University of Regina

ENEL/ENSE 865: Applied Machine Learning

Professor: Dr. Abdul Bais

Instructional Designer/TA: Dr. Muhammad Hamza Asad

Submitted by:

Name: Nandish Bakulkumar Bhatt

Student ID: 200441204

Programming Assignment: 03

Introduction to Logistic Regression:

In Logistic Regression, the target variable is categorical. For example, to predict whether an email is spam or not.

.

Sigmoid (Score) Function:

$$\mathrm{P}(\mathrm{y}=+1|\mathrm{x},\hat{\mathrm{w}})=\frac{1}{1+e^{-\hat{\mathrm{w}}Th(x)}}$$

If $\mathrm{P}(\mathrm{y}=+1|\mathrm{x},\hat{\mathrm{w}})$>0.5, then y(predicted)=+1

If $\mathrm{P}(\mathrm{y}=+1|\mathrm{x},\hat{\mathrm{w}})$=0.5, then y(predicted)=+1 or -1 (unknown)

If $\mathrm{P}(\mathrm{y}=+1|\mathrm{x},\hat{\mathrm{w}})$<0.5, then y(predicted)=(-1)

Likelihood function:

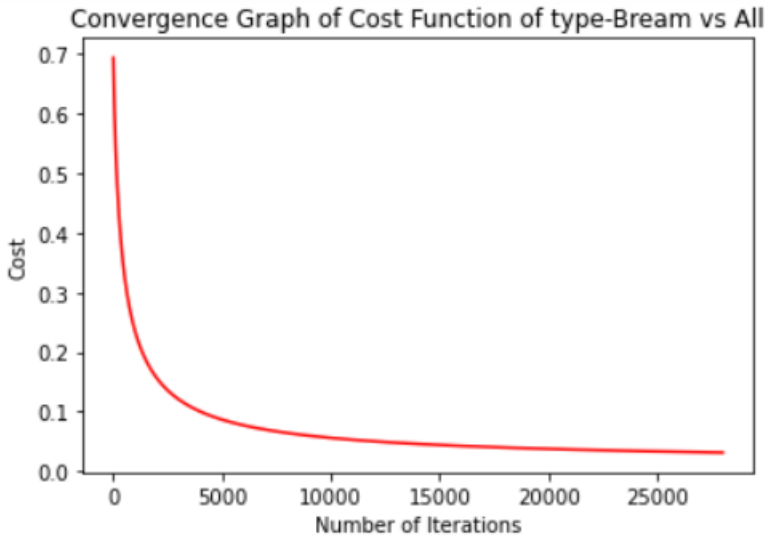
$$l(\mathbf{w})=\prod_{i=1}^{N}\mathrm{P}(y_i|\mathbf{x_i},\mathbf{w})$$

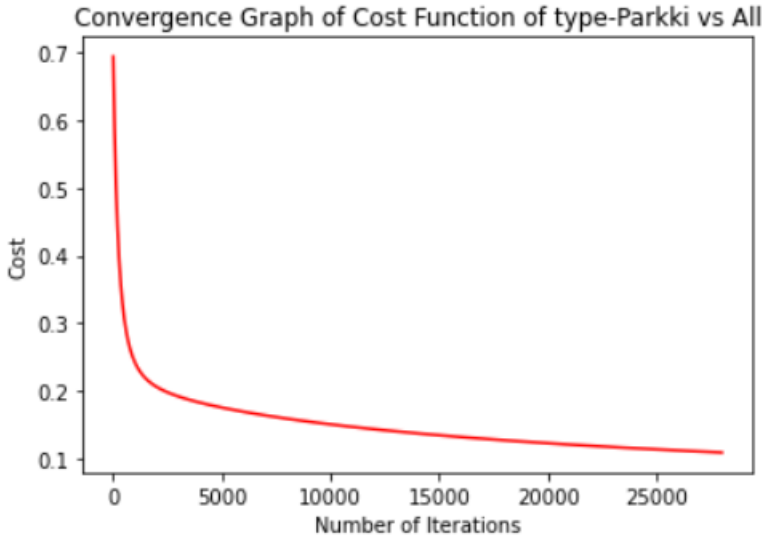
Log Likelihood function:

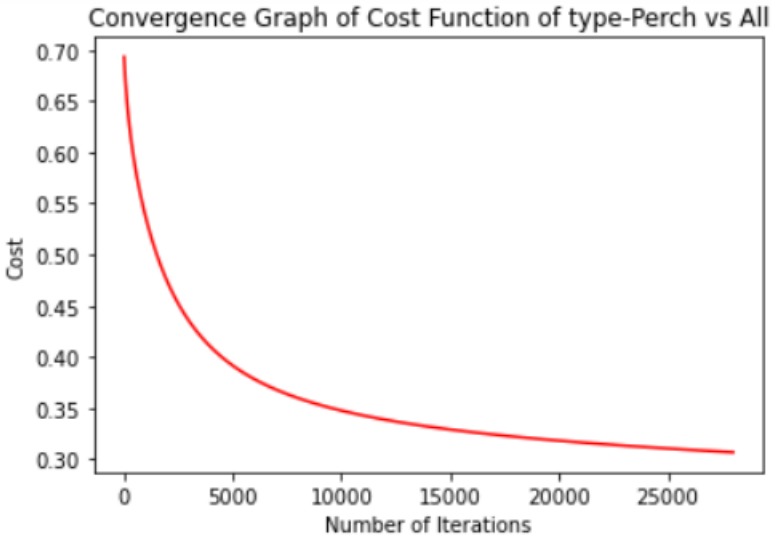
$$\frac{\partial ll(\mathbf{w})}{\partial \mathrm{w}_j}=\sum_{i=0}^{N}h_j\,(\mathbf{x}_i)\big(\mathbf{1}[y_i=+1]-\mathrm{P}(y=+1|\mathbf{x_i},\mathbf{w})\big)$$

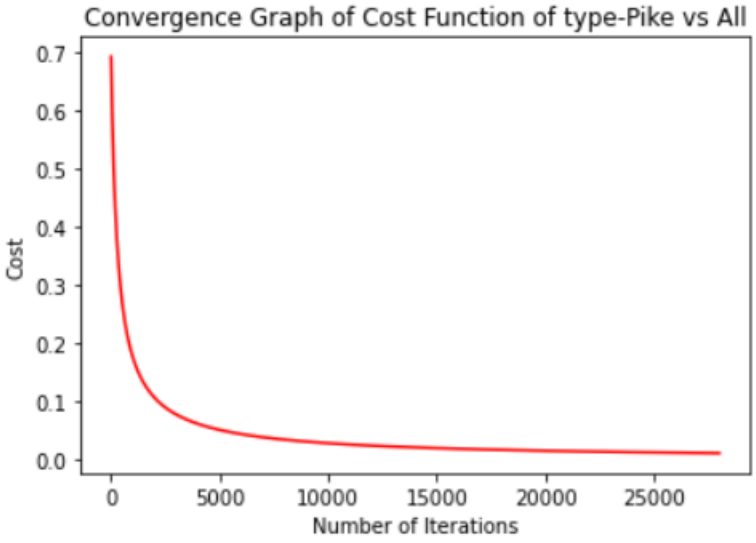
Results:

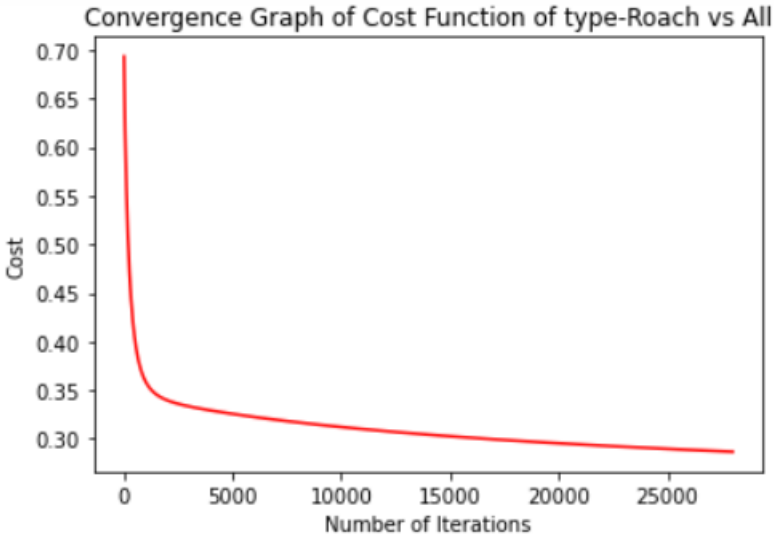
```
Training accuracy of the model is  0.84251968503937
Testing accuracy of the model is  0.65625


Fitting the given dataset..


Training accuracy of the model is  0.8267716535433071
Testing accuracy of the model is  0.75


Fitting the given dataset..


Training accuracy of the model is  0.8188976377952756
Testing accuracy of the model is  0.78125


Fitting the given dataset..


Training accuracy of the model is  0.8110236220472441
Testing accuracy of the model is  0.78125


Fitting the given dataset..


Training accuracy of the model is  0.84251968503937
Testing accuracy of the model is  0.84375


Fitting the given dataset..


Training accuracy of the model is  0.8110236220472441
Testing accuracy of the model is  0.8125
```

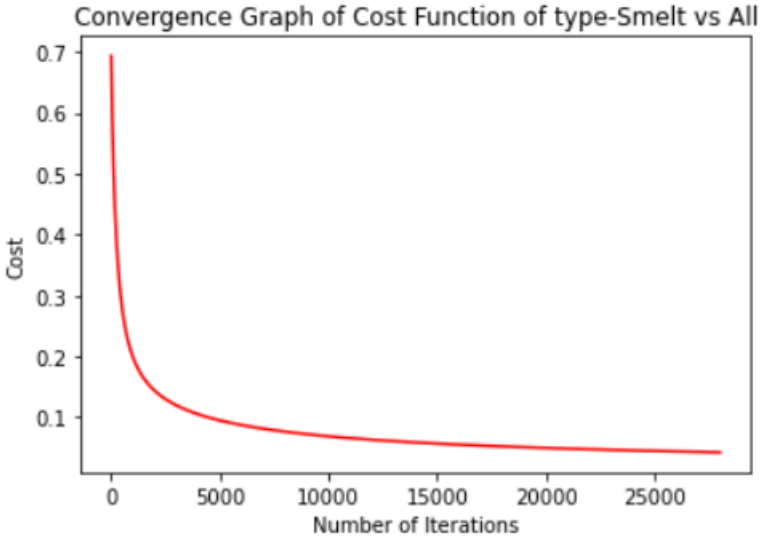

Convergence Graph of Cost Function of type-Bream vs All
Cost
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
5000
10000
15000
20000
25000
Number of Iterations

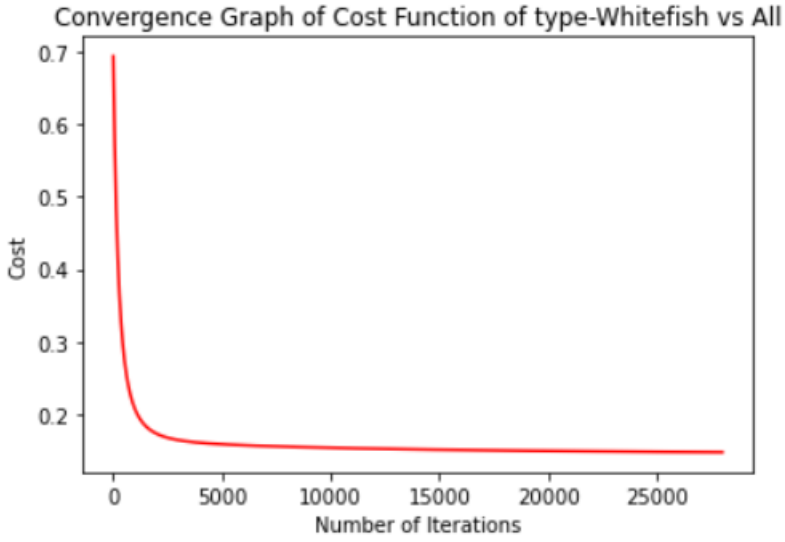

Convergence Graph of Cost Function of type-Parkki vs All
Cost
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
5000
10000
15000
20000
25000
Number of Iterations


Convergence Graph of Cost Function of type-Perch vs All
Cost
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0
5000
10000
15000
20000
25000
Number of Iterations


Convergence Graph of Cost Function of type-Pike vs All
Cost
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
5000
10000
15000
20000
25000
Number of Iterations


Convergence Graph of Cost Function of type-Roach vs All
Cost
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0
5000
10000
15000
20000
25000
Number of Iterations


Convergence Graph of Cost Function of type-Smelt vs All
Cost
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
5000
10000
15000
20000
25000
Number of Iterations



Acknowledgment:

I would like to express my gratitude to Dr. Abdul Bais and Dr. Muhammad Hamza Asad for their continuous guidance and encouragement during this course. Without their support, this work would not been possible. The notes prepared by them is really comprehensive and has a deep learning.