Big Data, Innovation, and Business Strategy
Trevor Huang (EAS 2019)
Advisor: Anoop Menon

My project was to design and write a program that could take any patent from the United States Patent and Trademark Office (USPTO) and evaluate its quality through machine learning. The goal of the project was to create the deep neural net architecture with some preliminary results so that in the future others could easily refine the machine and easily modify its structure. I worked under Dr. Anoop Menon with two other underclassmen at Penn.

Machine learning is a fast-growing branch of computer science that is showing promise in analyzing problems that were normally considered too complicated for a machine to understand and solve. These problems include image recognition, language translation, and many others. In particular, my research was about using deep neural networks to evaluate patents. Deep learning is a special type of machine learning where the input (in this case the patent itself) is represented as some matrix of numbers and given to the machine. The machine then transforms this matrix in any way the programmer sees fit to try to reduce the matrix to a single value, which is the output (in this case, some number that represents the quality of the patent). The matrix transformations can be as simple as matrix multiplication to the more complex dimensionality reduction through Principal Component Analysis. During training, the machine sees if its answer is the same as the training data’s answer and changes numbers to try to get more of the training data correct.

Through this research I greatly improved my computer programming ability and communication skills, and learned a lot about research itself. The programming knowledge will help me in the future as I continue software engineering. Because I had never worked in a team environment on such a large project, I learned to rely on others for areas I am not as specialized in and figured out how to cooperate when programming. The project made me understand that research is very different from an assigned school project. To our knowledge, no one has done quite the project we were working on, which meant that we could not rely on other’s experiments and work to know if we were on track or what the best solution was. Research is about figuring out what works and what doesn’t, and learning the most from the failures. Many things broke when we were...
programming the machine, but that’s how the team bonded and could cooperate to make something better.

I gained insight on how research is conducted through my experience with PURM. While my team was unable to hit as high an accuracy in prediction as we were hoping for, the groundwork has been set for how patents can be analyzed by a computer. I enjoyed my research with Dr. Menon and hope in the future to be part of the team to get this project even farther.