Shraddha Shekhar

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

Summary

Graduate student with 2 yr experience as a Software Engineer. My expertise are in Python, databases (SQL/ NoSQL), cloud, ML, NLP and delivering scalable and efficient solutions. I am a quick learner with strong problem-solving skills.

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: Machine Learning, RL, Natural Language Processing, Information Retrieval, Distributed Systems

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, Java, SQL, C++, GoLang, JavaScript, HTML, CSS.

Libraries & Frameworks: RestAPI, Flask, Django, Dask, Spring Boot, Maven, JUnit testing, ReactJS, Firebase, JSON.

Machine Learning: Numpy, Pandas, Pytorch, Tensorflow, Transformers, LLMs, LoRA, OpenAI, SkLearn, Matplotlib

Databases & Tools: MySQL, MongoDB, PostgreSQL, Elasticsearch, Google Cloud Platform, AWS, Kafka, Apache Solr, Linux command line, Git, Postman, Docker, Object-oriented design, SDLC, Jira, Tableau, CI/CD, Agile, Scrum.

Work Experience

Software Engineer Intern, JerseySTEM | New Jersey Mar 2024 - Present
- Analyzed real-time and historical employee data to identify trends and anomalies, conducted exploratory data analysis (EDA) and visualized insights with Matplotlib, and PowerBI, which reduced HR operational inefficiencies by 11%.
- Developed and implemented predictive models for employee engagement and attrition using Logistic Regression, Random Forest, SVM, XGBoost, AdaBoost, and LightGBM, resulting in improved retention strategies.
- Collaborated with stakeholders to understand business requirements and resolving over 30 data inconsistencies.

Software Engineer (Data work)**, Persistent Systems** | Pune May 2021 - Aug 2022
- Developed multiple microservices, analysing website traffic and improving security by predicting and identifying malicious actions using machine learning.
- Led the team in automating large-scale data generation saving 12hr/day and developed a robust Data Simulator microservice (Data Lake) using **Elasticsearch** (NoSQL)**, Kafka and Docker**, achieving **97%** data integrity rate.
- Integrated a python **ETL** pipeline for streaming data, handling **40 million** logs daily using Apache Kafka.
- Enhanced system performance by implementing parallel processing with Dask, resulting in 15% faster asynchronous data generation across 15000 user.
- Executed advanced SQL queries to detect anomalies and crafted Kibana dashboard improving user engagement by 15%.
- Developed **RestAPI** to run data simulator on specified **Google Cloud** instance using IP addr and log specifications.

Data Scientist Intern, Verzeo | Pune Dec 2020 - Mar 2021
- Performed data extraction, preprocessing, feature engineering and visualized insights through Tableau dashboards.
- Devised machine learning models using random forest, KNN, and apriori for Market Basket Analysis.

Projects

Generative Chatbot: Empathetic Multi-topic Conversational AI | GPT2, LLMs, LoRA, NLP, GCP, API, TF
- Developed a Generative chatbot using GPT2, RoBerta on 200K reddit posts reflecting emotions and factual accuracy.
- Finetuned GPT-2 model with NER and Sentiments. Deployed flask Chatbot on GCP, 15+ turns of coherent responses.

Educational Web App to Visualize RL Methods | AWS, RL, Gym, Pytorch, Flask, Python, REST API
- Deployed Flask app on AWS EC2 to enable students visualize and interact with RL algorithms, boost understanding.

Fault-Tolerant Distributed Key-Value Storage | Go, gRPC, Distributed Systems
- Built a reliable distributed key-value storage system in Go. Implemented the Raft Consensus algorithm to elect leaders and replicate logs, achieving consensus in leader election and log replication for a system consisting of five nodes.

Natural Language to SQL Conversion with LSTM and AI | TensorFlow, NLP, MySQL, Word2vec
- Built deep learning models using LSTM and vector embedding for converting english sentences into SQL queries.
- Designed Dash app and integrated with MySQL, achieving 83% accuracy in fetching correct query results. [paper link].

Co-curricular & Achievements

Awards: Awarded by Peer's Choice, acknowledging dedication and excellence in delivering educational support as a TA.
Graduate Assistant: CSE 546 - Reinforcement Learning (Sp 23), CSE 574 - Intro to Machine Learning(Sum 23, Fall 23)