

Shraddha Shekhar

+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

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: 3.95/4

TECHNICAL SKILLS

Languages: Python, R, SQL, Matlab, Java, C++, Javascript

Machine Learning: Pandas, Numpy, Pytorch, Keras, Tensorflow, Transformers, HuggingFace, LLMs, LoRA, PEFT, BERT, NLTK, Scikit-Learn, Matplotlib, OpenAI Gymnasium, Flask, Dash, CNN

Databases & Tools: MySQL, MongoDB, PostgreSQL, AWS, S3, RDS, Google Cloud Platform, Elasticsearch, Hadoop, HDFS, Spark, Apache Solr, Kafka, Oracle, RestAPI, Linux, Git, Postman, Docker, Jira, Tableau, PowerBI, Agile, Scrum

WORK EXPERIENCE

Software Engineer Intern, JerseySTEM | New Jersey

Mar 2024 - Present

- Built an AI agent for converting discussion to actions, leveraging **LangChain** and fine-tuning LLama2 LLM.
- 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%.
- Implemented predictive models for employee engagement and attrition using Logistic Regression, Random Forest, **SVM**, **XGBoost**, AdaBoost, resulting in improved retention strategies.

Software Engineer (Machine Learning Engineer), Persistent Systems | Pune

May 2021 - Aug 2022

- Implemented ML models like **KMeans** for clustering similar logs and DBScan for anomaly detection to identify malicious logs in website traffic. Used **NLP** techniques and TF-IDF vectorization for log processing.
- 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 and enhanced system performance by 15% with parallel processing using Dask across 15,000 users.
- Analyzed data with advanced **SQL** queries and crafted Kibana dashboard improving user engagement by 15%.
- Developed RestAPI to run data simulator on specified **AWS EC2** instance using IP addr and log specifications.

Software Engineer Intern, Verzeo | Pune

Dec 2020 - Mar 2021

- Performed data extraction, preprocessing, feature engineering and visualized insights through Tableau dashboards.
- Optimized customer purchasing patterns by 23% using Random Forest and Apriori models for market basket analysis.

PROJECTS

Generative Chatbot: Empathetic Multi-topic Conversational AI | GPT2, LLMs, NLP, GCP, API, JavaScript

- Engineered a Generative chatbot utilizing GPT2, RoBerta, Apache Solr on 200K reddit posts, understanding and reflecting emotions and factual accuracy in responses while switching topics, achieving 73% accuracy.
- Optimized model performance by finetuning GPT-2, BERT model using LoRA-PEFT with Named Entity Recognition, Sentiment and query optimization. Deployed Chatbot as restAPI on **GCP**, 15+ turns of coherent responses to users.

Educational Web App to Visualize RL Methods | AWS, RL, Gym, Pytorch, Flask, Python, REST API

- Deployed Flask app on AWS EC2 to enable students to visualize and interact with RL algorithms, enhancing their understanding of complex RL concepts. Solved multi-agent env using DQN, Double DQN, Actor-Critic.

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)