, eager learner, discriminative, greedy learner, bottom up separate and conquer. **Feature Data:** In most scenarios Categorical. **Target Variable:** Mostly Categorical . **Noisy Data:** Decision rules are very sensitive to the noise in the data and even a small change in the data can throw the classification/prediction off . **Missing Data :** Is handled as per the package implementation for instance CART provides the functionality of assigning the missing data examples to the partition which has maximum examples already assigned to it. **Correlated Data**: Typically, it will not be affected by correlated variables and the split will occur as per the best feature in terms of info gain. **Parameters and Hyperparameters** Gini index, information gain, chi square ,pre pruning ,post pruning, assigning levels of trees **Evaluation Measures:** Prediction Accuracy**,** ROC Curve**,** Confusion Matrix ,Sensitivity, Specificity **Package Implementation:** CART,C5.0,ID 3
Regression	**Type:** Mostly Supervised Prediction , Parametric and Non-Parametric, eager learner, discriminative. **Feature Data:** Numerical and Categorical. **Target Variable:** Mostly Numerical . **Noisy Data:** Regression is very sensitive to the noise in the data and even a small change in the data can throw the prediction off . **Missing Data :** Is handled as per the package implementation for instance caret provides the functionality of assigning the missing data examples to the partition which has maximum examples already assigned to it. **Correlated Data**: Typically, it will be affected by correlated predictor features and techniques like stepwise Regression or PCA will have to be implemented. **Parameters and Hyperparameters** Variables retained, lambda(Ridge Regression or Lasso Regression),Polynomial degree, tau(Regularization Parameter). **Evaluation Measures:** RMSE Root Mean Square Error, R squared. **Package Implementation:** caret
Neural Networks	**Type:** Supervised Classification and Numerical Prediction , Non-Parametric, eager learner, di, generative. **Feature Data:** Categorical and/ or Numerical. **Target Variable:** Categorical and /or Numerical.

	Noisy Data: Neural networks are not very sensitive to the noise in the data though a small data set can cause issues of overfitting the training data. **Missing Data :** Is handled as per the package implementation for instance CART provides the functionality of assigning the missing data examples to the partition which has maximum examples already assigned to it. **Correlated Data**: Typically, for a large enough data it will not be affected by correlated variables and the classification will not be affected. **Parameters and Hyperparameters:** weights(Parameters), Hyperparameters: size of hidden layers, iterations, epochs, back propagation or feedforward, learning rate, activation function. **Evaluation Measures:** Confusion Matrix ,Sensitivity, Specificity, loss function, **Package Implementation:** neuralnet, caret. Note: Hyperparameters effect how parameters effect the new learning
SVM	**Type:** Supervised Classification and Numerical Prediction , Non-Parametric, eager learner, di, generative. **Feature Data:** Categorical and/ or Numerical. **Target Variable:** Categorical and /or Numerical. **Noisy Data:** Decision rules are very sensitive to the noise in the data and even a small change in the data can throw the classification/prediction off . **Missing Data :** Is handled as per the package implementation for instance CART provides the functionality of assigning the missing data examples to the partition which has maximum examples already assigned to it. **Correlated Data**: Typically, SVM is affected by correlated data. **Parameters and Hyperparameters:** different kernels, degree of the kernel,c the regularization parameter. **Evaluation Measures:** Confusion Matrix ,Sensitivity, Specificity, Loss function, error function. **Package Implementation:** caret,e1071