Reflection Essay

Data mining is a vital tool in data analytics. As a learning analyst in the future, I am always eager to learn the strategies of discovering patterns in large educational data sets. However, the term “data mining” sounded very unfamiliar to me at the first place, because I had no previous experience in coding and data processing except building mathematical models by using MATLAB. Even though I had some experience of writing a computer language, all I have done was coping and pasting the code that my professor has written. My personal learning objectives in this course would be understanding the basics of writing code in Python and able to evaluate and analyze data. Moreover, I hoped to be master of writing code at the end of the course.

In the first class, my goals for this course have changed. I realized that I am not a professional computer scientist but a learning analyst who should apply the knowledge of data mining in the field of education. What makes I competitive is my knowledge of education instead of my coding techniques. Since there are a great number of students who have many years of coding experience and have written code for most of their academic life in college, I should rather focus on the application of data mining skills in education to make myself unique and competitive.

I have accomplished eight individual coding exercises, four grouped analysis challenge assignments, and a creative EDM assignment with team members. For each week, the individual coding exercise has different topics and involves the contents of that week. The assignments contain data cleaning, regression in prediction, classification, clustering, PCA, diagnostic metrics, SNA, correlation mining and casual mining. In addition to these main topics, we have also discussed behavior and affect detection, knowledge tracing and data visualization. My biggest take-aways from this course are regression in prediction and data visualization. The regression prediction is one of the most major functions in data analysis, and data visualization is a very useful tool which helps people who read codes understand clearly and easily. The most challenge part for me were PCA and diagnostic metrics.

My learning goal for this course is to identify inquiry questions and be able to solve problems using the techniques. I have learned to analyze the questions and understand data before I operate it. There are a lot of techniques taught in this course which are very complicated, but I have remained all my work and notes in a file in case I need them. Besides the goals, I have also learned how to look for hints and codes online and in Python, which facilitate me in the progress of coding.

Furthermore, I definitely improve my awareness about new issues related to education from this course. In the group work, I participated in the engagement from various aspects, such as arranging meetings, writing codes, inviting professional computer scientists as guests, motivating group member etc. In my future career, I believe that these skills, including technical skills and soft skills, will help me to do my job well and get along with others in the career filed as well.