Reykjavik University

Master Degree in Computer Science

Course of Data Mining and Machine Learning

Theoretical Answer - K-Nearest Neighbours

Author

Davide Parente

Instructor

Eyjólfur Ingi Ásgeirsson

A.A. 2022/2023

1 What explains this difference in accuracy between kNN and wkNN when k increases?

In the kNN algorithm the predictions are made by averaging across the k neighbours while in the wkNN the predictions are made between neighbors also considering their distances. When k is larger, the distance could become larger which goes against the principle behind the kNN (i.e. that neighbours that are nearer have similar classes). Contrary to kNN in the wkNN algorithm the distances play an important role so the accuracy always remains constant.

Is it possible to see, from the figure that until a certain value there is a good accuracy

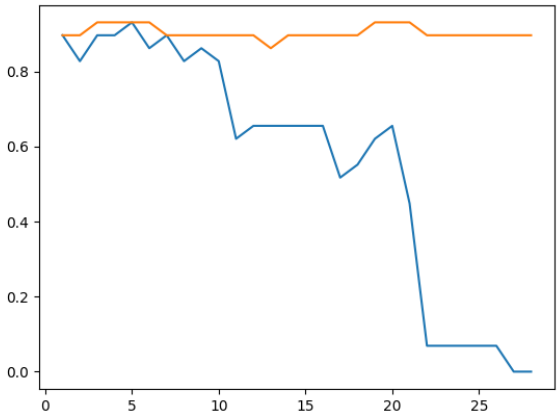

Figure 1: Compare kNN with wkNN

for kNN algorithm but when the k increases the algorithm loses its effectiveness.
In the following picture[1] I will show an example where an high value of k makes the prediction wrong, unlike wKNN which would have made the correct prediction with the same k value.

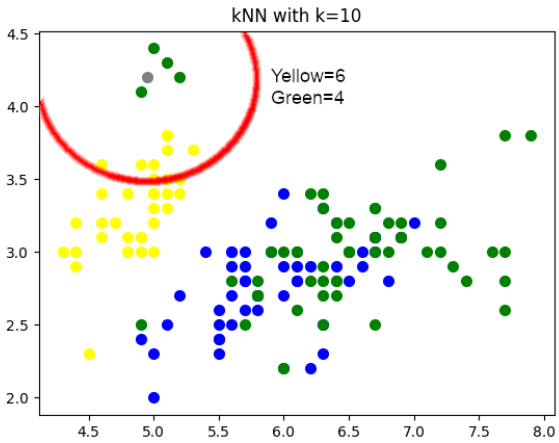

Figure 2: Example of kNN with not optimal k

[1]The dataset used has been modified just to show this case, it is not the real plot