Machine Learning HW

Team Member:

0516017 李柏毅

0516059 劉嘉豪

0516306 尤健羽

0516319 傅信瑀

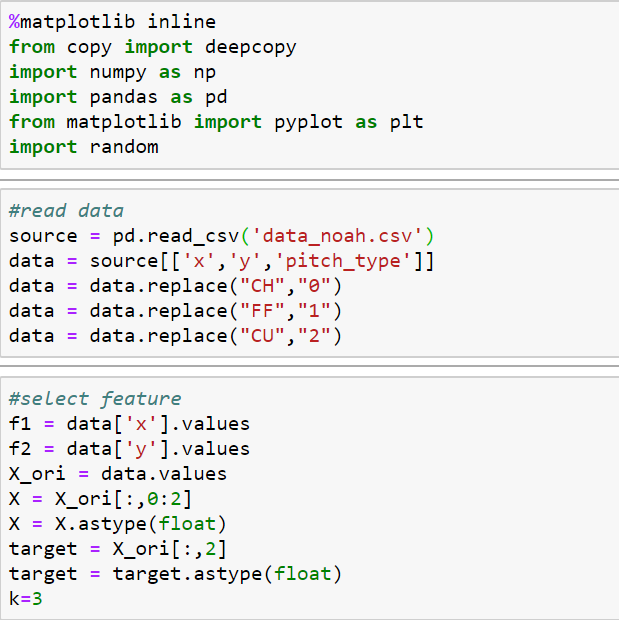
0516322 朱蝶

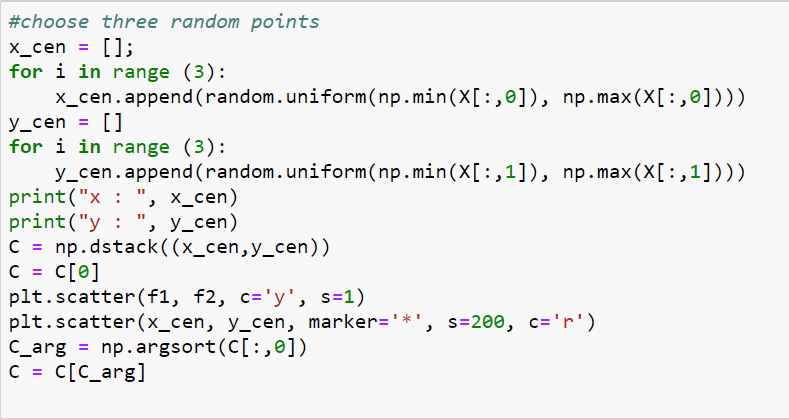
1. What environments the members are using:

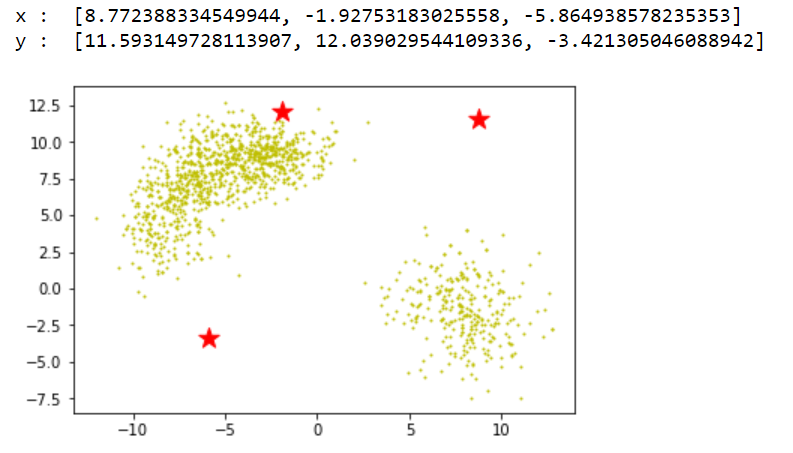
Use Jupyter as the environment

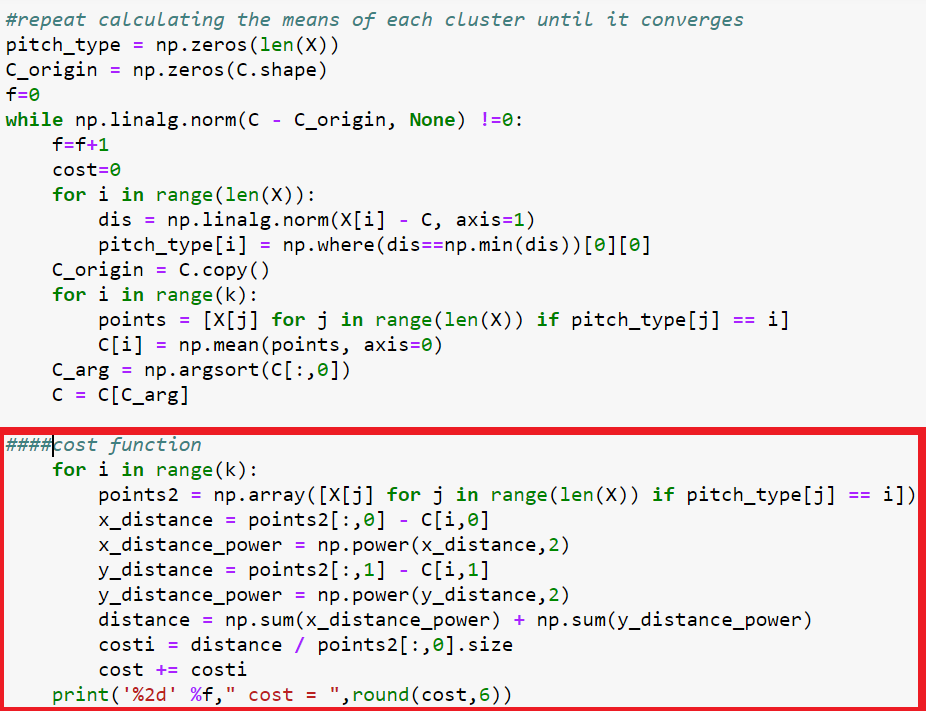
1. K-means code and result:

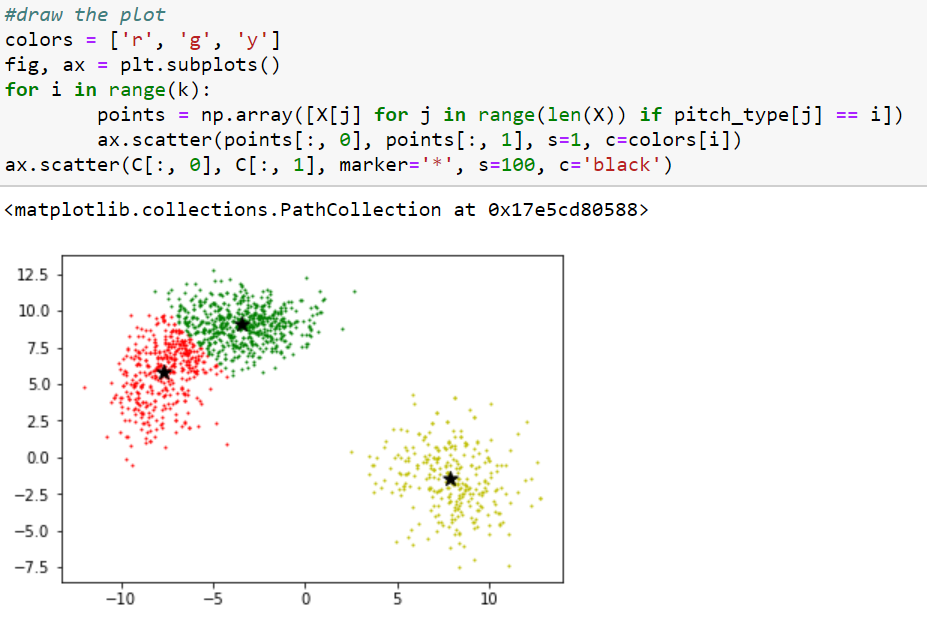
\*Feature : X , Y Target : pitch\_type



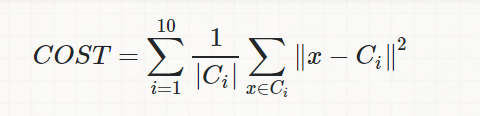




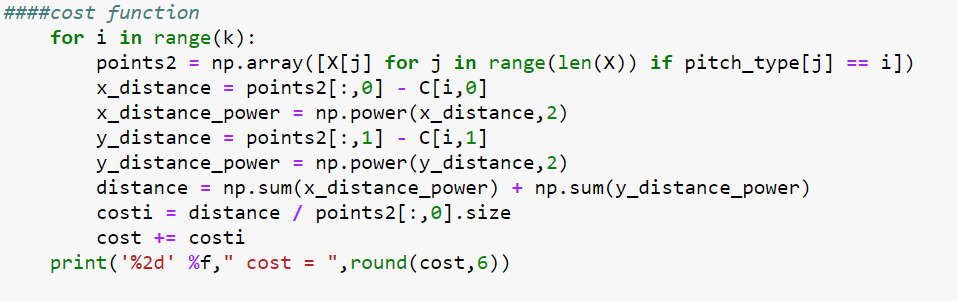


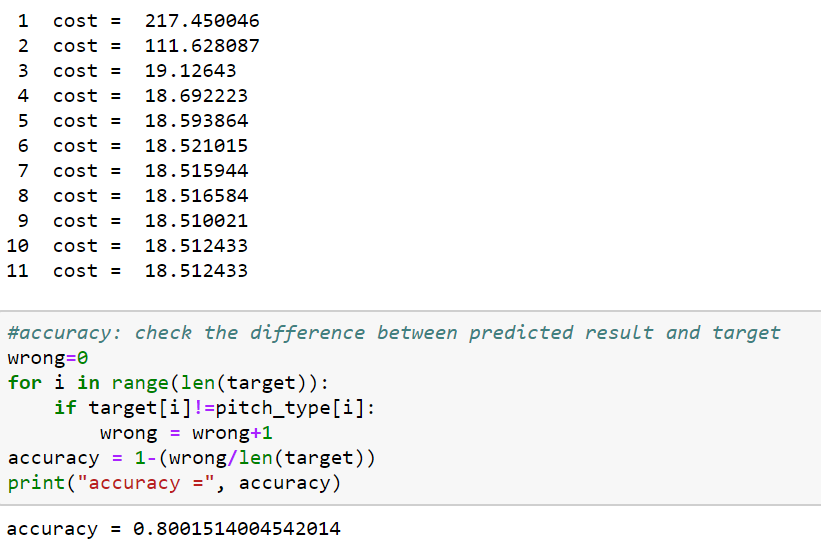


1. Cost function and accuracy:



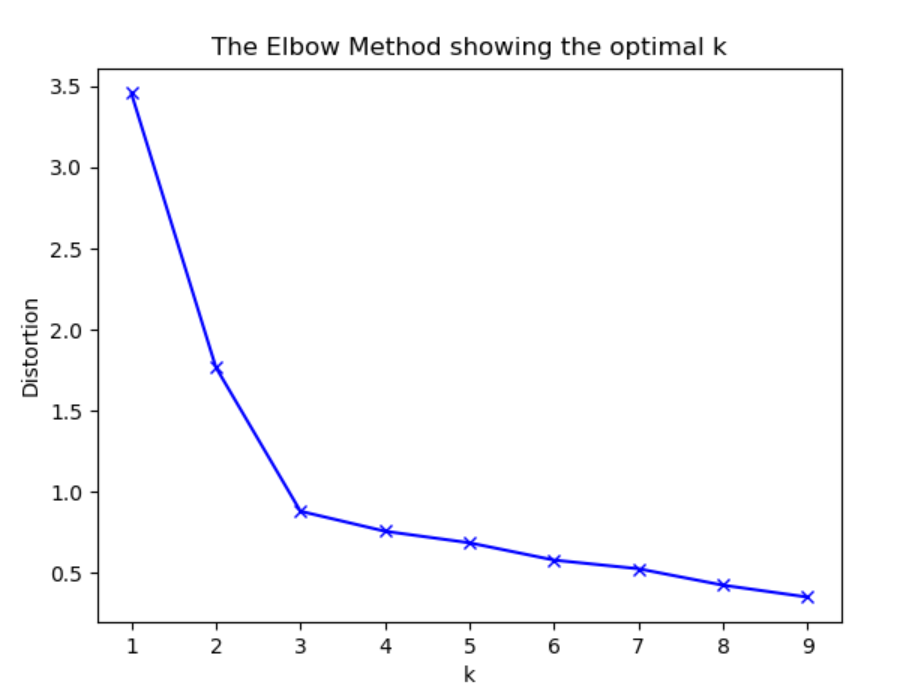
Calculate every distance between every point and their own center, and sum them which are in the same cluster, then divide the cluster’s size.





1. Why choose k=3:

The technique to determine k called the elbow method.

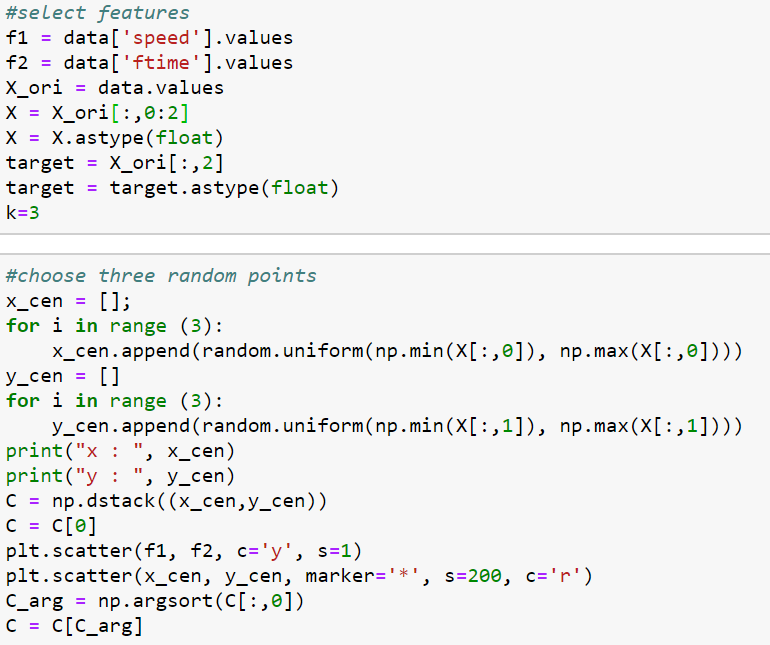


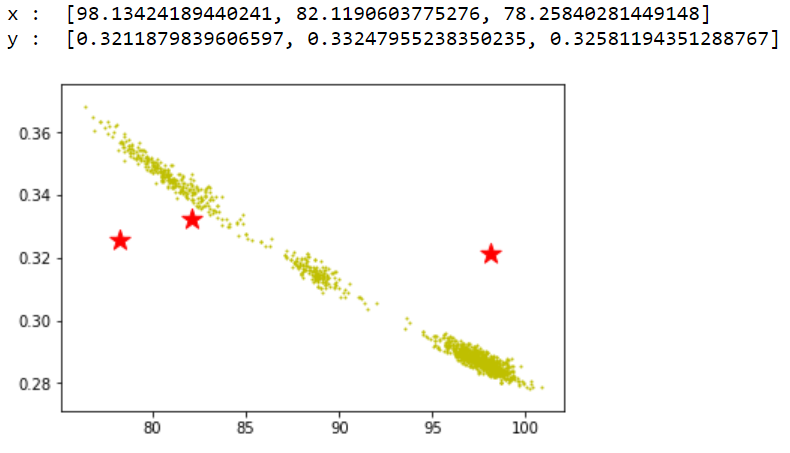
We can see that the improvement will decline, at some point rapidly, creating the elbow shape. That point is the optimal value, and it is 3.

1. Another two attributes to partition:

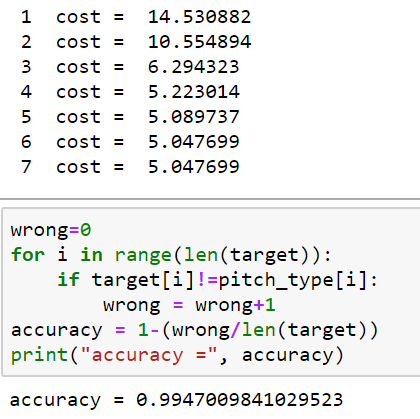
\*Feature: ftime, speed Target: pitch\_type

Only screenshot the code differ from above

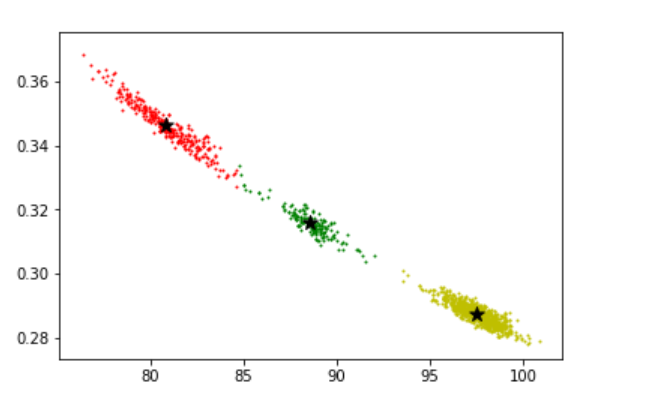




\*Cost function result and accuracy:

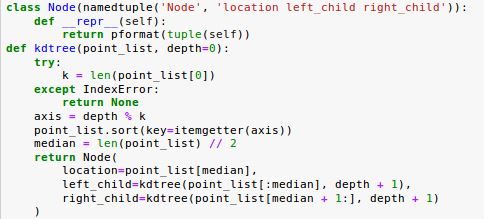


\*Result:

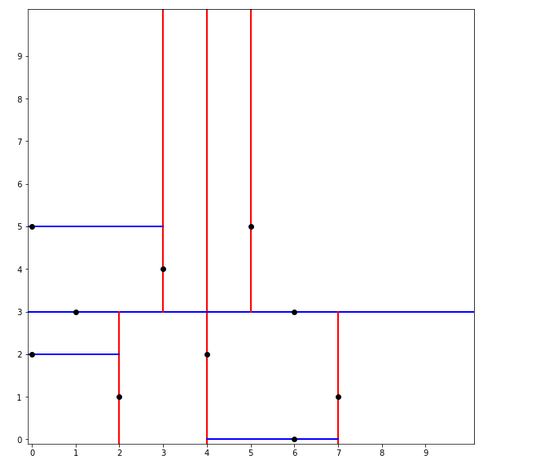


1. KD-tree code

Divide into two parts, and do it recursively, we use median to be divided point.



1. Result of KD-tree



END