**Topic:** Classification using Neural Network Approach with 10-fold Cross validation

**Methodology:**

Phase1: Data Pre-Processing:

* Normalize the required feature values

Phase2: Classification:

* Separate the data set into Training and Test data set
* Applying Neural Network Model to train the training data set
* Classify test data based on our trained model
* Applying k-fold cross validation approach to validate classification process

**Description:**

In the first phase, the process is initiated by reading the data and keeps tracking them into structured array. As the Class label of the data set is predefined and represented as string, therefore it requires Label encoding process in order to convert the target variable value into different numerical labels. Label encoding process is made by importing LabelEncoder() function of sklearn library and passing the values of target variable onto it. Therefore, the target value is represented one variable and the other feature set values are represented as another variable.

In the next phase, the dataset is split randomly into the Training set and Test set based on defined threshold value and accordingly the train and test features set and target values are defined.

Feature Scaling is used to for fitting and transforming of the train data as well as test data set.

An user-defined function named create\_model()is defied to process the classification system. Sequential() function is used to initialize the neural network and adding the input layer, first hidden layer, second hidden layer, output layer to form the network and compile it based on defined optimizer argument.

After defining the classification system, a classification model is designed based on passing the actual data set with parameter values in order to fit it with train and test data set.

The problem with machine learning models is that it is quite harder to know how well a designed model performs until its performance can be carried out based on testing on an independent data set. Cross Validation is a very useful technique for assessing the performance of designed model. In this process, 10-fold cross validation approach is applied, therefore training data set are divided into 10 number of equal subsets and for every iteration, it keeps one fold as Validation set and keep all the remaining k-1 folds in the Cross validation training set and calculating the accuracy of the model by validating the predicted results against the validation set.