Deep Learning Course Assignment Report

Bicheng Gao,[1] *Songyu Ke,*[2] *Yuhao Zhou*[3]

[1]*volz.kz.g@gmail.com*
[2]*songyuke@sjtu.edu.cn*
[3]*yuhao_zhou95@sjtu.edu.cn*

14 ACM class, Shanghai Jiao Tong University

ABSTRACT ******edit(accuracy)

The purpose of this research is to classify the tone using the neural network and compare the differences between three neural network frameworks — `Torch`, `MXNet` and `Theano`. Since it seems hard for the neural network to extract the feature from raw data automatically, it is necessary to preprocess the data with some proper methods. We've tried two ways of data preprocessing: eliminating the noise in all datasets or add some noise in training data to fit the environment of test data. With the processing data and Convolutional Neural Networks, we finally achieved the accuracy of _____% in `test_new` dataset. After reaching a better performance, we keep working on comparing the frameworks and provide some major factors listing on this report.

Introduction *****edit

The project of *Deep Learning Course* aims at constructing a proper neural network model to solve tone classification with different frameworks and discuss factors which may have an in-

fluence on the performance of each framework.

With only the raw data, the file: `train.engy`, `train.f0`, `test.engy`, `test.f0`, `test_new.engy`, `text_new.f0`, it's a rather tough task for neural network itself to extract proper feature that may have a good performance in the test dataset. Therefore, preprocessing the data is an appropriate way to change the data, which is much easier to train a better network. We eliminate the low energy noise and use cubic spline interpolation to fix the data with a proper length. Then with those data (in directory `data`) ****edit(data processing procedure)
Since the test data is recording in the noisy environment, which is totally different from the situation of training data, the other thought is to add some white noise in the training data to fit the environment of the test data. We also tried this method to evaluate the performance.
And in the 2nd *Section*, we will provide all the details in our data preprocessing.

After preprocessing the data, another major work need to be decided is the model of neural networks. With different frameworks, the implementation of each framework is different. Therefore, with the same parameters, the performance of three frameworks may have some slight changes. To choose an appropriate neural network model is complicated in many works. In this project, the total size of training data set is 400 with the fixed data shape: $[120, 1]$ ([width, height]). Therefore a small model may work in a small collection of data. Fully-connected Networks and Convolutional Neural Networks are our choices, since it has excellent performance in many tasks. Meanwhile, we also try some complicated models with more layers, and the results of models comparison are also shown in the report.
All the information details of this part is shown in the 3rd *Section* **Model of Neural Networks**

In the 4th *Section* **Background of the Frameworks**, we will provide the elementary infor-

mation of three frameworks: `Torch, Theano, MXNet`. We will first give out a fundamental analysis of three frameworks through their own description and then compare with our own evaluations by the experiment of the tone classification in some aspects, which shown in the 5th *Section* **Evaluation**. ********edit(some factors, reason)**

For the 6th *Section* **Future Work** Part, we're going to show some

Data Preprocessing

Model of Neural Networks

Background of the Frameworks

Evaluation

Future Work

Conclusion