Team Viserion

Shasha Lin, Julie Helmers, Millie Dwyer

1 Team Members

Millie Dwyer

Millie is a second year MS in Data Science student at NYU's Center for Data Science (CDS). Prior to entering the program, she received a BA in Math from NYU. Afterwards she worked for 3 years in data analytics at an adtech company. This past summer, she interned at Aetna with the Analytics and Behavioral Change department. At CDS she has taken the core data science classes, as well as electives in Deep Learning and Optimization and Computational Linear Algebra. She is currently taking Natural Language Processing (NLP). She is interested in exploring deep learning applications through this capstone project.

Julie Helmers

Julie is also a second year MS in Data Science student. She has about 1 year of experience as a data science intern at Major League Baseball and 2+ years of experience in a computational cognitive neuroscience lab conducting EEG, genetic, behavioral, and pharmacological studies. Prior to NYU, she received her BS in Cognitive Neuroscience from Brown University. Her past/current CDS electives include Deep Learning, Optimization and Computational Linear Algebra, Statistics and Machine Learning of Genomics, and Data Structures and Algorithms.

Shasha Lin

Like Julie and Millie, Shasha is in her second year at CDS. Before CDS, she obtained a Master's degree in Research and Experimental Psychology and worked as a graduate assistant in psychiatry labs at Mount Sinai Hospital, where she did quantitative research on cocaine addiction, honed her statistical skills, and developed a deep understanding of human behavior data. While a student at CDS, Shasha has been exposed to and excited by the application of deep learning in a diverse range of domains, such as language modeling/machine translation, text classification, music analysis, and biomedical predictions. This past summer, she worked with Applecart LLC, a start-up that builds machine learning models to predict political voting behaviors, where Shasha prototyped the company's first semi-supervised learning model.

2 Team Motivation

Millie, Julie, and Shasha collaborated on three class projects while taking Deep Learning class, including image classification/generation and natural language modeling, and have a shared interest in further developing their deep learning skills and applying them to challenging problems in the real world, such as machine-based language understanding, frontiers of biomedical research, and problems in the financial/business sectors.

3 Project Bids

Princeton University: Word Embeddings for Quotes

Since being an English major in college, Shasha has been very interested in the intersection between human language and artificial intelligence, especially with human vs machine translation, and she furthered this interest last semester with her work on legal document classification using doc2vec models. Both Julie and Millie also have deep learning and NLP experience from their coursework,

and they possess basic JavaScript and HTML skills, which are required for this automated question-answering task. The entirety of Team Viserion is especially excited to have a production quality chrome application resulting from the Capstone project.

Vector Institute for Artificial Intelligence: Decoding the Genome

We all have some degree of experience with biomedical data, which will be applicable for this exciting project. Julie has been especially interested in bioinformatics and genetics since her time in a research lab at Brown, where she studied the role of certain genes in human learning and decision-making and attended an intensive workshop in biostatistics. She is currently exploring the application of statistical learning to large-scale biological data in Professor Satija's Statistics and Machine Learning of Genomics class. Millie and Shasha's experience and interest in health care and biomedical data stem from their time at Aetna and Mount Sinai, respectively.

Industrial Classification

Team Viserion are also motivated to expand our academic and work experiences into the business/finance world. Both Shasha and Millie are currently learning about different word embedding methods from Professor Cho's NLP class, and collectively we feel confident about accomplishing this project in a high quality manner. Furthermore, Shasha is familiar with the distributed skip-gram model mentioned in the project proposal. We enjoyed Dr. Gill's guest lecture last week and are excited about using advanced data science techniques for economic research and the financial profiling of different companies.