Cover Letter

This is Yingjun Guan, a graduate student majors in Information Management from the School of Information Science, UIUC. I am writing to express my great interests in the Whole Tale 2017 Summer internship opportunities. I am very happy to have the chance to apply for the program, and I am that confident that my skills and previous experience is a great match for the project. Among all three projects, I’m mostly interested in the first project: Online Materials Science Modeling and Analytics.

During my undergraduate study and graduate study and research, I had a lot of opportunities to programming with various languages and tools such as Python, R, MatLab, Fortran, etc. (For more information, you can view my webpage and GITHUB repository listed on my resume). In my master study of Information Management, I get myself familiar with data analytics packages in R, such as Pandas, Numpy, Yt, cImage, Plotly, Scipy, etc. I have also applied Python to my course work in *Data Visualization*, *Data Mining*, and *Network Analysis*. I also learned and applied R in my course project of *Data, Statistics and Information* for text mining and statistical modeling.

I also have a bachelor and master degree from Civil Engineering, which enable me to have basic material knowledges. I then further researched on the non-linearity properties of construction materials such as concrete, cement and PMMA for one year in Construction Material Group in PSU. I have got two publications for my previous research, one of which is about the non-linearity characteristics of the concrete material in the field of Structural Health Monitoring (SHM).

During my previous experiences, I employed data analytics techniques into a lot of projects. I enjoy the process of transforming the knowledge I learnt in class to real world projects. In my study in UIUC iSchool, I have participated in the project of Natural Language Processing (NLP) with professor Jana Diesner in which I’m responsible for data cleaning and parsing and collaborate with the emotional analysis from abundant email information. I have also participated in the project of machine learning for affiliation analysis with Professor Vetle Torvik, where I’m responsible for the analysis for the spatial distribution and data visualization of academic publication affiliation. In my future PhD study in iSchool, I plan to explore more about machine learning and data visualization, and participate in more projects that is relevant to data analytics.

For the Whole Tale Projects, I wish I could have the chance to learn the process of this project and the function of programming and data analytics in it. I would also like to collaborate with my future teammates to improve and clarify our goal as we are working on the project. In a broader vision, I would help to employ multiple machine learning models for various database within MDF. For instance, I would suggest to research on the data in detail to determine the machine learning type (supervised and unsupervised methods) and various instance-based or attribute-based classifiers. Finally, although Jupyter notebook is easy to demonstrate and reproduce the project results, I wish I could contribute to diverse ways of data visualization.

I’m very eager to participate in this research. I wish I could color my summer with the project, also help with the project with myself.