Cpts 315 HW3

Hongqi Guo

011552159

1.

For the voted perceptron, there are nonlinear boundaries. The voting perceptron recorded the weight w.k obtained after the K.th update and the number of times c.k correctly classified the samples in the subsequent training process. Although the voting perceptron improves the generalization ability of the perceptron, it needs to save K weight vectors. Clearly this is nonlinear.

For the averaged perceptron, we have a linear boundary because we just average the two weight vectors. Clearly this single weight vector defines a linear decision boundary.

2.

If the weight of a training sample is very high, it means that it is more important than other samples, so we can infer that this sample should have a greater impact on the current weight of the perceptron. We can increase the sample weight by increasing the sample loss of the perceptron. The way to increase the sample loss is to multiply the sample loss by the sample weight and apply the results.

3.

We can solve this problem by increasing the number of samples for a few classes. A few classes are used to train the perceptron multiple times. In this way, the input of a few people is amplified.

4

5

6.

Summary of a Few Useful Things to Know About Machine Learning

The authors claim that machine learning is the driving force of future innovation, but there are still some problems with machine learning. Machine learning projects take far longer than the required time, and the results of machine learning are not ideal. This paper summarizes 12 important lessons that researchers and practitioners of machine learning have learned. These include pitfalls to avoid, important questions to focus on, and answers to common questions.

Machine learning consists of three components, which are: Representation, Evaluation, and Optimization. The basic goal of machine learning is to generalize beyond training set examples. The data itself is the biggest driver of machine learning accuracy. The best criteria for selecting data is what knowledge is easy to express. Overfitting occurs when we do not have enough knowledge and data to fully identify

the correct classifier. Cross validation can help overcome overfitting, and the most popular method is to add regularization terms to evaluation functions. In machine learning, algorithms with large amounts of data are more valuable for reference. Machine learning is often applied to observational data. However, predictive variables and experimental data are not controlled by learners, and some learning algorithms will extract causal information from the data. The existence of causality is a philosophical question, which has no clear answer.

Machine learning is esoteric and challenging. It has a lot of 'folk wisdom’. Each solution has its advantages and disadvantages.