

Education

Columbia University

Master of Science in Computer Science

New York, NY

Aug. 2018 – Dec. 2019

Relevant Courses: Machine Learning, Deep Learning, Computer Vision, NLP, Algorithms, Cloud Computing

Vellore Institute of Technology

Bachelor of Technology in Computer Science

Vellore, India

Jul. 2013 – May. 2017

Relevant Courses: Probability, Statistics, Artificial Intelligence, Data Mining, Database Systems, Data Structures

Experience

Tempus Labs (on contract through Egen Solutions)

Software Developer

Chicago, IL

Mar 2020 - Present

Project 1 - Completeness Application

- Implemented an end-to-end data pipeline to process Tempus's clinical and molecular data, and determine the steps required to finish client deliveries, saving the team 7-10 hours of manual investigation work per week.
- Designed Splunk dashboards to track and monitor the progress of client deliveries in real-time and provide an accurate estimate of the time needed for the completion of deliveries.
- Key technologies used - Docker, Airflow, AWS (S3, Lambda, Cloudwatch, Redshift), Splunk

Project 2 - Backfill Application

- Led a team of 3 to design and develop an application to trigger the backfill actions required to generate and ingest the clinical and molecular data required by clients into database.
- Wrote automation scripts using Python to read data from S3, process it, and trigger the data pipelines to generate the missing data. The automation reduced turnaround time of client deliveries by more than 60%.
- Key technologies used - Python, GCP (Cloud Composer, Cloud Storage), AWS (S3, Batch)

Indian Institute of Technology - Delhi

Machine Learning Intern

Delhi, India

Jul 2017 - Jan 2018

- Proposed a deep learning solution for bone age assessment using Tensorflow and Python.
- Built a preprocessing module to extract a region of interest from hand radiology images using several thresholding and region-growing image segmentation techniques.
- Created a deep learning model with an accuracy of 89% using transfer learning on the VGG16 CNN model.
- Key technologies used - Python, Tensorflow, Matlab

Projects

Detection of Cancer Metastases in Pathology Images

- Designed a custom multi-input deep neural network for detection of breast cancer tumors present in pathology images using Tensorflow and Python.
- Used transfer learning to train the model on Google Cloud Platform and achieved 98% test accuracy.

Virtual Grocery Store Assistant

- Developed a cloud-based mobile application that analyzes a user's buying patterns to recommend groceries.
- Created a custom RESTful API using Amazon AWS services such as Rekognition, Lambda, Cognito, S3, and Lex.

Hybrid Techniques in Text Mining and Analysis of Social Networks Media Data ([link](#))

- Contributed a book chapter in a Springer Journal providing a detailed analysis of all the hybrid and traditional text mining techniques used for the analysis of social media and network data.

Face Recognition using Neural Networks

- Developed a hybrid system for face recognition using artificial neural networks on MATLAB with 95% test accuracy.
- Used PCA algorithm in conjunction with Neural Networks to reduce redundancy in input data resulting in higher accuracy.

Technical Skills

Domain and Interests: Machine Learning, Deep Learning, Computer Vision, Cloud Computing

Languages and Frameworks: Python, Tensorflow, SQL, C, C++, Java, MATLAB, AWS, GCP, Airflow, Git