Capstone Form

UNIVERSITY OF WISCONSIN
DATA SCIENCE

The following form is to capture an overview and objective of your capstone project for publication. *Sensitive information can be withheld as long as the overview and objective of your project is provided.*

*Fields are required however, may be optional for publishing.

1. Choose the semester:

   ☑ Spring   ☐ Summer   ☐ Fall

2. What year was the project completed?

3. What is your first and last name?* Mukund Raghav Sharma

4. What is your capstone project title?* Comparative Case Study Between Gated Graph Neural Networks vs. Relational Graph Convolutional Networks for the Variable Misuse Task Using Deep Learning and Python.

5. Describe your project that includes the objective in 250 words or less.

The purpose of the comparative case study based project was to discern the better model between models based on Gated Graph Neural Networks (GGNN) and Relational Graph Convolutional Networks (RGCN) on the Variable Misuse Task, a prediction task involving discerning the correct variable to be used in a particular spot amongst all variables of the same type in the particular scope. The comparison between GGNN and RGCN models involved computing the test accuracy on three experiments the source data of which is obtained by downloading the source code of the top 25 trending C# repositories on Github. These three experiments involved training and obtaining the test accuracy of all the repositories, an esoteric and popular repository to deduce which model was more performant across different types of source code.

The overarching goal for this project was to discover which model would generalize and perform better in the Static Analysis tooling space that's typically rule based by inculcating the representational power of Deep Learning to solve more state-of-the-art problems.

Capstone Form

UNIVERSITY OF WISCONSIN
DATA SCIENCE

The Data Science concepts related to this capstone are machine learning (specifically deep learning),

hyperparameter optimization, visualization techniques, graph theory and natural language processing. Also,

putting together the final report involved a lot of the writing techniques to present the data in a simple yet

compelling manner.

The results from the data highlighted that the Relational Graph Convolutional Network outperformed the Gated

Graph Neural Network across all experiments, although, within a margin of 5%. Training on more data resulted

a higher test accuracy for both models and a smaller difference between the two.

6. Provide a photo from your project or an image that best represents your capstone.*
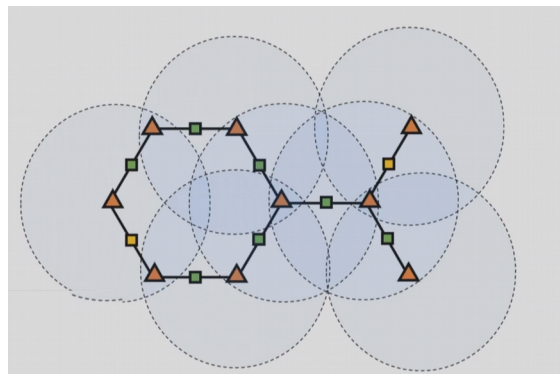
7. What is the name of the company or site?* N/A

   a. What is the city in which it is located?* N/A

   b. What is the state in which it is located?* N/A

   c. What is the country in which it is located?* N/A

   d. Would you like to have the company or site information withheld?*

8. Would you like to share a testimonial of your capstone experience?*

Capstone Form

a. Provide a testimonial in regards to your capstone experience.

Working on the capstone was probably one of the most fun times I have had in a while. Gaining proficiency over the two and a half years of this Master's program, I felt I had to take on a challenging topic pertinent to my work as a software engineering in the developer tooling space to research methods of improving the overall user experience. The specific area I wanted to make a difference in is static analysis tooling by incorporating deep learning in the field of detecting errors before they affect the users during program execution.

This project involved a deep understanding of the complex literature that's at the bleeding edge of this field. I recall, fully understanding the first page of the paper that this paper is based off took me almost a week but my intuition grew over time by constant repetition and by the end of the first month and a half, I had covered a majority of the literature my capstone was based off of. Next challenge was to identify and grok the abstractions of the code that the authors used to generate the results and then adjust it to tailor to the architecture of my pipeline and despite my decade worth of experience programming, it took effort to understand the rationale. There was no better joy when I finally saw the results as it demarcated the culmination of a full understanding of the training and testing processes involved.

I can attribute courses such as Data Mining & Machine Learning and Prescriptive Analysis as ones that directly helped me with this capstone, however, I also give credit to a lot of the other non-technical courses that helped me with thinking the way an ethical and well-written Data Scientist should.

To be honest, I barely scratched the surface on this extremely vast field but what this capstone project did help me with is truly understanding the nuances of this confluence of deep learning and static analysis, a field that's pretty much in its infancy. However, I relish the progress I made from scratch to a point where I understand the structure and methodology to conduct further experiments.

b. Provide a photo of yourself to be included with the testimonial.