INF283 – Exercise 4

Sindre Eik de Lange

# 1. Support Vector Machines (SVMs)

1.1

1. Code to find the dimensions of a vector **X** is already defined as *X\_shape*, which, according to *make\_blobs()* DocStrings is equal to (**n\_samples**, **n\_features**).

2. Code to find the dimensions of a vector y is already defined as *y\_shape*, which, according to *make\_blobs()* DocStrings is equal to (**n\_samples**).

3. Again, according to *make\_blobs()* DocString, vector **y** contains: 'The integer labels for cluster membership of each sample', aka. each tuples' (in **X**) label.

1.2

**m** and **b** are the values in each touple, contained in the list (of touples) iterated through in the for-loop. More specifically, the for-loop iterates through a list, containing touples, consisting of values (**m\_i**, **b\_i**), where **i** = iteration number. They represent the values **a** and **b** for linear functions, given as *y=ax+b*, where **xfit** then represents the **x**.

1.3

**d** represents the distance between each vectors' line and the nearest data point on each side.

1.4

1. Code to print out the value of support vectors: *model.support\_vectors\_* (after assigning **model** as the SVM classifier).

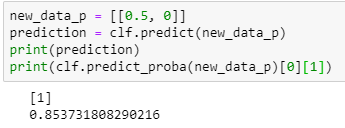
2. There are 3 support vectors.

1.5

They centered **r** in origo.

1.6

According to <https://matplotlib.org/examples/color/colormaps_reference.html>, when specifying *cmap=’autumn’*, then the colors will range from **red** to **yellow**, so I assume that it starts on **red** (left side), e.g. 0 = **red**, 1 = **yellow**.

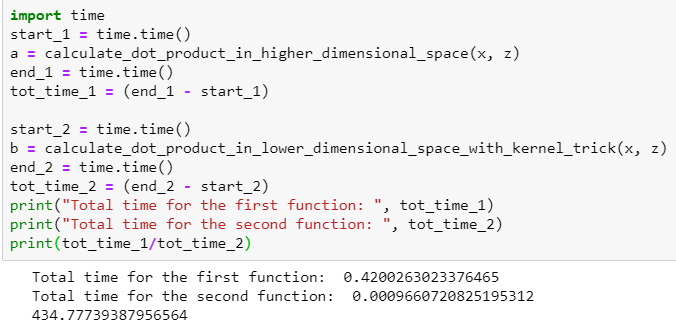


So, as one can see here, the new data point will be classified as 1 = **yellow**, and the probability of this happening is approximately 0.854.

1.7

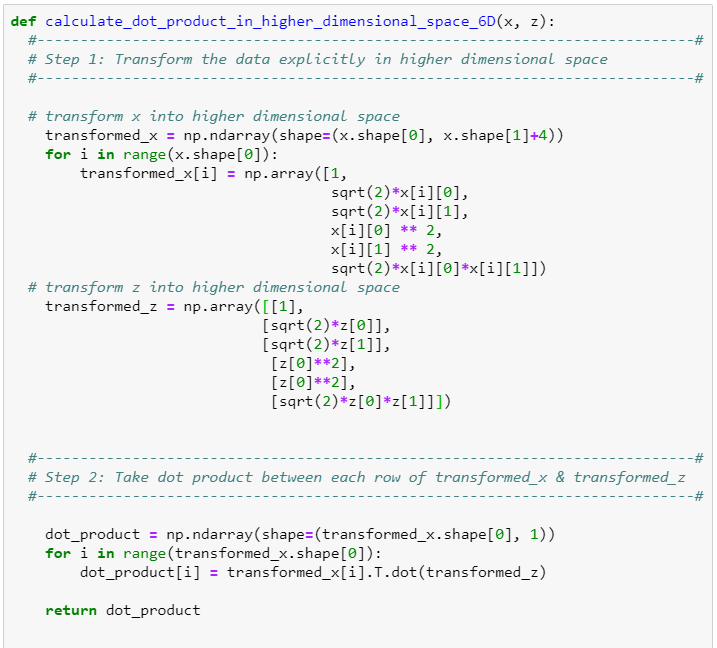
This is probably data dependent like everything else in machine learning, but I would assume one would utilize a validation set, combined with either GridSearch or RandomSearch, with or without cross validation (depending on available resources and the task domain), in order to get the optimal hyperparameter. Alternatively, one could utilize Bayesian Optimization with ‘skopt’.

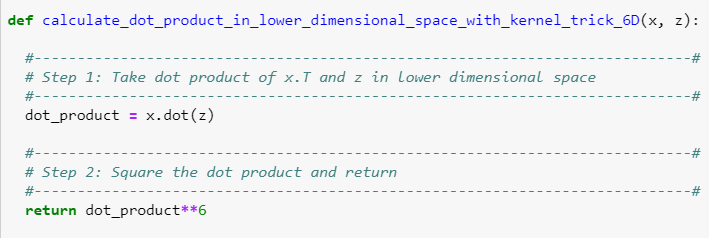
1.8

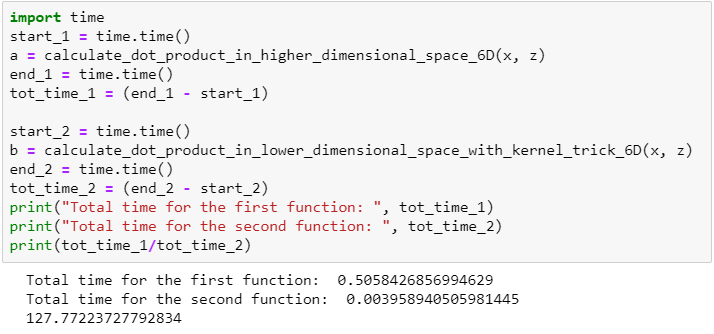


The **kernel trick** function is faster than the other one.

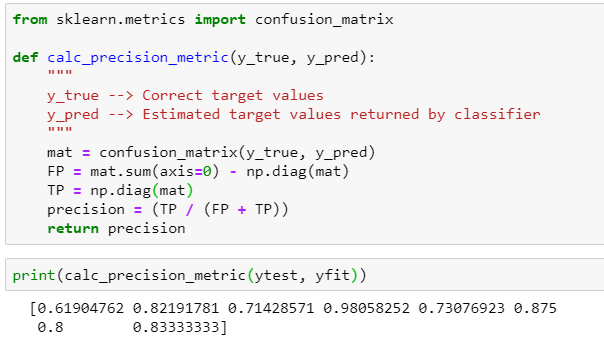
1.9







1.10

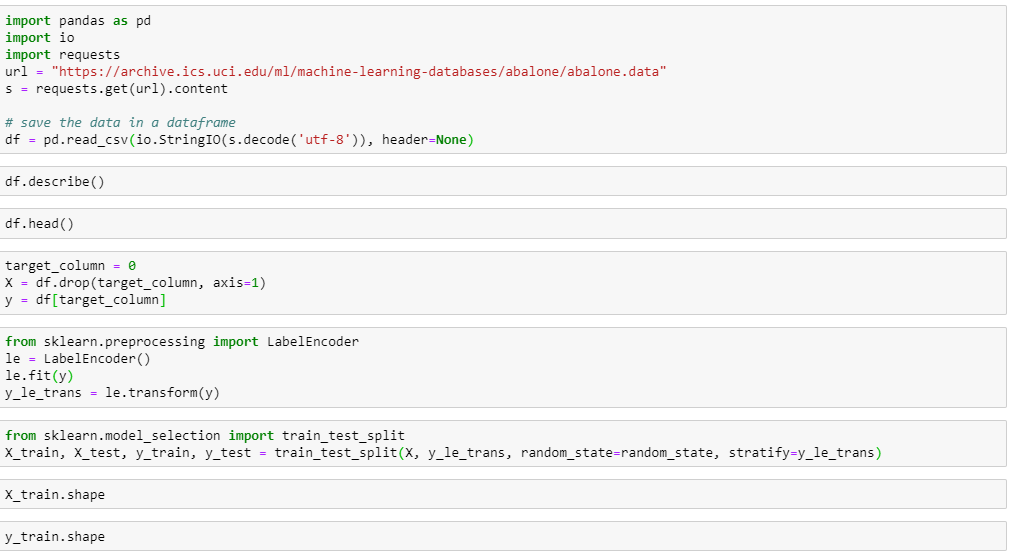


# 2. Ensemble Learning

2.1

Changing the ‘voting’ parameter from ‘hard’ to ‘soft’ slightly improves the **VotingClassifier**. This is because the ‘hard’ classifier only considers the binary input from each voter, while the ‘soft’ classifier takes into account the average of probabilities. This means that in our case, where each of the “weak” learners are fairly good (> .85), it helps to consider how certain each classifier is.

2.2







So, one can see that the **Bagging Classifier** is the best classifier. Important to note that this is based on running the models one time, and not utilizing any hyperparameter optimization tools like GridSearch(CV) or RandomSearch(CV).