CS200 Junior I.S. Software Description

Minhwa Lee

March 3, 2020

The objective of my Junior I.S. is to explore the application of feedforward neural network (FNN) to real-world data sets. The main language that will be used for the software is Python with its machine-learning library such as Scikit-learn in Tensorflow backend. Scikit-learn is an open-source machine learning library that provides efficient tools for supervised learning, which is an excellent choice for this project. Also, a statistical software R will be mainly used for reformatting the raw data and exporting it to a csv-format file.

The software consists of the two main parts: training and predicting the data set for the two case studies that I will research in the paper. To be specific, the main problem that the software will solve is to perform FNN with multi-layer perceptron and back-propagation and to compare the results of prediction with logistic regression model. In terms of data collection, Iris data is from UCI machine learning and Suicide dataset from Centers for Disease Control and Prevention in South Korea. The flow chart of the software is presented in the figure below. Scikit-learn has been widely used for data predictions, which has many online resources for how to build the FNN using Python. Therefore, the codes written in the software will be mainly referenced from those web sources, with some modifications to be optimized to these two data sets. The following url is a link for an example of the code. (`https://machinelearningmastery.com/implement-backpropagation-algorithm-scratch-python/`.)