Reflection Essay

My undergraduate degree was in electrical and computer engineering. I used to work with C++/C# mostly and just little experience in python. With a high interest on data science in the field of education, I apply the master degree in learning analytics and aim to extend my skills on the python programming, data analysis and data mining.

“Core Methods in Educational Data Mining” provides an overview of the use of new data sources in education with the aim of developing students’ ability to perform analysis and covers methods and technologies associated with Data Science, educational data mining and learning analytics, as well as discusses the opportunities for education that these methods present and the problems that they may create. Therefore, this course highly in line with my supplementary skills. As I expected, in this course, I learned the basic visualization of python, Data Wrangling, data network analysis, machine learning algorithm like PCA, several clustering algorithms, model evaluation and so on. The depth of all the course content is just in line with my background and my current knowledge.

Unlike my undergraduate engineering-related disciplines, I face the computer system and underlying principles every day. I prefer to use programming languages to directly perform wonderful data analysis and visualization. During my experience in this subject, I met a lot of difficulties in the analytics. For example, I am always unfamiliar with various libraries and data structures of python, and I need to check official documents every time I fetch a number or call an algorithm. And when I am finishing the problem of ICE7, I always fail to add annotation in my visualization of Elbow method to show my elbow point. After I search using google and check the document, I found the solution and resolved it. What I feel most challenging part is ICE4. Because I lack the experience in both the python and machine learning, it is difficult for me to modify the code in this assignment. When calling internal library functions, I always encounter errors of wrong data type or wrong dimension. However, this part is also my favorite in this course as I learned a lot both in python and machine learning algorithm. I really enjoy the process of encountering problems and solving them. What’s more, I love data science and python. I really enjoy the sense of accomplishment that the prediction accuracy rate has improved after my processing.

After this class, I learned the basics of data analysis (including python syntax and data processing, etc.), the use of some simple machine learning algorithms, the principle and evaluation of models, understand and apply the following diagnostic metrics to models: Kappa, A', correlation, RMSE, ROC. In a word, I can say that I can use python to proficiently perform data analysis! Most importantly, this course increased my interest in learning education data mining.

In my future objective, I will have an in-depth understanding of the principles of various data science algorithms, such as principal component analysis, clustering algorithms, and decision trees. I will not only be able to adjust parameters, but also know the specific principles, so that I can know the advantages and disadvantages of each algorithm and when should I use it or not. And I also want to focus on the deep learning algorithm. If I can, I want to be a profession that combines my undergraduate computer major and my master's direction——a data scientist in education. I can give full play to the programming advantages and mathematical background of my undergraduate computer science in this position.