

Group Practical 1

Shiyu Yi,2016141231175

Chuang Du,2016141462277

Jiali Shang,2016141462137

July 18, 2018

1 Problem overview

1. Using iris data to assess the classification performance by tuning the KNN classifiers:

- Splitting the data using different percentage
- Change cross validation folds
- Changing the value of K
- Normalise the data

1) Summarise the above classification performances of the above settings using tables/figures

2) Discuss the results.

Table 1: Splitting the data using different percentage		
Training proportion	Test Proportion	Evaluate Score
0.9	0.1	0.9333333333333333
0.8	0.2	0.9333333333333333
0.7	0.3	0.9555555555555556
0.6	0.4	0.9666666666666667
0.5	0.5	0.9733333333333334
0.4	0.6	0.9666666666666667
0.3	0.7	0.9333333333333333
0.2	0.8	0.9166666666666666
0.1	0.9	0.837037037037037

2 Problem one

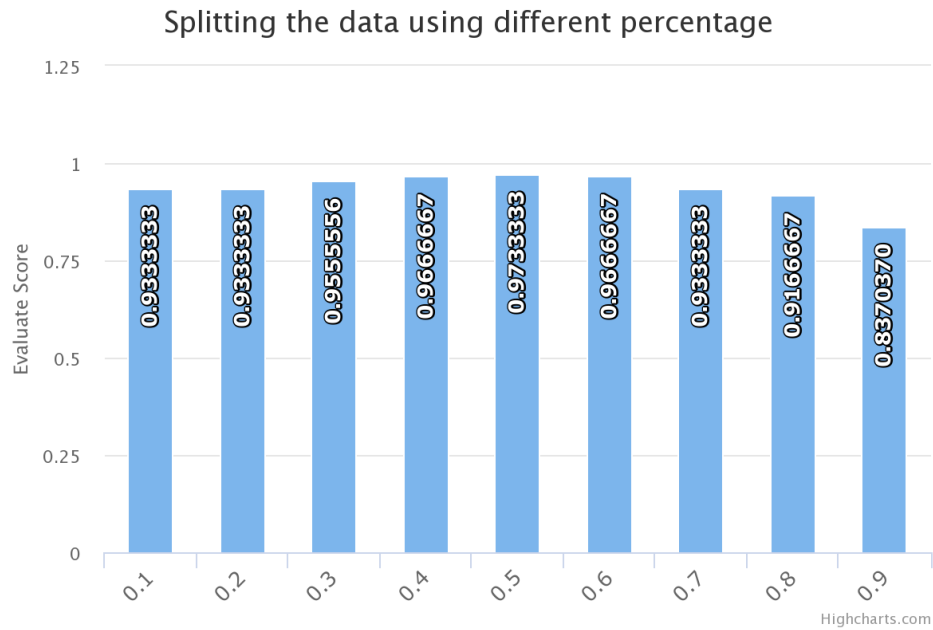


fig.1

Table 2: cross validation folds

cv's folder	Accuracy
2	0.94(+/-0.04)
3	0.99(+/-0.02)
4	0.97(+/-0.04)
5	0.97(+/-0.05)
6	0.97(+/-0.07)
7	0.97(+/-0.07)
8	0.97(+/-0.08)

3 Problem two/three

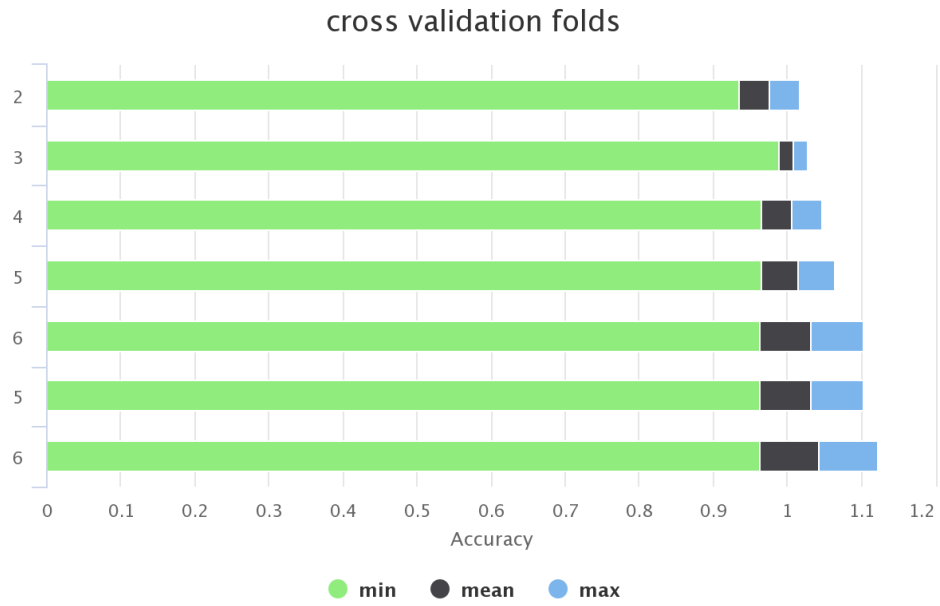


fig.2

Table 3: Normalise the data	
Method	Accuracy
Min-Max scaling	0.97(+/-0.05)
Standardization	0.97(+/-0.02)
Normalizer	0.97(+/-0.04)

4 Problem Four

5 Classification

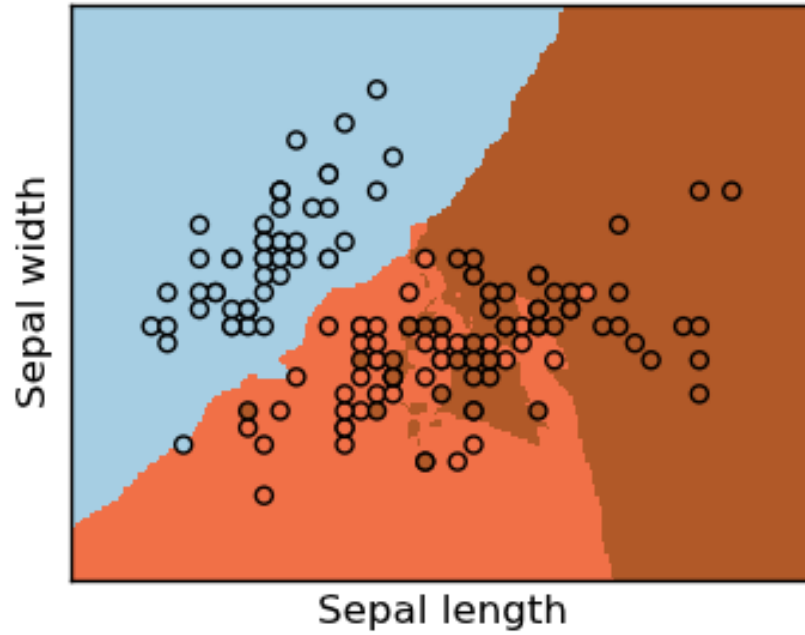


fig.2