Jinam Shah

(919)922-8481| jbshah@ncsu.edu | github.com/jinamshah | linkedin.com/in/jinamshah

EDUCATION

North Carolina State University, Raleigh, NC GPA: 4.00/4.00

Master of Computer Science, August 2021 – May 2023

Relevant Courses: Artificial Intelligence-1, Neural Networks, Automated Learning and Data Analysis, Design and Analysis of Algorithms, Software Engineering, Algorithms for Data Guided Business Intelligence.

SKILLS

- **Programming languages and tools**: Python, PyTorch, Tensorflow, PySpark, Keras, Pandas, Django, Flask, Spark, SQL, Git, C, C++, JavaScript, PHP.
- **Domain expertise**: Machine Learning, Deep Learning, Natural Language Processing, Image Processing, Distributed Training, Big Data.
- **Cloud technologies**: AWS, GCP, EC2, ECS, S3, Lambda, RedShift, Kinesis, API Gateway, CloudFormation, etc.
- **Design Principles**: Cost-effective architecture, Secure and Reliable Infrastructure Design, Serverless Architecture.

WORK EXPERIENCE

Machine Learning Intern, Cactus Communications, Princeton, USA May 2022 –Present

- Working with the director of data engineering on various big data problems and pattern recognition problems based on the company's datasets.
- Own the Machine learning track for the Name Disambiguation project.

Senior Software Engineer, Cactus Communications, Mumbai, India June 2020 – July 2021

- Lead the architecture planning and implementation of various products in the domain of ML/NLP and BigData, bridging the gap between the business and tech side.
- Created end-to-end machine learning pipelines in the production environment using cloud infrastructure in a fault-tolerant and cost-effective manner.
- Designed and implemented a BigData platform that ingests over 1.5Tb/week and generates over 4.5Tb/week. It manages over ~250Tb of data in the data lake.

PROJECTS

Document classification

- Ensemble of DL and ML models (based on BERT) performing document classification for over an unprecedented 1500 classes deployed using truly serverless architecture.
- Saved the organization around $1M per annum and reduced the TAT for the service from 8 hours to under 2 minutes (down by 99.6%).

Image recognition

- Image recognition software for determining ethical compliance of images added in research papers.
- Achieved state-of-the-art performance (99.8% accuracy) and deployed using serverless architecture.

Automated language correction

- Worked on the software engineering and "explainable AI" portion behind an NLP software that focused on automated grammar correction.
- This product led the efforts of the business to create a new vertical for generating revenue of around $1M.

ACTIVITIES

- **Courses**: Course series on Machine learning by Andrew Ng, Udacity course series on ML and Ethical AI.
- **Hackathons**: Top 25 percentile in Reply Code hackathon, consistent open-source contributor through HacktoberFest
- **Volunteer**: AI4Good foundation, working on solving UN's SDG.