Assignment-2

Dilukshi Charitha Dissanayake
Department of Computer Science
Blekinge Institute of Technology
Karlskrona,Sweden
didi18@student.bth.se

Keywords—Supervised Learning Algorithms, Performance Comparison

I. INTRODUCTION

In this assignment I have implemented three supervised classification algorithms which are, Support Vector Machines (SVM), Multilayer Perceptron (MLP), and K- Nearest Neighbors (KNN). After the implementation of algorithms, I compared computational (Training Time) and predictive performance (Accuracy & F measure) of each algorithm.

II. IMPLEMENTATION

First, I have implemented SVM, followed by MLP and KNN while returning accuracy, training time and F measure for each fold. Each algorithm was implemented with stratified K fold to achieve the same class distribution in each fold. Then I have implemented Table 12.4 [1] where I have calculated the average and the standard deviation for each algorithms' performance score. Next I have implemented Table 12.8 [1] where I have ranked my data from best performance to the worst performance and calculated the average rank for each algorithm. Lastly, I have implemented the Friedman test based on ranks for all 3 algorithms.

III. RESULTS AND INTERPRETATION

TABLE I. ACCURACY

Fold	*SVM*	*MLP*	*KNN*
1	0.7787(2)	0.9501(1)	0.7505(3)
2	0.8(2)	0.8935(1)	0.8(2)
3	0.8152(2)	0.9261(1)	0.7913(3)
4	0.8457(2)	0.9435(1)	0.8022(3)
5	0.8152(3)	0.9413(1)	0.8348(2)
6	0.8565(2)	0.9326(1)	0.8174(3)
7	0.8348(2)	0.9652(1)	0.8174(3)
8	0.8457(2)	0.9478(1)	0.8217(3)
9	0.75(2)	0.8457(1)	0.7283(3)
10	0.8022(2)	0.8413(1)	0.7217(3)
avg	0.8144	0.9187	0.7885
std	0.0332	0.0439	0.0405
avg rank	2.1	1	2.8

According to the results from the Friedman test, test statistic was 17.8974 while rejecting H0. This means that not all algorithms perform equally while comparing to the accuracies. According to the Nemenyi test critical difference (CD) is 1.047 since alpha is 0.05 and k=3(Algo 12.1 from [1]). Therefore, difference between MLP vs SVM and between MLP vs KNN exceeds the CD. However, difference between SVM vs KNN does not exceeds the CD which means SVM and KNN may perform equally by comparing accuracies.

TABLE II. FSCORE

Fold	*SVM*	*MLP*	*KNN*
1	0.7302(2)	0.9352(1)	0.6814(3)
2	0.7473(2)	0.8747(1)	0.7356(3)
3	0.7658(2)	0.9081(1)	0.7348(3)
4	0.8022(2)	0.9261(1)	0.7437(3)
5	0.7757(3)	0.9252(1)	0.7877(2)
6	0.829(2)	0.9186(1)	0.7778(3)
7	0.7841(2)	0.9553(1)	0.7572(3)
8	0.8097(2)	0.9333(1)	0.7796(3)
9	0.716(2)	0.8264(1)	0.6719(3)
10	0.7612(2)	0.7978(1)	0.6614(3)
avg	0.7721	0.9001	0.7331
std	0.0356	0.0512	0.0464
avg rank	2.1	1	2.9

Friedman test statistic was 18.2 for F measure which also means that not all algorithms preform equally while comparing the f scores. While comparing the CD (1.047) of Nemenyi test, difference between MLP vs SVM and between MLP vs KNN exceeds the CD. However, difference between SVM vs KNN does not exceeds the CD which means SVM and KNN may perform equally by comparing Fscores.

TABLE III. TRAINING TIME

Fold	*SVM*	*MLP*	*KNN*
1	1.0994(2)	1.8965(3)	0.0498(1)
2	1.0228(2)	1.6446(3)	0.0159(1)
3	1.0248(2)	1.7798(3)	0.0142(1)
4	1.0522(2)	1.1092(3)	0.0136(1)
5	1.0705(3)	0.9997(2)	0.0139(1)
6	1.1677(2)	1.3645(3)	0.0135(1)
7	1.0126(2)	1.4022(3)	0.0134(1)
8	1.0097(3)	0.9338(2)	0.0144(1)
9	0.9786(2)	2.0897(3)	0.0138(1)
10	0.9701(2)	1.8403(3)	0.0139(1)
avg	1.0408	1.506	0.0176
std	0.0594	0.4043	0.0113
avg rank	2.2	2.8	1

Friedman test statistic was 16.8 for training time which also means that not all algorithms preform equally while comparing the training time. While comparing the CD (1.047) of Nemenyi test, difference between MLP average rank vs KNN and between SVM vs KNN exceeds the CD. However, difference between SVM vs MLP does not exceeds the CD which means SVM and MLP may perform equally by comparing the training time.

REFERENCES

[1] P. Flach, Machine Learning: The Art and Science of Algorithms that Make Sense of Data. Cambridge University Press, 2012.