Homework 5 of Machine Learning

- 邱寶樟(26)

1. **Overfitting has many faces.**
   During the machine learning process, overfitting often occurs. Different learning methods applied to the same data can result in different reactions in terms of bias and variance. Therefore, using a powerful learner may not necessarily be the best solution. Hence, moderate reduction of overfitting can be achieved through techniques such as cross-validation, regularization, and significance testing.

2. **Feature engineering is the key.**
   Feature engineering requires a significant amount of effort because the process of collecting, organizing, cleaning, and preprocessing data is time-consuming and involves continuous trial and error. Additionally, this task involves domain knowledge. Therefore, finding suitable features for prediction in a large amount of data, while avoiding overfitting or underfitting, is a challenging endeavor.

3. **More data beats a clever algorithm.**
   If the classifier is not accurate enough, one can address this issue by designing better algorithms or collecting more data. However, collecting more data is often a preferable approach because more complex classifiers may be more time-consuming to operate and challenging to adapt to different situations. Therefore, simpler classifiers continue to be widely used to date.