4.02 k-Nearest Neighbours

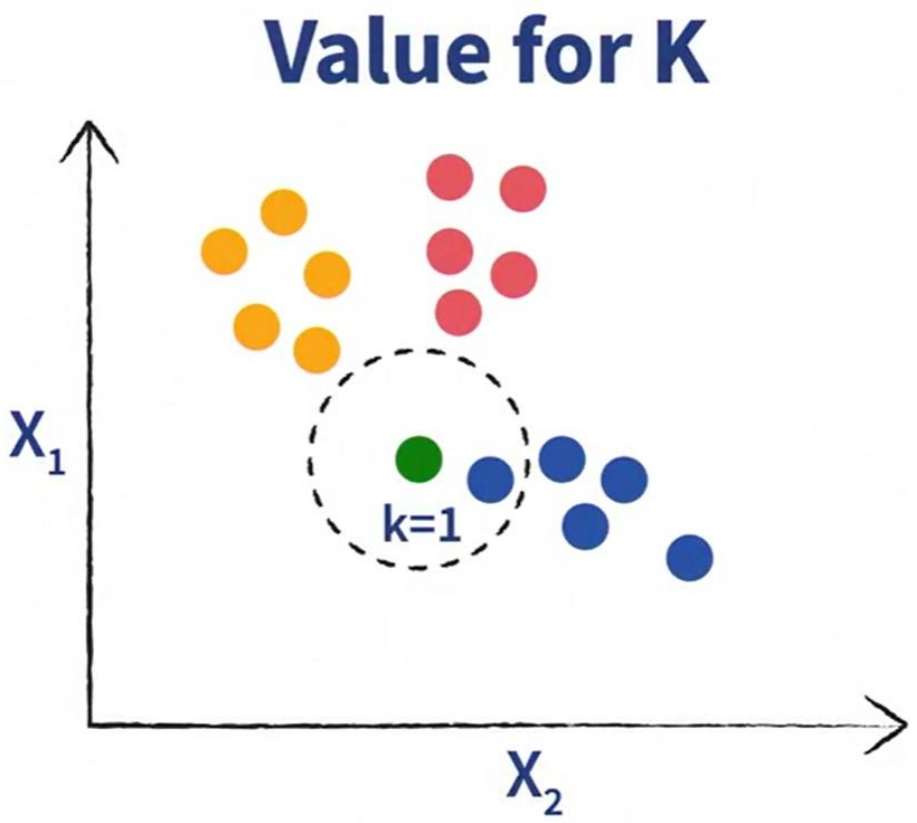
K-Nearest Neighbour Classification (KNN)

- Supervised ML classifier that memorises observations from within a test set to predict classification labels for new unlabeled observations
- KNN makes predictions based on how similar training observations are to the new incoming observations
- More similar the observation values, more likely they will be classified with the same label
- Use Cases include Stock Price Prediction & Product Recommendations

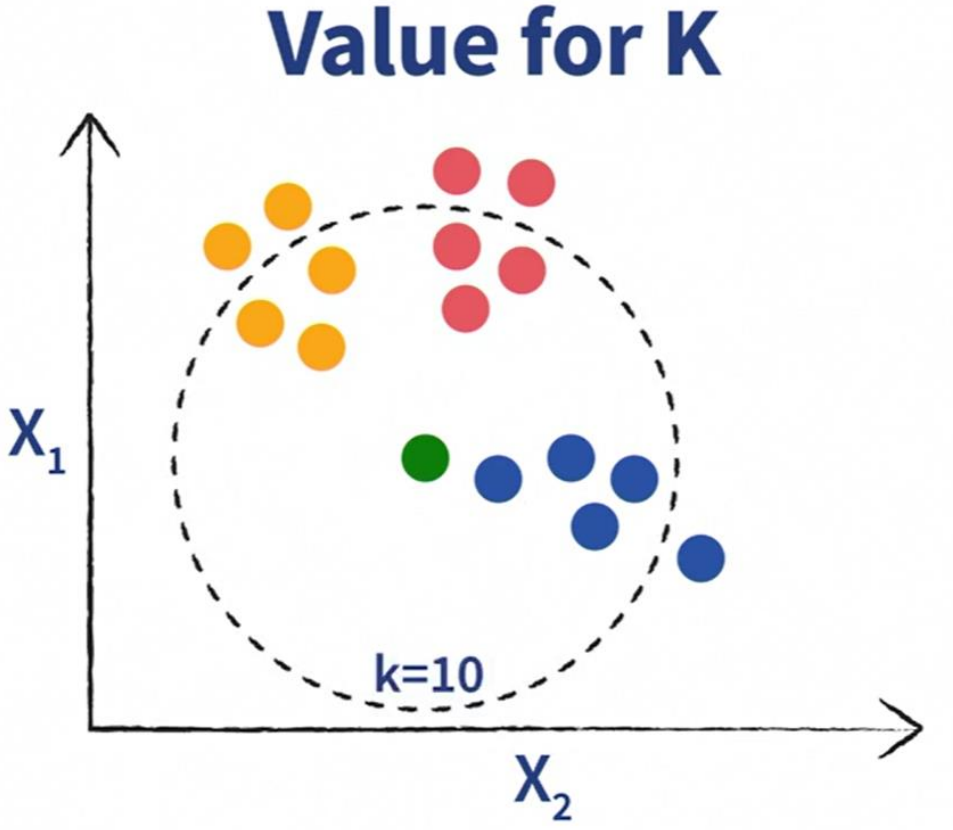
KNN Model Assumptions

- Data Set is labeled
- Data Set contains relevant features only (i.e., minimal noise)
- Data Set has distinct sub-groups
- Data is not very large (since KNN will take a long time)

KNN Model Examples



If k=1, then KNN algo will only use its **nearest neighbour** to make predictions



If k=10, then KNN algo will only use its **10 nearest neighbours** to make predictions