CBD Robotics

Machine Learning in Python

Week 1 (Unit 1 – Basics):

Session 1: Review essential parts of Python (Data collections, OO)

Session 2: Anaconda setup, Jupyter notebook, Pandas, numpy, scipy, Statistics

Project: Bartender project

Week 2:

Session 1: Probability Distributions, Central Limit Theorem, Hypothesis Testing, t-test

Session 2: A/B test, visualization, ANOVA, F-test

Project: Analysis report (using statistical models)

Week 3 (Unit 2 – Supervised Learning)

Session 1: Data Exploration, Feature Engineering, PCA, Feature Selection

Session 2: Linear Regression, Multivariate Regression, Residue, Homoscedasticity

Assignment: PCA computation (by hand)

Week 4:

Session 1: Holdout, Cross Validation, Overfitting, Class Imbalance, Error Types (I &II), Partial Least Square Regression (PLSR), Gradient Descent Algorithm

Session 2: KNN, Naïve Bayes, Logistic Regression

Project: Melbourne Housing

Week 5:

Session 1: Decision Tree, Entropy, ID3 Algorithm, Pruning, Random Forest

Session 2: SVM, Boosting

Project: Airline Arrivals

Week 6 (Unit 3 – Unsupervised Learning):

Session 1: K-means, Clustering, Cluster Evaluation, Silhouette score

Session 2: Mean Shift, Spectral, Affinity

Project: Boston Marathon

Week 7 (Unit 4 – Deep Learning):

Session 1: Neural Network, Back Propagation

Assignment: Construct 2-layers ANN and calculate adaptive weight (by hand)

Session 2: Supervised and Unsupervised techniques in Neural Network

Week 8 :

Session 1: Deep Learning, CNN, RNN, LSTM, Hidden Layer, Convolutional layer, Maxpooling layer, Sub-sampling layer

Assignment: Constructing CNN, RNN (by hand)

Session 2: Tensorflow, Keras & Current Deep Learning model

Project: Fruit classification

Week 9 (Unit 5 – Natural Language Processing):

Session 1: Bag of Words (BoW), BoW Features, Spacy, supervised technique

Session 2: TF-IDF, Lament Semantic Analysis (LSA), Sentence Similarity, Cosine similarity

Week 10:

Session 1: Word2Vec, sense2vec, n-grams, Textrank, Sequence2sequence

Session 2: pLSA, Latent Dirichlet Allocation (LDA), Non Negative Matrix Factorization (NNMF)

Project: Thousand texts – author classification

Week 11 (Unit 6 – Computer Vision):

Session 1: OpenCV, Processing Video, Tracking

Session 2: Corner Detector, Feature Transform, Geotagged Images

Week 12 (Unit 7 – Data Scraping):

Session 1: Scrapy

Session 2: API, Json, HTML scraping

Project: Temperature

Week 13 (Unit 8 – other topics):

Session 1: Big data, Hadoop

Session 2: Hidden Markov Model

Week 14:

Session 1: Time Series, Stochastic Modeling, ARIMA

Session 2: Auto Regression, ARMA

Project: Stock price

Week 15 & 16: Final Capstone project