**Suman Kalapatapu**

12805 Shaker Blvd, Cleveland, OH 44120|Open to relocation

suman.kalapatapu@gmail.com | (513) 307 - 6877 |

<https://www.linkedin.com/in/skin/>| <https://github.com/suman500bn>| https://suman-kalapatapu.netlify.app/

**Education**

**University of Cincinnati** **Fall 2019 – Dec 2020**

Master's in Computer Engineering, GPA: 3.7

Graduate Coursework in Intelligent Data Analysis, Machine learning, Big data analytics, Cloud computing, Deep Learning

**CVR College of Engineering** **Fall 2012 - May 2016**

Bachelor's in Electronics and Communication Engineering GPA: 3.8

Undergraduate Coursework in Object Oriented Programming, Operating Systems, Algorithms and Data Structures

**Areas of Expertise**

|  |  |
| --- | --- |
| Programming Languages | Python, Java, JavaScript, Linux shell, SQL |
| Cloud Services | AWS EC2, Lambda, S3, Docker, Azure, CloudFront, API Gateway, CloudWatch, VPC |
| Framework and Database | Node.js, Django, Flask, Serverless, PySpark, Hadoop, MySQL, Postgres, MongoDB |
| Machine Learning Libraries | Pandas, NumPy, Scikit-Learn, Seaborn, Matplotlib, PyTorch, NLTK |

**Backend Developer and ML Engineer Y apps Lab May 2020**

* Designed scalable social media app with serverless architecture using Node.js, Mongo for backend and Redis for caching which can handle high load of concurrent users.
* Developed ML algorithm to generate newsfeed based on personalized interests and a provision to insert advertisements.
* Developed payment system to batch load financial transactions in Stripe and PayPal with distributed load on system.
* Utilized FFMPEG to encode videos into HLS format which reduced 90% data consumption and created parallel processing of deep learning model to filter explicit content with 90% accuracy.
* Secured connection to database through VPC peering creating private subnet with NAT gateway and public subnet with internet gateway and placing lambda functions in private subnet.

**Data Science Internship Cincinnati’s Children Hospital Jan 2020 - Jun 2020**

* Implemented POC to find impact of chemotherapy in cancer patients at different phases using ML
* Visualized survival analysis of patients having different side effects through Cox Proportional Hazard in R
* Created a novel approach to predict survival for different patients (age, gender, ethnicity) from initial phases which can help physicians to change treatment procedure.

**Software Development Analyst NTT DATA Global Delivery Services Ltd. Oct 2016 – Jul 2019**

* Created Django REST API for social authentication, generated JWT tokens, multiform data upload with presigned urls and email verification using message queues.
* Integrated application with S3, API gateway, CDN with 60% reduced load time and load balancer to distribute the load.
* Created search autocomplete API by creating search index in database which resulted in 70% reduced querying time.
* Built, tested, and maintained ETL data pipelines integrating data from multiple servers into a cohesive data warehouse and other repositories for reporting and analysis.

**Projects**

* **Recommender system**: Built 5 types of recommender system based on popularity, classification, collaborative filtering, content filtering approach using Pandas, Scikit-learn on movie, bank, restaurant, housing datasets.
* **Hosted Docker application in Cloud**: Built CRUD application with upload and download functionality using Python Flask, MySQL for Backend and HTML, CSS for front end, containerized application using Docker and deployed it in AWS EC2.
* **Weather Data analysis**: Cleaned and collated TBs of datasets using SparkQL, Python and developed solutions for efficient analysis of weather data of Cincinnati region for last 10 years on Hadoop cluster.
* **Crime data analysis using Azure Databricks:** Explored parquet files data using optimized PySpark queries and visualized through Tableau integration. Extracted time Series data and forecasted crime count through ARIMA model

**Certification**

* Microsoft Azure AZ900 certification | Deep Learning Specialization by Nvidia| PySpark Certification from Datacamp

**Awards**

* Spot Award of excellence in one year at NTT Data and got promoted to grade 6 within one and half year of joining.