Yutong Shen

HUDK 4050 Reflection Essay

At the end of the semester, I will briefly discuss my takeaways and reflections for HUDK 4050.

This is my first semester at TC, and I am an M.S. student in the Learning Analytics program. I graduated in May from Boston University, and I majored in Psychology and Computer Science, so I have some Computer Science background. However, I did not know too much about the educational field. I am taking HUDK 4050 this semester since I am interested in learning about educational data mining, and it is a required course for my program. By taking this course, I expect to explore the ways data mining is applied to solve educational problems. I also hope to exercise my ability to perform data analysis.

At the beginning of the semester, my personal learning goals for HUDK 4050 are to become familiar with data analysis methods and interpret the results. But as the semester progresses, I have more objectives. I aim to adopt the appropriate data mining methods to answer the identified educational questions and evaluate the implications of educational data mining. During the semester, I have learned various data mining methods, including Linear Regression, Logistic Regression, Decision Tree, Clustering, PCA, Diagnostic Metrics, and Social Network Analysis. I have also learned data cleaning, feature selection, and how to utilize different python packages. I realized the importance of data privacy issues. My biggest takeaway from this course is the understanding of how educational data mining methods allow us to explore the increasingly large-scale data from educational settings. By using the developing methods, we can better understand students and schools. Therefore, we can find ways to improve students' learning.

At this point, my favorite engagement experience for the course was when I worked on the ACAs. I can practice the data mining methods learned in class and connect them with real-world

examples. I found the Creative EDM Assignment the most challenging since we had to identify our own inquiry questions and craft an analytical plan. Since we worked in groups, we discussed and shared ideas to approach this assignment. Through the eight coding exercises (ICEs), I learned about data manipulation techniques and different methods to perform analysis. Through the four Analysis Challenge Assignments, I practiced these skills and abilities. In ACA1, I used data masking to access the specific data elements that are needed. In ACA2, I adopted logistic regression and decision tree to build the classifier for prediction. In ACA3, I tried PCA and Silhouette Coefficient to reduce dimension and identify clusters. In ACA4, I applied social network analysis to identify and analyze a network.

To sum up, after participating in this course, I believe that I have achieved most of my learning goals. However, I found that I am not very proficient in making interpretations based on the analysis I performed, so I can further improve my ability to interpret the results and practice more with real-world cases. In addition, I can learn more about the python packages for educational data mining. In the future, I might want to work on neural networks and Natural Language Processing, and I am looking forward to learning their implications in the educational field.