Shruti Singh

Dayton, Ohio | +1- (937) 716-8664 | 97shruti@gmail.com | Linkedin

EDUCATION

University of Dayton — **Dayton, Ohio**
PhD student in Computer Science — *August 2023 - December 2027 (Expected)*
Masters of Computer Science — *January 2021 - August 2022*

WORK EXPERIENCE

University of Dayton — **Dayton, Ohio**
Graduate Research Assistant — *August 2022 - Present*

- Developed a deep reinforcement learning environment for the Lunar Lander case study and trained using Proximal Policy Optimization (PPO) algorithm for feature selection.
- Composed a facial recognition attendance system for online meetings, with a 99.39% accuracy rate of facial recognition, utilizing Python and libraries with OpenCV, and dlib.
- Combined the facial recognition algorithm in Python, and a screen capture method, in the online class platform system.
- Experimented with six multi-class multi-output algorithms to predict career paths based on historical background, achieving an accuracy of 91.21% and 95.97% for labels domain and positions respectively.
- Spearheaded a Unity 3D application to create a roller coaster environment in augmented and virtual reality.

Democracy Lab — **Bentonville, Arkansas**
Software Engineer — *March 2023 - August 2023*

- Incorporated Python scripts to automate the data migration process from one container in Azure CosmosDB, to another optimizing the runtime by 70%.
- Implemented the replacement of ‘moment’ library with ‘date fns’ in the front-end JavaScript, resulting in improved performance and reduced bundle size by 30%.
- Refined a formatting solution for the Contact email field in Volunteer Applicants, ensuring the accurate formatting of string.
- Transformed a ‘reloader’ dependency for faster updates of rendered pages, eliminating the need to rebuild the Docker container for code changes to take effect.

Walmart — **Bentonville, Arkansas**
Intern — *June 2022 - August 2022*

- Engaged with the development, logistics, and supply chain team to analyze and assess barcode on GitHub.
- Executed the replication and validation of the barcode logic in Python, leading to optimized scanning.
- Researched the GS1 barcode standards documentation to include conditions for 10 specific case exceptions such as over sized items, perishable items, and hazardous materials.
- Initiated a standard barcode logic using Python in 3 phases; defined a function to (1) validate the barcode structure,(2) to generate the barcode checksum, (3) to encode the data into a barcode symbology.
- Compiled a big query search using SQL to collect over 100 test inputs for the standard logic, including barcode types, product categories, and case exceptions.
- Verified the standard logic's accuracy to be 100% as compared to the legacy logic and visualized the results.

Mindzcloud — **Nagpur, India**
Software Engineer — *July 2019 - September 2020*

- Exported and imported Salesforce account, contact, and opportunity data in bulk using Data Loader, and performed data cleaning on CSV and Excel files to ensure data accuracy and consistency.
- Collaborated across 4 multi-functional teams (business, product, testing, and development) to correspond on user stories for new features, generate data reports, and review workflows for onboarding new customers.
- Improved 30 trigger cases written in Apex language to automate the creation of new leads on contact creation.
- Resolved user story problems related to the approval process for new accounts, the creation of customer support flows, and the publishing of Knowledge Articles.

RESEARCH PROJECTS

Deep Reinforcement Learning-Based Feature Selection

- Modified a case study for Lunar Lander with the gym library, using the state variables such as the lander's position, velocity, and angle to train a PPO model and land the lander safely.
- Conducted experiments under the Proximal Policy Optimization (PPO) learning technique to evaluate feature importance in the Lunar Lander environment.
- Streamlined 2 filter methods; first, information gain and then, correlation analysis to systematically evaluate and select relevant features within the observation space.
- Calculated entropy and joint entropy to assess the information content and relationships among observations, enhancing the reliability of feature selection.
- Defined 2 key performance metrics, "time" and "distance," to quantitatively measure and improve the reliability of cyber physical systems.

Career Prediction based on Historical Background

- Gathered and processed 420 LinkedIn profiles, each containing 23 features in Excel, including education, career history, skills, and endorsements.
- Estimated variance, entropy, and joint entropy analysis techniques to reduce the initial 23 features down to 11, achieving feature analysis and reduction.
- Trained with six multi class multi output classifiers, where random forest classifier emerged as the top-performing model, delivering an accuracy of 91% and 95% in predicting career trajectories.
- Evaluated the model's effectiveness by employing a real-world test set, achieving an accuracy rate of more than 90% for domain and position class labels.

Retail Recommendation System

- Modernized a recommendation system using e-commerce data consisting of rating, product id, and timestamp.
- Devised popularity based and collaborative filtering recommendation algorithms, and optimized algorithm parameters using grid search.
- Designed and led a novel collaborative filtering algorithm that incorporates user feedback to improve recommendation accuracy.
- Achieved a 10% improvement in mean average precision (MAP) and normalized discounted cumulative gain (NDCG) scores over traditional collaborative filtering algorithms.
- Validated a new user feedback mechanism that allows users to rate and review recommendations, and used this data to train the collaborative filtering algorithm.

Search Strategies : Gridworld Environment

- Spearheaded the development of a gridworld environment for a goal-seeking task, featuring single agent obstacles, and an ultimate goal location.
- Applied 4 search algorithms to navigate the gridworld environment, differentiating between uninformed and informed search strategies to find optimal paths.
- Computed the DFS, BFS, UCS with a cost matrix and A-star algorithm with Manhattan distance heuristic, to plan and execute optimal paths within the learning environment.
- Revised the performance with three reinforcement learning algorithms; value iteration, policy iteration, and Q-learning with epsilon greedy strategy.
- Traced and displayed action trajectories for all the 7 different agents, showcasing differences between uninformed and informed methods for the same agent and environment.

TECHNICAL SKILLS

Skills: Python, MATLAB, SQL
Frameworks & Technologies: NLP, Gym, OpenCV, Tableau, Tensorflow, Keras, GCP
Publications: "Multi Output Career Prediction - Dataset, Method, and Benchmark Suite"
Achievements: Walmart Data Science Bootcamp, Advanced Software Engineering, Black Wings Hackathon