A Learning Algorithm for Improving the Precision of the Minimum Classification Error Learning

Jun Rokui

School of Information Science ,Shimodaira laboratory,
Japan Advanced Institute of Science and Technology

February 13, 1998

Keywords: minimum classification error, generalization, a penalty term ,the objective function,overlearning.

1 Introduction

Some research to search for the most suitable data on recognition model has been studied. To classify exactly a lot of data in this field of study is in great demand.Until now, Maximum Likelihood extimation is widely used. However, now more highly efficient discriminative learning is demanded. So Minimum Classisication Error has appeared from a different point of learning. In this paper, we take up the problem of Minimum Classification Error which is a highly efficient discriminative learning, and present a solution.

2 Minimum Classification Error

Minimum Classification Error is highly efficient discriminative learning published as the result of Katagiri's and Juang's research work in 1992. All data are learned on MCE, and every parameters are ajusted to a minimum errors in recognition. This differ from method which assumes class form until now, is a method which assumes class boundaries to minimize classification error.

MCE gives a good recognition ability for training data, but does a lowering of recognition ability for generated data. this cause is no information for data distribution in a learning argorithm of MCE. In this paper, we put information for generalization in a learning argorithm of MCE.So we expect improvement in recognition ability, and study possibility.

3 Generalization

We have to deal with questions of generalization separately,on a case-by-case basis. We define as generalization ability which can recognize data added trainig data change as it. In this paper, we define as a fixed index for genaralization ability the ratio of input change to output change. If we could keep output for input as little as possible,generalization ability would improve. In this paper, we define as the new objective function with a penalty term of a fixed index for generalization ability.

4 Proposal

4.1 Abstract

In this paper, we introduce a derived function of input and output a fixed index for generalization ability to a learning argorithm of MCE. We define as the penalty term a fixed index proposed, and add it the objective function for MCE. This proposal is a learning argorithm which minimizes a penalty term and a objective function.

4.2 MCE Multi-layer Feed-forward Neural Networks

MCE can be applied to many new classifier structure such as the multilayer perceptron, learning vector quantizer. In this paper, we use multilayer perceptron. Because there are researches in a area of multilayer perceptron which added a penalty term the objective function, and some of them are confirmed available.

In this paper, multi-layer perceptron type neural network is employed as a platform of recognizer to evaluate the performance.

4.3 Preparation and Experiments

MCE has a probrem that has a strong influence for initial value because MCE is a method which assumes class boundaries to minimize classification error. To decide suitable initial value is very important. So it is necessary to decide suitable initial value to give rough class form before experimentation. In this paper,we use back propagetion before experimentation for MCE and proposed method.

Two types data, artifical data and real world data, were used to evaluate the performance of the learning algorithms.

5 Conclusion

In this paper, we make a comparative study of NN(MSE), MCE, and modification MCE for artificial data. As a result, modification MCE was the best recognition rate for generated data of all. In case of the experiments on real-world data, databases in UCI Machine

Learning Repository from University of California Irvine were used. As a result, modification MCE was the best recognition rate for generated data of all. Modification MCE goes better to improving generalization ability by dominateing to decide the sharpness boundary as a purpose of MCE learning.

The parameter of a sigmoid function as a loss function influences the generalization ability. So we have experimented in an influence of changing the papameter. In case of the experiment on generated data,modification MCE was better than MCE, or as well as it. As a result, we gave the case the proposed method was available regardless of an influence of the parameter.

In this paper, we proposed a novel approach to improve the generalization ability of MCE. The proposed method giving a solution to essentially probrem of MCE can be expected to effects on various methods using MCE.