CS7313 Advanced Machine Learning and Pattern Recognition

Project Proposal

Submitted by Angeline Lawrence (a_l492)

Project title:

Analysis and comparison of Linear Regression, Logistic Regression, Support Vector Machine (SVM) and Deep Learning Algorithms

Team member:

Angeline Lawrence

Description of the problem:

Over the past two decades Machine Learning has become one of the mainstays of information technology and with that, a rather central, usually hidden, part of our life. With the ever-increasing amounts of data becoming available, it is important for us to choose the best algorithm for a given data set. This will not only make the work easier for the data scientist but will also result in best predictions.

The purpose of this project is to compare some most commonly used machine learning algorithms such as *Linear Regression, Logistic Regression, SVM and Deep Learning* using the well-known housing prices data set. Towards the end, we will be able to understand the advantages and disadvantages of each Algorithm for the data set under test.

The knowledge gained through this project is not limited to the current data set, but also can be extrapolated to various other data sets related to stock market, biomedical etc.

Dataset Used: House Prices (Kaggle)

Preliminary plan (milestones) :

- Implement Linear Regression
- Implement Logistic Regression
- Implement SVM
- Implement Deep Learning algorithms
- Analyze the results produced

Reference

http://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1540&context=etd_projects

https://www.kaggle.com/c/house-prices-advanced-regression-techniques