Shraddha Shekhar

Buffalo, NY | +1 716-994-5448 | iamshraddhashekhar@gmail.com | www.linkedin.com/in/shraddha-shekhar/ | **Portfolio**

Education

University at Buffalo, The State University of New York Aug 2022 - Jan 2024
Master of Science in Computer Science and Engineering GPA:**3.92/4.00**
Courses: Natural Language Processing, Information Retrieval, Distributed Systems, Machine Learning, RL

Savitribai Phule Pune University Aug 2017 - July 2021
Bachelor of Engineering in Computer Engineering GPA: **9.32/10**
Courses: Data Structures, Computer Architecture, Computer Networks, Operating Systems, Database

Technical Skills

Languages: Python, C++/C, Java, GoLang, SQL, JavaScript, HTML.

Libraries & Frameworks: Numpy, Pandas, NLTK, Pytorch, Keras, Tensorflow 2, HuggingFace, LLMs, BERT, OpenAI Gymnasium, Scikit-Learn, Matplotlib, Dash, Flask, Django, Dask, Spring Boot.

Databases & Tools: MySQL, MongoDB, Elasticsearch, Google Cloud Platform, AWS, Kafka, Apache Solr, Linux command line, Git, Postman, Docker, Jira, Tableau, Agile, Scrum.

Work Experience

Software Development Engineer, Persistent Systems, Pune, India May 2021 - July 2022
- Built a project in collaboration with the Government of India, driving website traffic analysis and improving security through predictive identification of malicious actions.
- Led the team in automating large-scale log generation saving 12hr/day and developed a robust Data Simulator (Secure Data Lake) using Python, Elasticsearch, Kafka and Docker, achieving 97% data integrity rate for load testing.
- Collaborated with red team to simulate security threats and evaluating multiple edge cases, resulting in a 20% improvement in threat mitigation strategies.
- Developed a python ETL pipeline handling 40 million logs daily using Apache Kafka.
- Enhanced system performance by implementing parallel processing with Dask, resulting in 15% faster asynchronous log generation across 15000 user.
- Optimized the open-source Wazuh agent-simulator with 6 new features expanding it's functionalities and robustness.

Software Intern (Data Science and Python), Verzeo, Pune, India Dec 2020 - Mar 2021
- Performed data preprocessing, feature engineering and visualized insights through responsive dashboards.
- Devised machine learning algorithms such as random forest, logistic regression and apriori for Market Basket Analysis.

Projects

Empathetic Generative Chatbot: Multi-topic Conversational AI | GPT2, LLMs, NLP, GCP, Reddit API, TF
- Spearheaded development of a generative chatbot utilizing GPT2, RoBerta and information retrieval techniques on 200K reddit posts to efficiently understanding emotions and holding a conversation with context change.
- Fine-tuning GPT2 base model with NER tags and Sentiment Analysis to generate empathetic responses along with BERT and RoBerta to enhance factual responses.
- Deployed a flask app for the chatbot on Google Cloud Platform, serving 15+ turns of coherent responses.

Educational Web App to Visualize RL Methods | AWS, RL, Gym, Pytorch, Flask, Python
- Developed an interactive Flask web application for RL course to demonstrate intricate workings of RL algorithms, Q-Learning, SARSA, Double DQN, to solve a custom multi-agent environment.
- Deployed the app on AWS EC2 to ensure accessibility and elevate students' understanding of RL concepts through visualization of algorithms and effect of various parameters on agent's learning process.

Natural Language to SQL Conversion with LSTM and AI | TensorFlow, MySQL, Word2vec
- Built deep learning models with LSTM, encoding, and vector embedding for converting natural language sentences into SQL queries, leveraging NLP and fundamental AI techniques for decision making.
- Designed an interactive Dash (plotly) front-end and integrated MySQL, achieving a 83% accuracy in fetching correct results for generated queries. (http://www.jetir.org/papers/JETIR2106077.pdf)

Co-curricular & Achievements

Awards: Awarded by Peer's Choice, acknowledging dedication and effectiveness in delivering educational support as a TA.
Graduate Assistant: CSE 546 - Reinforcement Learning (Spring 2023), CSE 574 - Intro to Machine Learning (Summer 2023, Fall 2023)