

Swapnil Sharma

+1-551-998-8215 | swapnil.sh@nyu.edu | linkedin.com/in/swapsharma | github.com/swappysh

EDUCATION

New York University

Master of science in Computer Science

- Machine Learning • Natural Language Processing • Natural Language Understanding • Computer Vision

New York, NY

Aug. 2021 – Ongoing

Indian Institute of Technology

Bachelor of Technology in Computer Science and Engineering

- Advanced Data Structures and Algorithms • Pattern Recognition • Deep Learning

Mandi, India

Aug. 2013 – May 2017

EXPERIENCE

Microsoft

Applied Data Scientist

May 2022 – August 2022

- Created a ground truth data-set for JavaScript unit test cases.
- **Prompt tuned OpenAI GPT3 Codex** to generate and evaluate test cases.
- Achieved **45%** high quality test cases on an average per input JavaScript file.

Nference

Senior Data Scientist | Data Scientist

May 2019 – August 2021

- Shaped a rule-based service to translate search queries to structured queries using **Named Entity Recognition, syntactic dependencies** and domain rules.
- Developed a highly available **hybrid (structured + unstructured) medical data** search analysis platform.
- Expanded **document signal extraction** and **signal ranking algorithm** using a mix of techniques including TF-IDF, sigmoid, PMI, sampling.
- Investigated techniques for **user intent identification** and devised a **BERT** based product recommendation system.

Directi

Platform Developer | Zeta – Payments And Accounting Team

August 2017 - May 2019

- Maintained highly consistent and available, multi-tenant payments system (~ **15m TPD**)
- Improved end-to-end **transaction latency** across payment services (~**5x**) via migrating multiple system levels operations to database level, introducing Redis based fallback mechanisms.
- Revamped entire payments system into the generic framework as a plug-and-play product for banking processes.

Flipkart

Software Development Intern

June 2016 – August 2016

- Upgraded **Apache Kafka** and **Apache Storm** keeping in view the requirement of the promotions division.
- Introduced **Batch processing** and **dependent Kafka consumer support** to Kafka spouts

Google Summer of Code

Student Software Engineer

May 2016 – August 2016

- Added Git support for large binary file version control and automatic uploading using git-annex as backend.

PROJECTS

Stereotypical biases in BERT models

Academic Project

Feb 2022 - May 2022

- Evaluating newer BERT models (Albert, CogLTX, Electra) for stereotypical biases on StereoSet Benchmark.

Automated Discrimination of Monocentric and Dicentric Chromosomes

Major Technical Project

August 2016 - May 2017

- Investigated different approaches to solve **chromosome discrimination problem**.
- Achieved **accuracy up to 70%** under low image resolution and limited data volume.

TECHNICAL SKILLS

Languages: Java, Python, C/C++, SQL (Postgres), Golang

Frameworks: Apache Storm, Apache Kafka, Zookeeper

Developer Tools: Git, Docker, Jupyter, Jenkins, Kibana, Grafana, Phabricator, Hadoop, AWS, L^AT_EX

Libraries: pandas, NumPy, Matplotlib, TensorFlow, Keras, PyTorch