Deep Learning

2021 Spring

Project 1

Measure of Machine Learning Models

The purpose of this homework is to observe and measure the effectiveness and features of different machine learning models we learned including Decision Tree, K-NN, SVM, and Neural neteorks.

Data sets

Please refer the three datasets from UCI data repository (https://archive.ics.uci.edu/ml/datasets.html):

1. Iris
2. Breast Cancer Wisconsin (wdbc)
3. waveform
4. MNIST (https://scidm.nchc.org.tw/dataset/mnist) [.csv is recommended]

Except MNIST dataset, the others can be found in the project file.

Learning Models

1. Decision Tree (C.4.5, C5.0)
2. K-NN (K nearest neighbors)
3. SVM
4. Neural networks (3-layer)

Measure and Comparison

1. Except MNIST dataset, separate the others datasets (Iris, wdbc, waveform) into training set (2/3) and test set (1/3). Build your models and test them, then show the precision, recall, F1, and accuracy for each datasets on different models.
2. Repeat the measure of 1. instead of 10-fold cross validation and show the average precision, recall, F1, and accuracy for each datasets on different models (except MNIST).

Hand in Report:

1. Report the models you build and your results for each dataset and the result in .ppt file.

2. You will share your methods and results with your .ppt file on class.