Deep Learning Assignment-0

Sri Sruthi Chilukuri

Sjsu ID 014514565

1. a.Github account has already been created with username sruthi1996.

* <https://github.com/sruthi1996>

d.Chapter-1

* Machine Learning: The art of making computers learn from data.
* Popular applications: Spam email filtering, image analysis, classifying articles, content filtering, credit card fraud detection, product recommendation etc.
* The types of ML systems are supervised, unsupervised, semi supervised & reinforcement learning ; based on whether they are trained before or not.
* They can also be classified as instance based learning and model based running based on whether they learn new points from already present data points or if they simply learn from observing patterns.
* Examples of supervised learning systems:
  + KNN
  + Linear regression
  + Logistic regression
  + Svm
  + Decision trees
  + Random forest
* Unsupervised learning examples:
  + Clustering: KMeans, DBScan, HCA
  + Anomaly detection: one class SVM, isolation forest
  + PCA, t-Distributed Stochastic Neighbor Embedding (t-SNE)
  + Association Rules: Apriori

Challenges of Machine Learning:

* When there is no enough Data available
* When the quality of data is poor
* When we overfit or underfit the train data
* Data mismatch

Chapter-2

Approach to solve a machine learning problem:

* Understand the business problem ; frame a clear problem statement of what exactly the ML problem would address.
* Selecting the right performance measure; to measure the efficiency of the ML model developed is important. It can be RMSE, confusion matrix or any other performance evaluation metric.
* Check for the proper assumptions and hidden meaning of each of the parameter in the dataset.
* Fetch the data and setup the workplace
* Visualize the dataset; understand the insights from the data
* Look for correlations between features
* Clean and process the data by handling text and categorical columns(can be done via one-hot encoding
* Select a suitable model and train it
* Use various cross validation techniques to measure the performance.
* Fine tune the model further until we come up with a promising one.