Karan **Shah**

🏠 Stony Brook, NY 11790 | 📱(+1) 631-438-6950 | ✉ kvshah@cs.stonybrook.edu | 🌐 the-karanshah.github.io | 💻 the-karanshah | in the-karanshah

Education

Stony Brook University *Stony Brook, NY*

MS in Computer Science | GPA : 3.68/4 *Aug. 2019 - Dec. 2020*

Coursework: Theory of Database Systems, Big Data Analytics, Machine Learning, Computer Vision, Data Science, Probability and Statistics

Gujarat Technological University *Ahmedabad, India*

B.E. in Computer Engineering | GPA : 8.03/10 *Aug. 2014 - May 2018*

Coursework : Data Structures, Algorithms, Software Engineering, Object Oriented Programming, Distributed Operating System

Work Experience

Software Developer Intern *Sept. 2020 - Dec. 2020*

Quantiphi *Marlborough, MA*

- Building software tools in **Python and React.js** with Machine Learning Engineers to accelerate production-grade Conversational AI use cases in **Docker Container** over **GCP** cluster
- Using **Flask** based routing methods with **multiprocessing** to trigger underlying capabilities of **NVIDIA**'s end-to-end ML framework and **Socket+GRPC** for live audio transcription

Software Engineer *June 2018 - April 2019*

Simform *Ahmedabad, India*

- Developed an end to end **Video Streaming system**; allows user to upload and scan the content on the cloud, stream on cross-platform desktop application over globally separated clients; using **AWS services (Lambda, S3, Cloudfront, MQTT), Electron.js, and React.js**
- Built a software system to semi-automate Company's marketing work with Chrome Extension and **REST** API Web backend using **Python and Node.js**, efficient in reducing ∼ **40 %** of Marketing team's manual work
- Designed and automated **ETL** jobs to process a huge realtime data using AWS services (Glue, Lambda, S3) and **Spark**

Research Intern, Machine Learning *July 2017 - May 2018*

ISRO - Indian Space Research Organization *Ahmedabad, India*

- Designed generalized **CNN** architecture to handle multimodal and/or multispectral imagery along with adaptablity to varying application requirements; trained on Potsdam benchmark dataset producing State of the art results to **Semantic Segmentation**
- Built software tools in **Python** for preprocessing of geographic remote sensing data, visualization, real-time performance evaluation, post-processing and output generation as a semi-automation approach
- Manipulated raster GIS data as a part of the pre-processing task using Python and QGIS application

Skills

Programming	Python, JAVA, Javascript, C/C++, Matlab
Cloud/Big Data	AWS, Google Cloud Platform, MapReduce, Spark
Tools/Technologies	Docker, Kubernetes, Git, REST, GRPC, SQL, NoSQL
Frameworks/Libraries	Node.js, Django, Flask, React.js, PyTorch, Tensorflow, Keras, OpenCV

Publication

Band-wise Independent Pansharpening Using Neural Networks with Shared Weights

Pankaj Bodani, Karan Shah, Shashikant Sharma (ISRO) Submitted to **Geocarto International Journal**

Projects

Medical Search Engine, RadSearch (Java, Python, Javascript, Elastic Search) *Jan. 2020 - May 2020*

- Created a distributed, multitenant-capable full-text search engine with an HTTP web interface and schema-free JSON documents using **Elastic Search** to handle user queries and **Java Servlets, Python, Javascript** to handle the backend
- Currently being used regularly by the doctors at the Stony Brook Hospital to find the most relevant information about patients

Company's Internal Communication System (Java, Javascript) *Sept. 2018 - Dec. 2018*

- Created a web-based software for employees of a Software Company to send/receive messages, share files in person, groups and broadcast channels with RBAC - Roll based access control
- Utilised **SQLite** to store data of each user with **Java** Servlets to query the database and applied Software Engineering approaches

Retail Sales Data Analysis (Python, Pandas) *Oct. 2019 - Dec. 2019*

- Analyzed the market basket dataset to provide sales insight for a chain of retail stores.
- Identified reasons for sales drop with the help of external datasets
- Implemented **Word2Vec and LSTM model with k-means clustering** to improve Shelf Space Management and Inventory management

Video Action Classification - Recognition (Python, PyTorch, Matlab) *Nov. 2019 - Dec. 2019*

- Trained CNN for human action recognition task on the UCF101 data
- Achieved **86%** (1st in class) on image test data, and **84%** (2nd in class) on video frames test data
- Trained LSTM in RNN to actions classification on data collected by Kinect v2
- Used Transfer Learning to compute features for 60000 video frames with limited compute resources.

Desktop Application to Synchronize Local Storage with Cloud (AWS, Node.js, Electron.js) *March 2017 - Apr. 2017*

- Built cross-platform application leveraging the Electron.js framework
- Used **Node.js and AWS services** to fully synchronize local storage and AWS S3 storage