Best and worst case for NN and Logistic Regression
Laukik Upadhye (01833608)

When to use Logistic Regression

Example:

Diabetes dataset is attached with the file:
This data set involves predicting onset of diabetes within 5 years. It has two classes as (1 for +ve and 0 for -ve). The data for each class is not balanced. It has 768 observations and 8 variables. Some values which are missing are filled with zero.
Following are the variable names:
- Number of times pregnant.
- Plasma glucose concentration 2 hours in an oral glucose tolerance test.
- Diastolic blood pressure (mm Hg).
- Triceps skinfold thickness (mm).
- 2-Hour serum insulin (mu U/ml).
- Body mass index (weight in kg/ (height in m) ^ 2).
- Diabetes pedigree function.
- Age (years).
- Class variable (0 or 1).

Neural Network output:
Mean accuracy is ~70% with the maximum number of hidden layers which is obtained by 20 runs (Fig 2.2: Size vs. Accuracy). Mean is calculated after 100 runs.

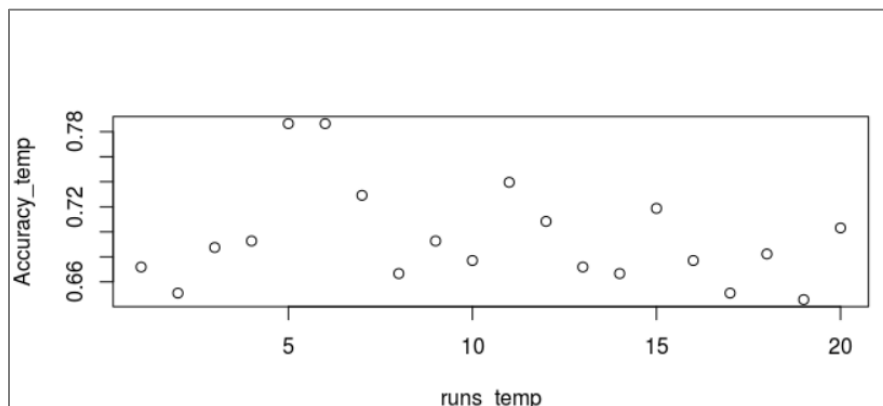*Figure 1.1: Accuracy for neural network*

Figure 2.2: Size vs Accuracy

Logistic Regression:

Mean accuracy is ~77% for logistic regression where it outperforms neural network

```
> source('/cloud/project/Logistic regression.R')
The max accuracy: 0.8333333
Mean accuracy:  0.7655729
```

Figure 3.3: Accuracy for Logistic regression

Example:

Dataset: Ionosphere dataset

It has 351 observations and 34 variables and one dependent variable (with 2 classes).

For testing and training I am splitting data in 1:4 ration (25% testing and 75% training). With these specifications it is giving **average accuracy as ~88%** and **maximum accuracy: ~96%** for neural network.

Neural Network:

Mean accuracy is ~88% with the maximum number of hidden layers which is obtained by 10 runs (Fig 5.2: Size vs. Accuracy). Mean is calculated after 100 runs.

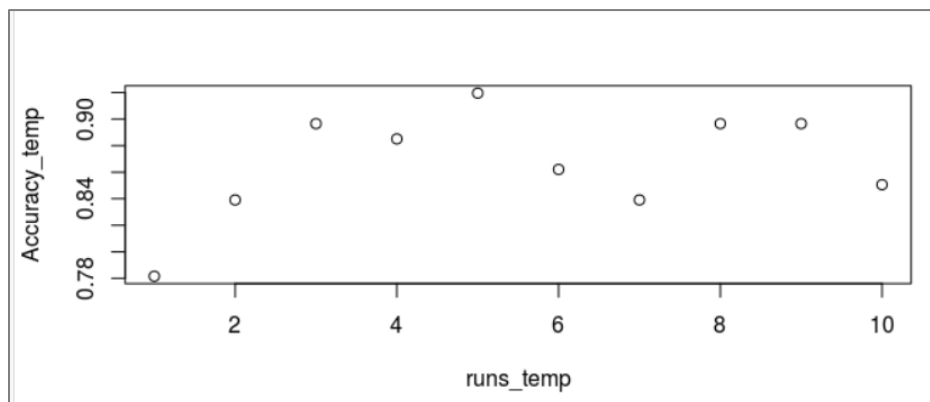Figure 4.1: Accuracy for Neural Network

Figure 5.2: Size vs Accuracy

Logistic Regression:

Mean accuracy is ~86%, for logistic regression, which show it has less accuracy than neural network

```
> source('/cloud/project/Logistic regression.R')
The max accuracy: 0.9310345
Mean accuracy:  0.8651567
```

This clearly shows, Neural Network performing better than Logistic Regression.