Multilayer perceptron for clustering to give a higher accuracy,

the deep learning neural networks learn how to map inputs to outputs so we try to apply it, afterload data need to make some operation on it prepare it so the pre-processing is required because If a data value is higher the model will be exploded. Data preparation involves using techniques such as normalization or standardization to rescale input and output variables prior to training a neural network model that enable us to work on data. After scaling the data with min and max scaler we apply MLP with 2 hidden layers and every layer consists of 15 neurons gives an accuracy of 0.863.

try to improve the accuracy by applying one of the wrapper methods on data to give the best number of new features that improve the performance of the model. we try to use Recursive feature elimination based on importance weights. But don’t work with MLP, so we use sequential feature selection and gives us the best number of features is 9 features with an accuracy 0.871 it is higher than the old, It follows a greedy search approach by evaluating all the possible combinations of features against the evaluation criterion

We use the data come from feature selection and apply the SOM cluster with a number of the cluster from 2: 8 cluster and add the prediction as a feature and give to MLP classifier. and repeat it 10 times and take the average to find the best accuracy it give accuracy 0.866 as we see the accuracy decrease so when we try to make tuning to the model we use the data come from feature selection and make tuning on it.

We make tuning on a hidden layer and on the number of neurons on each layer,

after applying the tuning in the number of hidden layers with 20 neurons in each layer.we start from a number of the hidden layer is 2 to 8 and give us best number of hidden layer is 7 with accuracy 0.90

we get this accuracy and apply tuning on the model with number of neuron in eavry model from 10 neuron to 30 neuron and the best number of neuron in each layer is 16 with accuracy 0.912 this is the best accuracy

* The data in real life is not clean we can’t take any information without making preprocessing
* Preprocessing operation different by different of data, so cleaning it is important
* Before training the model we need to check data is balanced or not because that effect model prediction and give a good accuracy, perception, and recall
* Every model can give good accuracy is dependent on the shape of the data
* This Sequential Feature Selector is a greedy algorithm that takes so much time in big data
* MLP cluster needs more tuning to reach the best number of neurons and number of hidden layers and this is cost and time effort
* MLP gives a good result with huge data rather than small