

# 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: **3.92/4.00**

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

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

**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 Engineer Intern, JerseySTEM | New Jersey

Mar 2024 - Present

- Created a system to transfer files from Google Drive using API, authenticate users, and upload files to an SFTP server.
- Developed and optimized PL/SQL packages for data processing and storage.

### Software Development Engineer, Persistent Systems | Pune, India

May 2021 - July 2022

- Developed multiple microservices, driving website traffic analysis 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 (Secure Data Lake) using Python, Elasticsearch, Kafka and Docker, achieving 97% data integrity rate.
- Collaborated with multiple teams to simulate security threats and evaluating multiple edge cases, resulting in a 20% improvement in threat mitigation strategies.
- Built a python ETL pipeline handling 40 million logs daily using Apache Kafka and Elasticsearch API.
- Enhanced system performance by implementing parallel processing with Dask, resulting in 15% faster asynchronous data generation across 15000 user.
- Optimized the open-source Wazuh agent-simulator with 6 new features scaling it's functionalities and robustness.

### Software Intern (Data and python), Verzeo | Pune, India

Dec 2020 - Mar 2021

- Performed data extraction, 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, LoRA, NLP, GCP, API

- Spearheaded development of generative chatbot utilizing GPT2, RoBERTa and information retrieval techniques on 200K reddit posts to efficiently understand emotions.
- Fine-tuning GPT-2 model with NER Tags and Sentiment Analysis using LoRA for empathetic responses, BERT and RoBERTa for factual accuracy. Deployed flask Chatbot on Google Cloud Platform, 15+ turns of coherent responses.

### 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 for a system consisting of five nodes.

### Educational Web App to Visualize RL Methods | AWS, RL, Gym, Pytorch, Python

- Developed and deployed interactive Flask app on AWS EC2 for RL course to visualize intricate RL algorithms.

### 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)