Chapter 7

Aly Khaled

August 14, 2021

1 What is Ensemble Learning?

- Ensemble Learning in using more than one predictor then it aggregate the prediction of each predictor and predict the class that gets most votes.
- **Ensemble Learning** works best when the predictors are **independent** form each other as possible and we can achieve this by **training them using very different algorithms**.
- **Soft voting** is to predict the class with the highest class probability averaged over all the individual **but this need all classifier to be able to estimate class probabilities which means that they all have a *predict_proba()* method**
- **Soft voting** is often achieve higher performance than **hard voting** because it gives more weight to highly confident votes
- When we train all the predictors we can make prediction for a new instance by simply aggregating the predictions of all predictors using:
  - **Statistical mode** for classification
  - **Average** for regression
- Each individual predictor has a higher bias than if it were trained on the original training set but **aggregation** reduces both bias and variance and the **net result** is that the ensemble has a similar bias but a lower variance than a single predictors trained on the original training set

2 Bagging VS Pasting

- There are two approaches for ensemble learning.
  - **First:** Use different training algorithms.
  - **Second:** Use the same training algorithm for every predictor and train them on different random subsets of the training set.
- There are two types of sampling:
  - **Bagging:** When sampling is performed with replacement.
  - **Pasting:** When sampling is performed without replacement.
- Only **Bagging** allows training instances to be sampled several times for the same predictor

- In **Scikit–Learn** the **BaggingClassifier** class automatically performs soft voting instead of hard voting if the base classifier can estimate class probabilities which means that it has a *predict_proba()* method