Report on Homework 3
CS420, Machine Learning, Shikui Tu, Summer 2018
Zelin Ye 515030910468

1 SVM vs. Neural Networks

1.1 Introduction

In machine learning, SVM (Support Vector Machine) is a commonly used classfication method due to its high efficiency and accuracy. Recent years, the neural network has been attracting more and more attention, and also used to solve classification problems. In this homework, I would compare the performances of SVM and neural network (e.g. MLP) on some classification datasets under different experimental settings (e.g. pass).

1.2 Methodology

In this section, I would introduce the datasets and models in my experiments.

1.2.1 Datasets

In my experiments, I use two datasets that are from **LIBSVM Data** [1]. One is **madelon**, a binary classification dataset with 500 features, 2000 training samples and 600 testing samples. Another is called **satimage**, which is for multi-class classification and has 36 features, 6 classes, 3104 training samples and 2000 testing samples.

1.2.2 Models

1.3 Experiments and Results

1.3.1 Preprocess

Most datasets are likely to have missing data, and those in LIBSVM Data are no exception. Therefore, I first make up for the omission in the datasets, replacing empty data with corresponding mean values. Afterwards, I convert the labels from numbers to one-hot vectors,

A Appendix

References

[1] "Libsvm data: Classification, regression, and multi-label." `https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/`.