BITS F464 – Machine Learning (2019-2020 Sem II)

Assignment 2 Report

Neural network:

Experiment with different configurations of the number of hidden layers, the number of hidden neurons, activation functions, learning rates, and weight initializations (Gaussian, uniform, etc.).

- The error of the model did not change much from 3500 to 4000 epochs .
- Weight initializations: Tried random zeroes , zeroes with Gaussian distribution.
- Random and Gaussian gave better results, So used Gaussian
- Activation Function: Used Sigmoid and Relu.Sigmoid had better accuracy
- No. of hidden layers:2.
- No. of hidden neurons:5
- Learning rates: best results came from 0.1 (Had lower accuracies for both 0.3 (>0.1) and 0.01(<0.1))

Final Values: Epoch=5000, Learning rate = 0.1, 2 layers, Sigmoid Activation

```
Scores: [89.72602739726028, 89.72602739726028, 90.06849315068493, 89.38356164383562, 89.38356164383562]
Mean Accuracy: 89.658%
```
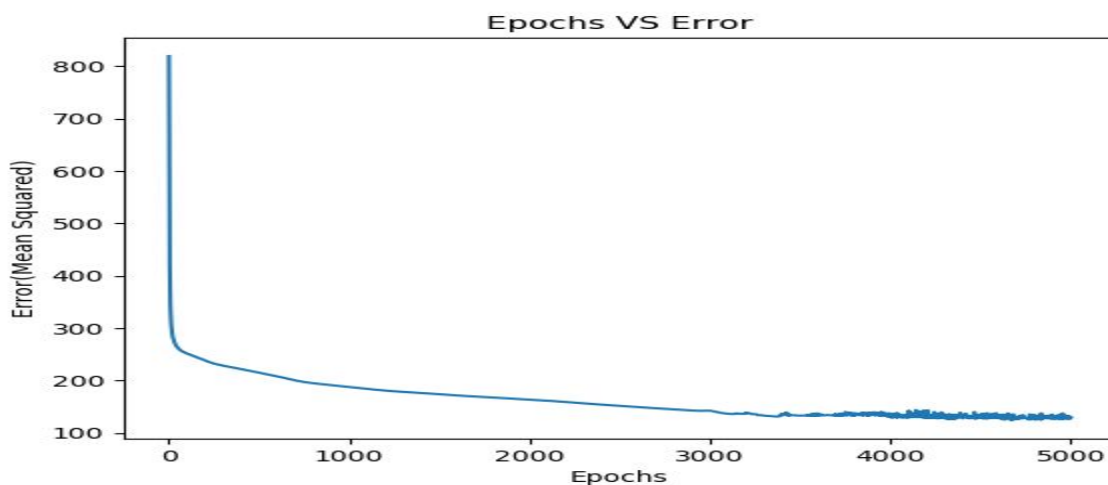
Use of Relu:

```
Scores: [46.917808219178085, 53.082191780821915, 52.3972602739726, 89.04109589041096, 46.57534246575342]
Mean Accuracy: 57.603%
```

Plot a graph illustrating how the loss function's value decreases as training proceeds.

Use an 80-20 train-test split to train/evaluate your model, and report the accuracy and F-score for the test set. (Results of 5-cross-fold validation)

1. Accuracies: [89.38356164383562, 89.72602739726028, 91.78082191780823, 88.6986301369863, 89.72602739726028]
2. F-Scores: [0.8896797153024911, 0.9013157894736842, 0.9215686274509803, 0.8925081433224755, 0.8863636363636364]
3. Mean Accuracy: 89.863 %
4. Accuracy stddev: 1.151
5. Mean F-Score: 0.898
6. F-score stddev: 0.014

Logistic Regression:

Attribute Information:

1. variance of Wavelet Transformed image (continuous)
2. skewness of Wavelet Transformed image (continuous)
3. kurtosis of Wavelet Transformed image (continuous)
4. entropy of image (continuous)
5. class (integer) (Target variable)

Experiment with different learning rates and weight initializations (Gaussian, uniform, etc.).

Tried Gaussian and Random. The results weren't much different.

Interpretation of Model(Without Regularisation)

The interpretation of the weights in logistic regression differs from the interpretation of the weights in linear regression since the outcome in logistic regression is a probability between 0 and 1.

$$log\left(\frac{P(y=1)}{1-P(y=1)}\right) = log\left(\frac{P(y=1)}{P(y=0)}\right) = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p$$

$$\frac{odds_{x_j+1}}{odds} = exp\left(\beta_j(x_j+1) - \beta_j x_j\right) = exp\left(\beta_j\right)$$

This log ratio is also called as odds.

A change in xj by one unit increases the log odds ratio by the value of the corresponding weight(bj). The highest weight (excluding w0) of 1.843 indicates that increasing the respective feature by one unit multiplies the odds by exp(1.843) (~6.315). It means that the probability for y=1 is 6.315 as high as y=0. So changing the entropy by one unit will increase the probability of y=1 by 6 folds.

Similarly, variance (feature) has the least weights on avg of -12.00 => exp(-12.00)(~6.144 8 10^-6) that is a change in variance by one unit will not have a significant impact in changing the class.

Reference Book:

Interpretable Machine Learning Book by Christoph Molnar

It is suggested that you scale your features before training/interpreting the model so that they share the same mean, variance, and/or bounds. Mention how you scale them in your report.

Rescaling the data points by dividing them

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Use an 80-20 train-test split to train/evaluate your model, and report the accuracy and F-score for the test set. Train a model (i) without regularization, (ii) with L1 regularization, and (iii) with L2 regularization.

Without Regularisation

```
Weights:[13.298888139266248, -12.260053712970961, -9.223596787777723, -10.625494840566944, 1.8436176412629206]
Weights:[13.296882209738868, -12.144377058563574, -9.279281280882907, -10.692859682775872, 1.8278119325147388]
Weights:[13.240956981378016, -12.436190599571047, -9.086766990986387, -10.543136418443074, 1.9052011409926162]
Weights:[13.24264505532951, -12.188647840049128, -9.14599829961288, -10.526692534148399, 1.8064701241709933]
Weights:[13.278420842193393, -12.393154453147464, -8.994366373142674, -10.569675540547795, 1.7726244228897379]
Accuracies: [97.44525547445255, 98.90510948905109, 97.08029197080292, 97.8102189781022, 97.8102189781022]
F-Scores: [0.9737827715355806, 0.9879518072289156, 0.968, 0.9747899159663866, 0.9743589743589743]
Mean Accuracy: 97.810
Accuracy stddev: 0.683
Mean F-Score: 0.976
F-score stddev: 0.007
```

With L1 Regularisation (lamda=0.0001)

```
Weights:[20.41022356355661, -19.774406763519288, -15.19930982988379, -18.551978791618286, 5.414101277485589]
Weights:[20.404327788931653, -19.567587353883876, -15.212573673270267, -18.731109008549993, 5.299819550896212]
Weights:[20.312194847627637, -20.02876240108717, -15.000895325354966, -18.541983178663596, 5.536668826576175]
Weights:[20.38185208155101, -19.615527507835907, -15.055454200132544, -18.545106813405983, 5.251339958868714]
Weights:[20.41410707462287, -19.95628019083401, -14.89190246713324, -18.513056071237774, 5.2583948087231]
Accuracies: [96.71532846715328, 98.17518248175182, 96.71532846715328, 97.8102189781022, 97.8102189781022]
F-Scores: [0.9660377358490565, 0.979757085020243, 0.9635627530364373, 0.9747899159663866, 0.9743589743589743]
Mean Accuracy: 97.445
Accuracy stddev: 0.683
Mean F-Score: 0.972
F-score stddev: 0.007
```

With L2 regularisation (lamda=0.0000001)

```
Weights:[23.931925978037007, -22.18549925368713, -16.478773544803577, -19.95566273411982, 3.349683019727485]
Weights:[23.897299732303985, -21.928769931793415, -16.53295991192256, -20.11876507311197, 3.2665530023224014]
Weights:[23.79372705777293, -22.610832802283653, -16.12151337830241, -19.911032028944238, 3.509549052621545]
Weights:[23.83412443981536, -22.06788425185178, -16.25053078279412, -19.81787666184026, 3.186243283030901]
Weights:[23.90845591511604, -22.608031874073426, -16.02326703183446, -19.853864760876778, 3.2561108914062444]
Accuracies: [97.44525547445255, 98.90510948905109, 97.08029197080292, 97.8102189781022, 97.8102189781022]
F-Scores: [0.9737827715355806, 0.9879518072289156, 0.9677419354838709, 0.9747899159663866, 0.9743589743589743]
Mean Accuracy: 97.810
Accuracy stddev: 0.683
Mean F-Score: 0.976
F-score stddev: 0.007
```