Machine Learning
(BITS F464)
Assignment – 2

Shivaank Agarwal	2016B4A70675H
Parth Anand	2016B4A70873H
Garvit Arora	2016B3A70462H

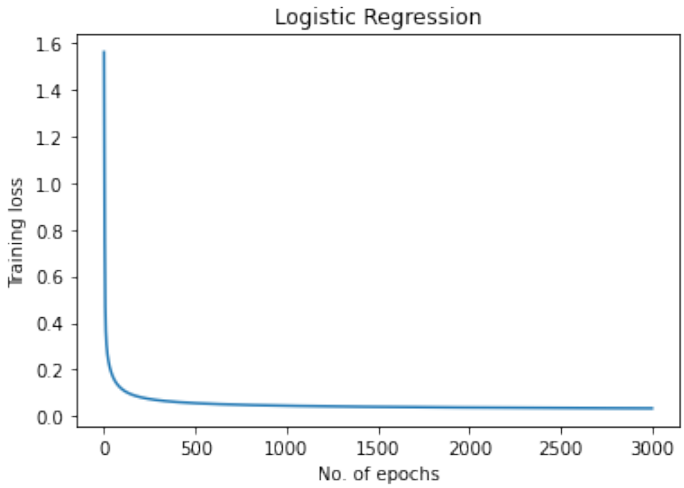
Logistic Regression

We will first start with the normalization of data to bring all the features to the same scale. We also split the data into training (80%) and test data (20%).

A) Without Regularization

We will follow the Gradient Descent method to minimize the loss function of the logistic regression.
We take the initial random 'w' values from a gaussian distribution.
No of iterations: 3000
Alpha: 0.5

After 3000 iterations, we get the following loss graph:



We see that the loss function reaches an almost constant value post 3000 iterations.

The final values of coefficients:

```
[-6.82723704 -7.07207098 -6.60751005  0.24366309 -2.57874144]
```

Using these coefficients, we calculate the accuracy and F-score on the test data set:

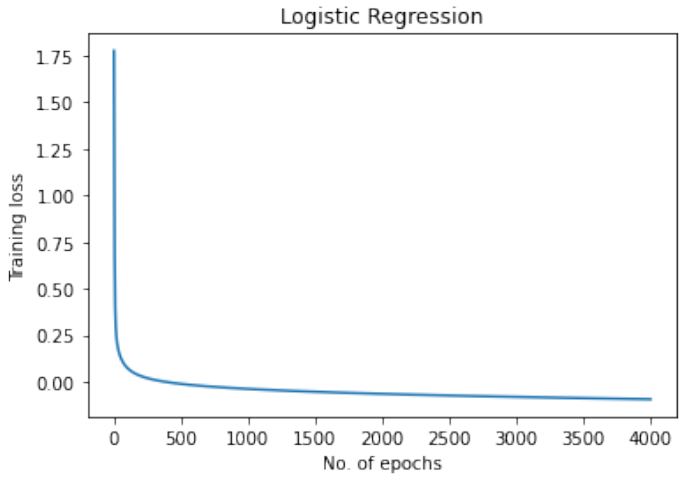
```
F-score:  0.987
Accuracy:  0.989
```

We can conclude that the trained model fits really good on the test set.

B) With L-1 Regularization

Following the same gradient descent method with the new loss function of L1 regularization, we calculate the coefficients.
No. of iterations: 4000
Alpha = 0.5
Lambda = 0.01

We get the loss function graph as:



The final values of coefficients:

```
[-7.34134042 -7.68424744 -7.16435608  0.20726468 -2.90009401]
```

Using these coefficients, we calculate the accuracy and F-score on the test data set:

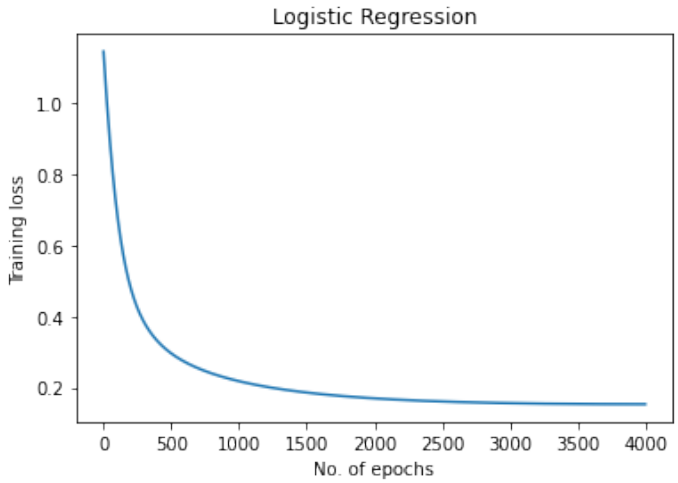
```
F-score:  0.991
Accuracy:  0.993
```

We can conclude that the trained model fits really good on the test set.

C) With L-2 Regularization

Following the same gradient descent method with the new loss function of L2 regularization, we calculate the coefficients.
No. of iterations: 4000
Alpha = 0.02
Lambda = 0.005

We get the loss function graph as:



The final values of coefficients:

```
[-3.40153378 -2.55991453 -2.46764246  0.30988807 -0.62717731]
```

Using these coefficients, we calculate the accuracy and F-score on the test data set:

```
F-score:  0.979
Accuracy:  0.982
```

We can conclude that the trained model fits really good on the test set.

Neural Networks

We start with normalizing the data to bring all the features to the same scale. We then split the data into a training set (80%) and test set (20%).

We define a maximum of 5 layers in our model (Meaning we will try to vary the hidden layers from 1 to 4 in our model).

The layers are initialized with coefficients drawn randomly from a gaussian distribution.

We then define the activation functions we would like to use for the layers ahead! We define 3 activation functions:

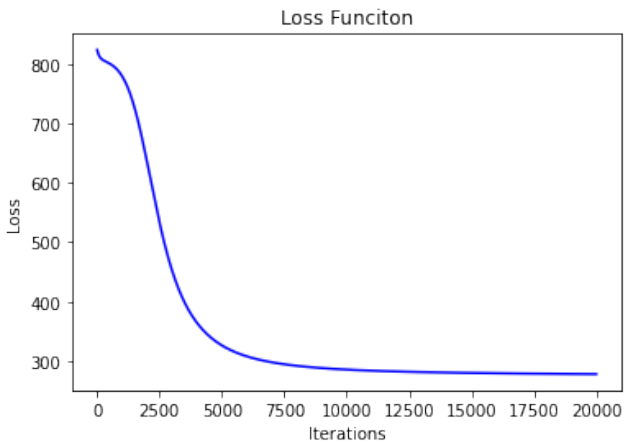
- Sigmoid
- RelU
- Tanh

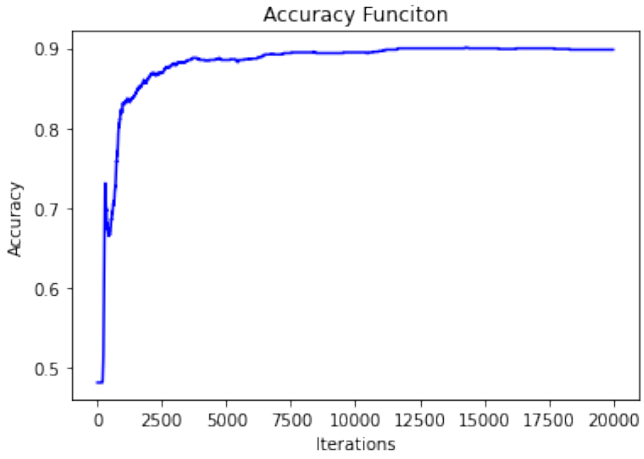
The loss function we have considered here is the cross entropy loss function for binary classification.

We tried the following models:

A) Model 1: Hidden layers: 1, Activations: Sigmoid, Sigmoid

No. of iterations: 20000
Alpha: 0.01
We get the following plots for the Loss and Accuracy function on the training data:



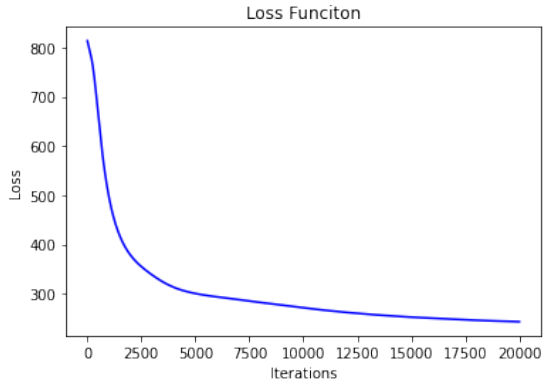


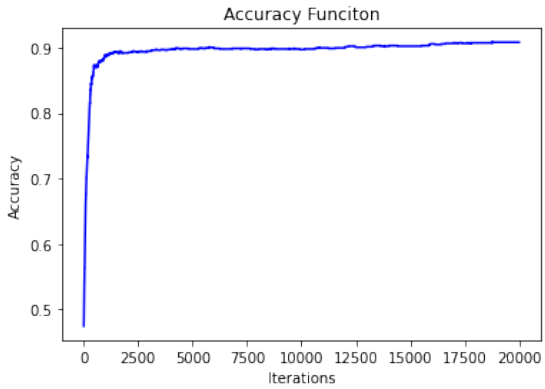
Finding the Accuracy and F – score on the Test set:

```
Test set accuracy: 0.91
Test set F score: 0.92
```

B) Model 2: Hidden layers: 1, Activations: Relu, Sigmoid

No. of iterations: 20000
Alpha: 0.01
We get the following plots for the Loss and Accuracy function on the training data:



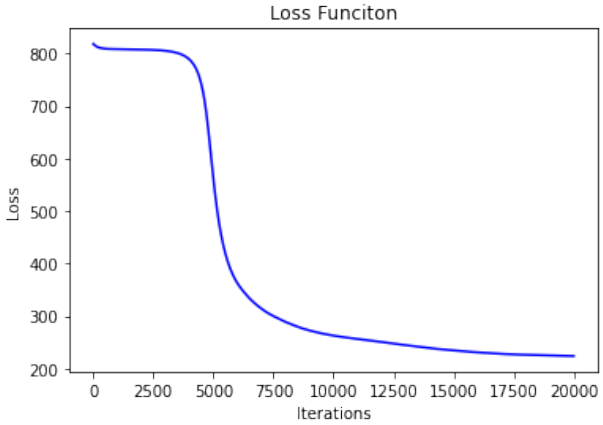


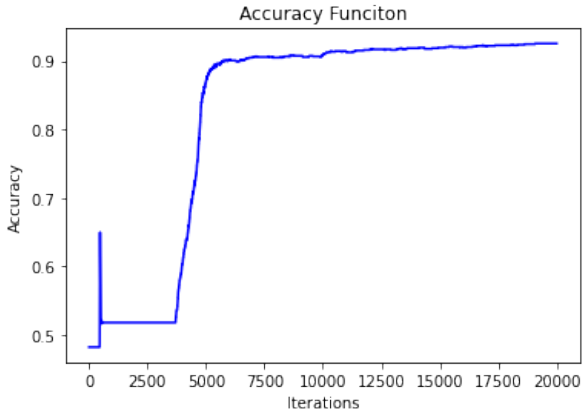
Finding the Accuracy and F – score on the Test set:

```
Test set accuracy: 0.91
Test set F score: 0.92
```

C) Model 3: Hidden layers: 3, Activations: Relu, Tanh, Relu, Sigmoid

No. of iterations: 20000
Alpha: 0.05
We get the following plots for the Loss and Accuracy function on the training data:





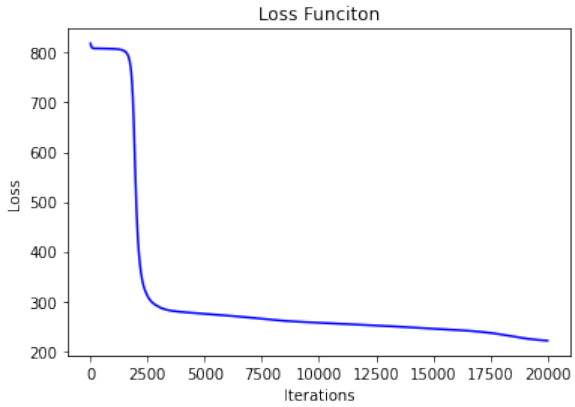
Finding the Accuracy and F – score on the Test set:

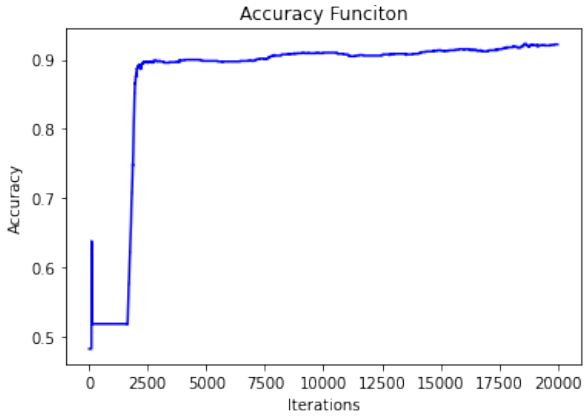
```
Test set accuracy: 0.90
Test set F score: 0.91
```

D) Model 4: Hidden layers: 3, Activations: Relu, Sigmoid, Relu, Sigmoid

No. of iterations: 20000

Alpha: 0.05
We get the following plots for the Loss and Accuracy function on the training data:



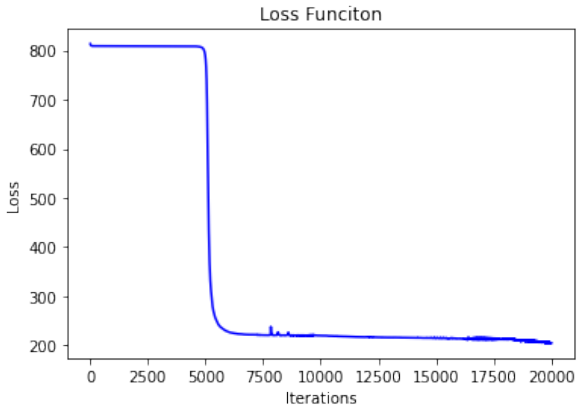


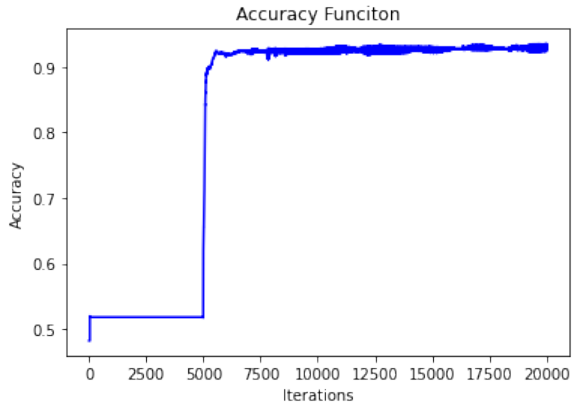
Finding the Accuracy and F – score on the Test set:

```
Test set accuracy: 0.91
Test set F score: 0.92
```

E) Model 5: Hidden layers: 4, Activations: Relu, Relu, Tanh, Relu, Sigmoid

No. of iterations: 20000
Alpha: 0.1
We get the following plots for the Loss and Accuracy function on the training data:





Finding the Accuracy and F – score on the Test set:

```
Test set accuracy: 0.91
Test set F score: 0.92
```

Overall, we consistently get an accuracy of 91%. This means that our neural network is performing really well for our test set with a F-score of around 0.92.