

**CAP- 5610 Machine Learning
Homework 6**

**Name: Nabila Shahnaz Khan
NID: 5067496**

Course Overview:

The course 'CAP- 5610 Machine Learning' is a basic course which is supposed to give preliminary idea about the field of machine learning. According to my understanding, this course should cover all the different dimensions of machine learning but at the same time it should be easy to grasp so that new students in this field and also students from different backgrounds can easily adapt and be benefitted. Personally, I think this course, conducted by Professor Yanjie Fu, was exactly that and at the same time it was a good combination of theoretical and practical implications of machine learning.

This course gave me a very clear idea about the concept of machine learning, its implementation procedure and possible applications. I learnt about different data categories (nominal, ordinal, interval, ratio) along with their properties and criteria such as noise, outlier, missing values, duplicity to examine the quality of data. Then I got to study about multiple data analysis techniques and data preprocessing approaches such as data aggregation, sampling. Based on these analysis, I learnt how to perform feature engineering which includes methods like feature selection, new feature creation, discretization and binarization, attribute preprocessing (normalization, standardization, scaling). I was also introduced to different distance and similarity metrics such as Euclidean distance, Minkowski distance, Hamming distance, Mahalanobis distance, Binary similarity, Cosine similarity etc. which are the basis of most machine learning algorithms. I learned about different supervised learning algorithms such as Rule-based Classification, Decision Trees, Naïve Bayes Classifier, Logistic Classifier, K-nearest neighbor and SVM. I also got to learn about different ensemble learners and the clear distinction between bagging, boosting and random forest. The difference between Generative models and discriminative models has been made pretty clear in the class. A thorough discussion was also done on different unsupervised learning algorithms such as K-means, Hierarchical Clustering, Density Based Clustering, Fuzzy Set & Fuzzy Logic, Fuzzy K-means and EM Mixture models. The dissimilarity between supervised and unsupervised learning and their situation based utility was explained acutely. Apart from that, different recommender system algorithms and ranking algorithms have also been discussed. Lastly, a preliminary idea was given about multiple types of deep learning algorithms where Reinforcement Learning, Sequential Learning and Representation Learning have been discussed.

As a beginner in the field of machine learning, I found this course to be very intuitive and helpful. Not only I learnt to implement different machine learning models, I also gained theoretical knowledge about how they work which will be really helpful in visualizing efficacy of different models in different environments. I learn to rationally compare the pros and cons of different machine learning algorithms to understand which one might work better in a certain scenario. Different techniques for analyzing and evaluating the machine learning algorithms were also shown in the class. Also, reading the core papers in the field of machine learning and constructive discussions on those papers were really beneficial in understanding the concepts

even better. The homework were suitable for developing the basic idea about the implementation of different machine learning models. And while doing the project, I gained some practical experience on how to process data and features step by step in order to improve accuracy and how to compare the results of different models.

Overall, as a newcomer in this field, I found this course really productive and convenient. Professor Yanjie Fu was really helpful, dedicated and understanding throughout the course. His comments and suggestions really inspired me to work harder and learn better.

Possible Improvement Scope:

I think this course is already very well-designed, so I cannot think of many possible improvements. But, as deep learning is the current emerging field, more emphasis can be given on that topic. I would have liked discussions on some current and possible application fields of different deep learning algorithms to get some insight about which neural network methods might work better in different situations. One simple homework designed on deep learning can be effective in making our understanding even better on this topic and will help students gather some experience regarding deep learning. Altogether, I really liked this course content and structure.