جامعة
الملك سعود
King Saud University

College of Computer and Information Sciences
Computer Science Department

[CSC462 Project]

[Gender Recognition by Voic]

Areej Almutairi
Sara Balatif
Noura Al-Asfoor
Lama Alosaimi

Overview

The database was created to identify a voice as male or female, based upon acoustic properties of the voice and speech. The dataset consists of 3,168 recorded voice samples, collected from male and female speakers.[CITATION Kor16 \l 1033]

The dataset is fairly distributed between males and females as shown in the figure below.

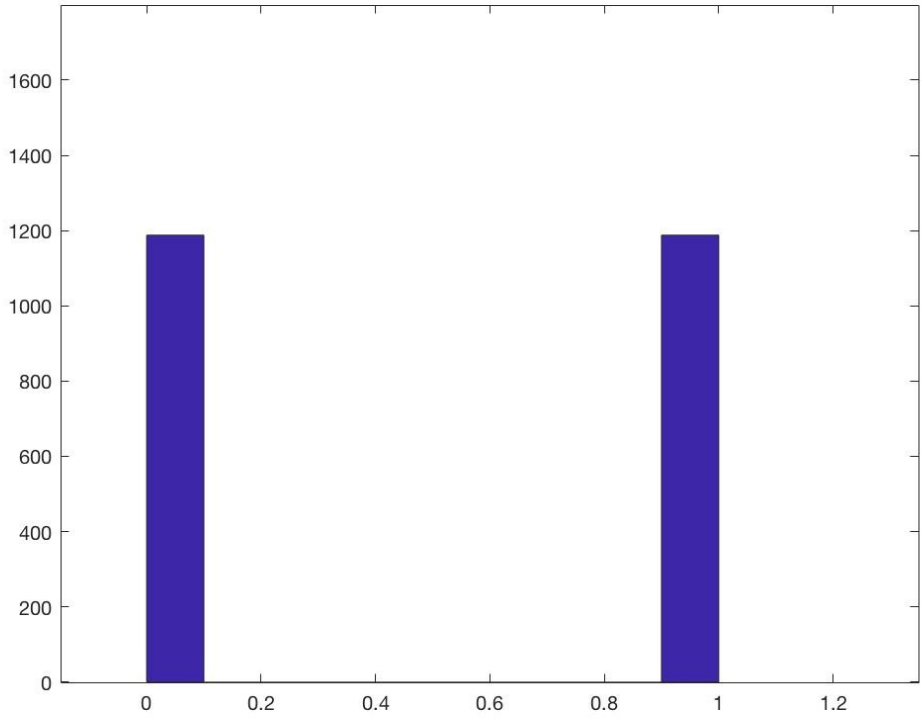

Figure 1: Distribution of Samples

Methodology

We implemented two classifiers, logistic regression and SVM. Since the dataset contains 3,168 recorded voice samples. Some of these voices are from male and the others are from female. The classifiers are obligated to recognize whether the voice is from male or female.

First we got two datasets, one for training the classifier, the other one to test it. Than we did three experiments:

First Experiment

We used logistic regression without feature normalization.

In training part we partition the dataset into 10 folds using **cvpartition** which is a pre-defined function in MATLAB. This function is responsible to divide the data set into training set (70%) and validation set (30%) with fairly distribution. So, for each value of lambda we need to call **cvpartition** for the sake of getting the training and validation error which is important to select the suitable lambda value. After getting the best lambda we needed to learn its theta's using **learnLRTheta.** The figure below shows the training and validation errors Vs. lambda values curve

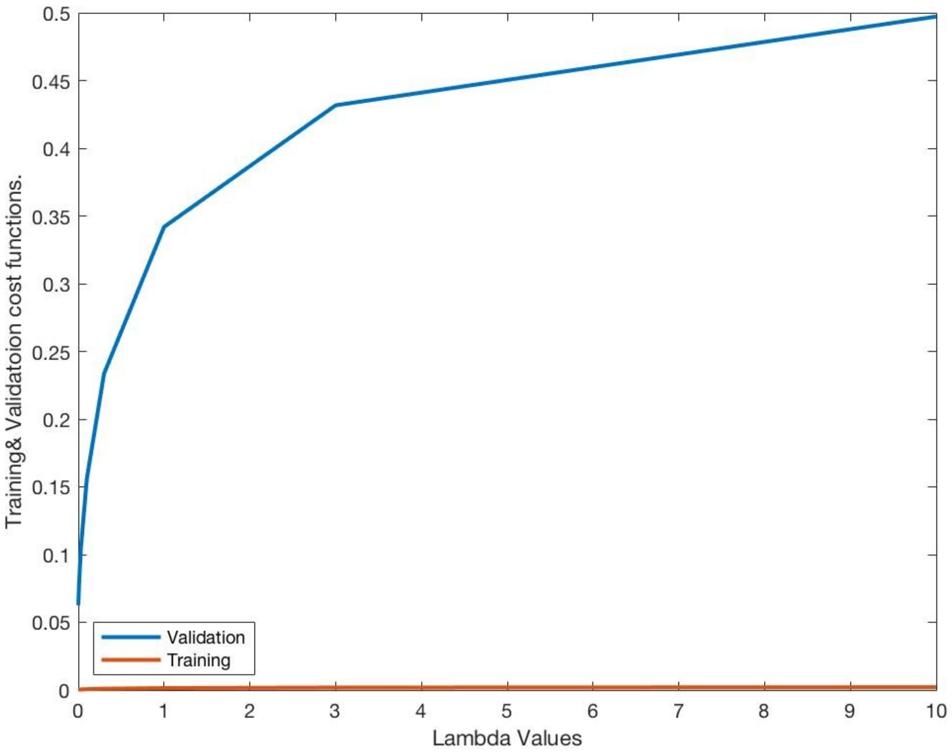

Figure 2: Training& Validation Error vs Lambda

After training we get the most suitable lambda value which is: ***0.0000*** .

The computed performance measures are shown in the table below:

Actual/Predicted:	0 (Male)	1 (Female)
0 (Male)	*376*	*20*
1 (Female)	*28*	*68*

Table 1: Confusion Matrix

Accuracy = *93.9394%*

Recall = *92.9293%*

Precision = *94.8454%*

F-Score = *93.88%*

Second Experiment

We used logistic regression with feature normalization. Before the training part we normalized the feature using **NormalizeFeature** function, to get the feature in similar range in order to fast up the gradient descent convergence. After that we did the same exact thing we did in the **First Experiment** by calculating the errors and lambda then plotting the results to help us make the decision whether the classifier is suffering from overfit or the opposite. The figure below shows the training and validation errors Vs. lambda values curve.

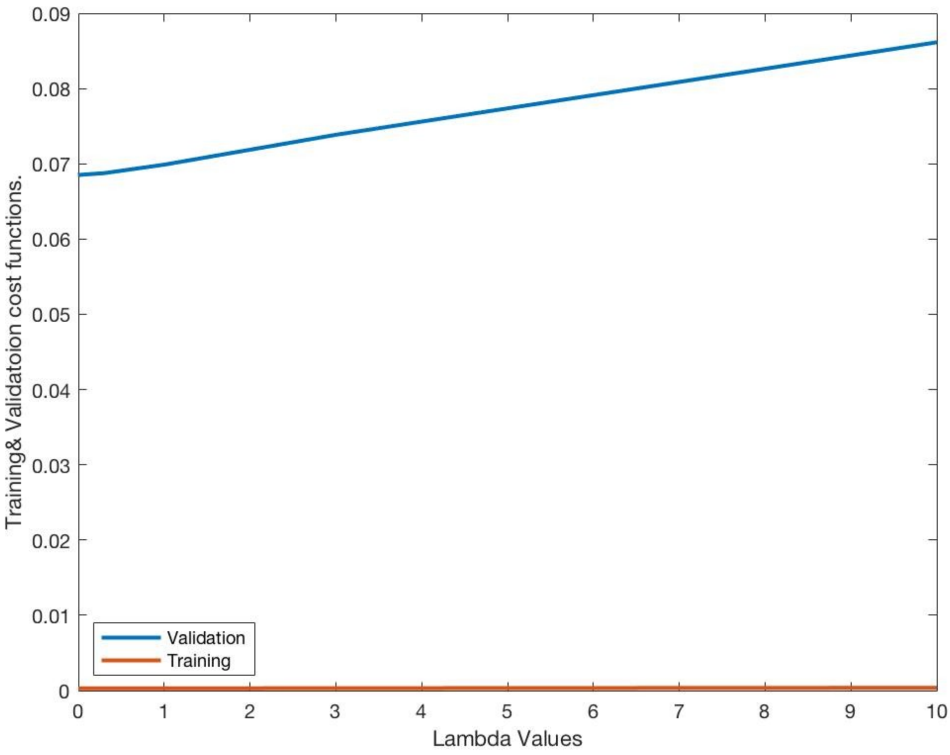

Figure 3: Normalized Training and Validation Error vs Lambda

After training we got the most suitable lambda value which is: ***0.0010*** .

The computed performance measures are shown in the table below.

Actual/Predicted:	0 (Male)	1 (Female)
0 (Male)	*371*	*25*
1 (Female)	*26*	*370*

Table 2: Normalized Confusion Matrix

Accuracy = *93.5606%*

Recall = *93.4343%*

Precision = *93.6709%*

F-Score = *93.55%*

Third Experiment

We used SVM without feature normalization.

First we used **fitcsvm** function with Train dataset to construct the model. Second we used **predict** function with Test dataset to predict new examples labels.

The computed performance measures are shown in the table below:

Actual/Predicted:	0 (Male)	1 (Female)
0 (Male)	*376*	*20*
1 (Female)	*130*	*266*

Table 3: SVM Confusion Matrix

Accuracy = *81.06%*

Recall = *67.17%*

Precision = *93.01%*

F-Score = *78.01%*

Conclusion

We concluded that Logistic Regression without normalization is the best classifier among the other classifiers as a result of its accuracy. The purpose of choosing the accuracy unit of measurement is for the fact that the data points are fairly distributed between the two classes.

Reference

[1] K. Becker, 2016. [Online]. Available: https://www.kaggle.com/primaryobjects/voicegender/home.