web: *https://daki7711.github.io/* email : davidyjk@colorado.edu

Statement of Purpose

of David Y.J. Kim

Research Statement: Understanding how an individual's behavior reflects his/her comprehension, intention, and cognitive state is an interesting topic that can be used in many ways, such as predicting future achievements, diagnosing one's learning processes, and more. My research goal is to create a solid framework that uses Machine Learning (ML) and Deep Learning (DL) to infer explanatory variables of human behavior from real-world data, then further apply them to support or refute hypotheses regarding the human mind. There are various and thorough studies on this topic that have yet to be conducted. Two critical problems come to mind. First, it is vague and uncertain about how to measure and quantify the mechanisms of human behavior. Second, previous models are not powerful enough to fully capture the various factors that determine human behavior. However, with the development of ML/DL, we may have a chance to overcome these limitations. Natural language processing research has opened the possibility of quantifying natural language semantics. With recurrent neural networks we are able to analyze sequences of human actions such as their reading patterns. Computer vision models can support neuroscience research that utilizes functional MRI (fMRI) to visualize the human brain. Despite this fascinating research, most of the results are based on fairly structured data. Most of the time, the input variables and optimization process are relatively straightforward, but human behavior data usually do not have this luxury. Hence there exists a gap in applying ML/DL models to analyze human behavior effectively. I intend to reduce this gap and model human behavior, which will enhance our practical ability to understand and predict the human mind.

Research Experience: During the past year, I collaborated with Dr. Michael Mozer to develop intelligent textbooks that infer student comprehension based on students' interactions with digital textbooks. Our focus is whether highlighting behavior can be used to predict a student's comprehension and retention of factual material and concepts.

We used observational data in a genuine educational setting with the aid of OpenStax, a nonprofit organization that supports open-access college-level digital textbooks. We collected data from 11,134 students in Biology, Physics, Sociology, and History classes for two full semesters. Students were able to highlight and add annotations to their e-textbooks while reading. Given these records, we attempted to infer student comprehension, as assessed by a quiz that students take at the end of each section.

Next, we focused on identifying the relationship between the pattern of highlighting and subsequent retention and understanding. While exploring alternative approaches to representing the pattern of highlights, the base fundamental we kept in mind is that the feature representation should be compact to reduce redundancy but still expressive enough to capture the signal in the data. We explore many different representations of the pattern of highlights and discovered that creating a vector where each element indicates which of the word is highlighted, then reducing the dimension using principal component analysis is most effective. We found that when students choose to highlight, the specific pattern of highlights can explain about 13% of the variance in observed quiz scores.

Currently, I am further exploring better highlight representations. The next strategy I intend to investigate is considering the semantics of the highlighted text. I plan to tackle this by building a framework that uses the highlight and quiz questions' textual content. For instance, using BERT (Bidirectional Encoder Representations from Transformers), it is possible to find sentence pairs with the most similar meanings.

Using this idea, feeding the highlighted sentence along with the questions into our model will produce a single evaluated score indicating how closely related the highlights are to the question. I will investigate if this score can improve the prediction of students' quiz performance under the Item Response Theory(IRT) Model[1].

Reason for applying: My current research experience has excited me about what is involved in analyzing human behavior, and I want to continue studying this topic as a Ph.D. student. While continuing my study of applied Machine Learning, I am looking for an interdisciplinary program that will allow me to engage with related academic fields. For example, a fundamental understanding of Cognitive Science and Psychology would be crucial in understanding how humans act. A stronger background in Neuroscience could open new pathways to find the connections between brain activity and behavior. I want to expand my research skills by gaining experience with data beyond textual material, e.g., visual data, time-series data, signal data. After my Ph.D. study I plan to continue as a research scientist in academia, contributing to improve machine intelligence capable of understanding and collaborating with humans. Considering all of this, I cannot think of a better choice than pursuing a Ph.D. as my next academic career.

[1] a family of mathematical models that attempt to explain the relationship between latent traits (such as ability, item difficulty) and their observed outcomes