Savinay Shukla

917-815-8731 | ss16924@nyu.edu | linkedin.com/in/savinayshukla | savinayshukla.github.io

EDUCATION

New York University   Brooklyn, NY
Master of Science in Computer Engineering   *Sep 2022 – May 2024*

Manipal University Jaipur   Jaipur, India
Bachelor of Technology in Information Technology   *Jul 2015 – Jul 2019*

EXPERIENCE

Graduate Course Assistant   New York City, NY
NYU Center for Data Science   *Jan 2023 – Present*
- Leading the charge in the deployment of sophisticated deep learning systems in distributed settings, leveraging **High-Performance Computing (HPCs)** with a focus on **Large Language Models (LLMs)**
- Mentoring a group of 40 graduate students in the practical application of **TensorFlow and PyTorch** on HPC clusters.

Software Engineer   Bengaluru, India
IBM   *Dec 2019 – Aug 2022*
- Led successful migration of legacy Apache Struts 2 web application to **Spring microservices**, enhancing scalability and performance
- Leveraged Java 8, Hibernate, and Oracle Database to create a robust and efficient back-end for the microservices
- Spearheaded code reviews and QA within an agile and test-driven development environment, elevating overall code quality by 40%

Data Science Intern   Pune, India
Morning Blaze Pvt. Ltd.   *Feb 2019 – Jul 2019*
- Engineered an **ETL pipeline** for extracting market indicators (commodities, forex, global markets), enhancing predictive modeling for BSE automotive stock opening prices
- Implemented **lightweight time-series models**, integrating market and technical indicators to boost accuracy
- Conducted comprehensive research on historical data, ensuring incorporation of extensive indicators to effectively capture stock market downturns
- Optimized the engine across 300+ BSE companies, resulting in an impressive **5% reduction** in prediction losses

CERTIFICATIONS AND BADGES

IBM Data Science Professional | *Authorized by IBM*   Dec 2021
IBM AI Associate | *Authorized by IBM*   Aug 2021
- Demonstrated proficiency in **data science**, utilizing **Jupyter notebooks, Python, SQL**, and relevant libraries for solving real-world problems

PROJECTS

Charlotte | *Ollama, Langchain, Python*   Feb 2023 – Present
- Developed an innovative personal assistant utilizing **Langchain** and **open-source LLMs** like Mistral and Llama 2 to optimize productivity workflows
- Engineered a speech-to-text input pipeline seamlessly integrated with Gmail API for efficient email management and Slack API for instant notifications

Distributed Dual-Discriminator GANs | *Pytorch, Generative Models, HPC*   Apr 2023 – May 2023
- Optimized DCGAN training pipeline by introducing an extra discriminator for faster convergence
- Realized a **40% reduction** in time for optimal FID and IS Scores across CIFAR, MNIST, and SVHN datasets
- Implemented a **parameter-server architecture** for distributed, multi-GPU training to scale the prototype

TECHNICAL SKILLS

Languages: Java, Python, C/C++, SQL, NoSQL, MongoDB, JavaScript, TypeScript, HTML, CSS
Data Science : Data Analysis and Visualization, Statistics, Predictive Modeling, Feature Engineering, EDA, ETL
Machine Learning : Neural Networks, Computer Vision, Natural Language Processing, Time Series Analysis
Frameworks : Pytorch, TensorFlow, Azure, Docker, REST, AWS, GCP, Cassandra, Kafka, Git, Kubernetes