UNIVERSITY OF VICTORIA

Department of Electrical and Computer Engineering

ECE 403/503 Optimization for Machine Learning

LABORATORY REPORT

Experiment No: 02

Title: Multi-Category Classification Using Binary Linear Classifiers

Date of Experiment: 11 June, 2019

Report Submitted on: 18 June, 2019

To: Mr. J. Zhan

Laboratory Group No.: B01 T 12

Name(s): Alvi Mahadi (V00912845)

1 Introduction and Objectives

In this experiment, we investigate a technique for multi-category classification based on binary classifications. The technique is then applied to Fisher's 3-class datasets of Iris plants to demonstrate its effectiveness. The dataset of Iris plants to be used in this experiment was created and published in 1936 by R. A. Fisher [1]. Fisher's paper is a classic in the field and is referenced frequently to this day, as a matter of fact the dataset is arguably the best-known in the pattern recognition literature [2]. The dataset includes features of 150 Iris plants of 3 species known as Setosa, Versicolor, and Virginica, where each sample Iris is represented by a 4-dimensional vector in terms of lengths and widths of the sepal and petal of the flower.

2 Implementation Steps and Results

2.1 Implementation Steps

We strictly followed the implementation steps stated in the laboratory manual [3].

2.2 Code

```
clc;
clear all;
close all;
D = load('/home/alvi/Documents/courses/ece503/labs/2/data/D_iris.mat');
D = D.D_iris(1:4,:);
X1 = D(:,1:50);
X2 = D(:,51:100);
X3 = D(:,101:150);
rand('state',111)
r1 = randperm(50);
Xtr1 = X1(:,r1(1:40));
Xte1 = X1(:,r1(41:50));
rand('state',112)
r2 = randperm(50);
Xtr2 = X2(:,r2(1:40));
Xte2 = X2(:,r2(41:50));
rand('state',113)
r3 = randperm(50);
Xtr3 = X3(:,r3(1:40));
Xte3 = X3(:,r3(41:50));
y = [ones(40, 1); -ones(80, 1)];
ones = ones(120, 1);
% 3.3 (i) calculate w1
Dtrain1 = [Xtr1'; Xtr2'; Xtr3'];
Dtrain1 = [Dtrain1 ones];
w1 = (inv(Dtrain1' * Dtrain1) * Dtrain1') * y;
% 3.3 (ii) calculate w2
Dtrain2 = [Xtr2'; Xtr1'; Xtr3'];
Dtrain2 = [Dtrain2 ones];
w2 = (inv(Dtrain2' * Dtrain2) * Dtrain2') * y;
% 3.3 (iii) calculate w3
Dtrain3 = [Xtr3'; Xtr2'; Xtr1'];
Dtrain3 = [Dtrain3 ones];
```

```
w3 = (inv(Dtrain3' * Dtrain3) * Dtrain3') * y;
mis_class = 0;
Yte1(1, 1:10) = 1;
Yte2(1, 1:10) = 2;
Yte3(1, 1:10) = 3;
Xtest = [Xte1 Xte2 Xte3];
Ytest = [Yte1 Yte2 Yte3];
Y = [1 0 0; 0 1 0; 0 0 1];
Yresult = int16.empty;
for i=1:30
f1 = (w1(1:4, :)' * Xtest(:, i) + w1(5));
f2 = (w2(1:4, :)' * Xtest(:, i) + w2(5));
f3 = (w3(1:4, :)' * Xtest(:, i) + w3(5));
F = [f1 f2 f3];
[~, I] = max(F);
Yresult = [Yresult Y(I, :)'];
if I ~= Ytest(i)
mis_class = mis_class + 1;
end
end
error_rate = mis_class / 30;
tp = sum(Yresult(1, 1:10)) + sum(Yresult(2, 11:20)) +
sum(Yresult(3, 21:30));
fp = sum(Yresult(1, 11:30)) + sum(Yresult(2, 1:10)) +
sum(Yresult(2, 21:30)) + sum(Yresult(3, 1:20));
fn = 30 - tp;
tn = 30 - fp;
confusion_matrix = [tp fp; fn tn];
fprintf("Confusion Matrix : \n");
disp(confusion_matrix);
fprintf("Error Rate: %f\n", error_rate);
```

2.3 Result

```
Confusion Matrix :
29     1
1    29

Error Rate: 0.033333
```

3 Discussion

Fishers dataset is a matrix of size 5×150 whose first 4 rows are the data of the 150 samples, and the last row contains labels of the 150 flowers. The matrix D_iris was constructed so that its first 50 columns are associated with Iris Setosa, the next 50 columns are for Iris Versicolor, and the last 50 columns are for Iris Virginica. For each data class of 50 samples, 40 samples are used for training and the remaining 10 samples are used for testing. Random selection of the samples is made for all three data classes as follows.

After going through the process stated in section 3.3 in [3], we've found our confusion matrix as

```
29     1
1    29
```

which essentially points that we have 29 out of 30 test cases as is classified True Positive, 1 out of 30 cases is classified as False Positive. The 12^{th} data which is actually "Versicolor" is classified as "Verginica". From the confusion matrix we can observe that, the Error Rate of our experiment is 0.03333 which is very accurate.

4 Conclusion

From our analysis we can see that Linear Binary Classifier works really well for multi-category classification. Calculating the optimal parameters of linear function $f(x, w, b) = w^T x + b$ that best separate class P from class N using the solution of linear equation $\hat{X}^T \hat{X} \hat{w} = \hat{X}^T y$ really increase the speed of the calculation rather than using "for" loops. Although the result was very promising, the experiment was conducted in controlled dataset thus limiting our assumption of performance with variety of other datasets. The performance of other classification techniques such as SVN is also worth exploring.

5 References

[1] R. A. Fisher, "The use of multiple measurements in taxonomic problems, Annual Eugenics, vol. 7, part II, pp. 179-188, 1936.
[2] UCI Machine Learning, http://archive.ics.uci.edu/ml, University of California Irvine, School of Information and Computer Science.
[3] LABORATORY MANUAL, ECE 403/503 - OPTIMIZATION for MACHINE LEARNING, Prepared by: Wu-Sheng Lu, Department of Electrical and Computer Engineering, University of Victoria.