SAI SIVARAM SAKETH VOOTLA

vootla.s@northeastern.edu | https://github.com/SakethSai31 | linkedin.com/in/saketh-vootla/ | (+1) 716-550-3721

EDUCATION

Northeastern University	Master of Science in **Information Systems**	May 2023
SRM University	Bachelor of Technology in **Electronics and Communication**	May 2017

TECHNICAL SKILLS

Programming Skills	**Python**, **JAVA**, C/C++, R, **SQL**, **Unix, Linux**, Shell Scripting, JavaScript, HTML, CSS
Databases	Oracle, DynamoDB, MongoDB, Hive, MySQL, Redshift, PostgreSQL, RDS, S3, NoSQL
Cloud Technologies	**AWS** (EC2, SES, DNS, RDS, Lambda, Cloudformation, API, IAM, CloudWatch, ECR), **GCP** (Big Query, Data Studio, Dataflow, DataLab, Cloud Functions, Pub/ Sub, Composer)
Tools	MS Excel, Power BI, Postman, Tableau, Git, CI/CD, Airflow, Docker, JIRA, Kubernetes
Concepts	Data Modelling, DevOps, Statistical Analysis, Machine Learning, Software Development

WORK EXPERIENCE

AI-Research Associate | ABECEDARIAN   June 2023 – Present

- Led a research project leveraging prompt engineering techniques and large language models, resulting in a **30%** improvement in image generation accuracy compared to baseline models. Implemented a novel approach combining ChatGPT and Google BARD, which achieved an average image quality score of **0.85.**
- Collaborated with Midjourney AI to develop a custom image generation pipeline, optimizing the process, and reducing the average generation time by **40%.** Implemented parallel processing techniques, enabling the system to generate **100** high-resolution images per hour, resulting in a **50%** increase in overall productivity compared to previous methods.

Software Development Engineer | Vodafone   Aug 2017 – Feb 2020

- Implemented large-scale data warehouse with data from over 10 sources, 1 **TB** data storage with an average query response time of **1 sec**. Generated **500 insights**/ day leveraging GCP, resulting in improved decision making.
- Reduced response time by **36%** through the implementation of efficient NLP techniques and data management strategies resulting in enhanced user experience and improved service delivery.
- Successfully architected and executed a migration strategy of **15 TB** of transactional data from an on-premises system to GCP Cloud Bigtable, completing the migration in **5 days** while maintaining **99.89%** data completeness
- Designed scalable ETL pipelines using Python, SQL, and Apache Airflow that processed **10 M** records per day with an error rate of **0.3%** and achieved a **99.99%** uptime. Integrated data from over **10** sources
- Executed performance tuning to utilize maximum efficiency of various systems pertaining to BI, Java, ABAP, FICO, Supply Chain by optimizing processing times by **~40%** utilization.

ACADEMIC PROJECTS

AWS Application Development & Deployment – *AWS, MySQL, Flask API, Microservices, Serverless*   Aug 2022 – Dec 2022

- Built a RESTful API using Flask and MySQL on AWS, implementing a 3-tier architecture for high availability and scalability.
- Designed a Python-based email notification system and deployed it on AWS using GitHub Actions.
- Created AWS CloudFormation templates to provision and automate various cloud resources, such as EC2 instances, RDS instances, VPCs, subnets, and security groups.
- Introduced dynamic-scaling of Application Load Balancer (ALB) by utilizing CloudWatch to monitor EC2 instances' CPU and memory utilization and setting triggers for scale-in and scale-out policies. This allowed the infrastructure to handle high user traffic, resulting in the ability to handle up to **10k** concurrent requests.

AWS DevOps and Automation – *AWS, CI/CD, DevOps, Jenkins, GITHUB, SNS*   Aug 2022 – Dec 2022

- Implemented a CI/CD pipeline for the RESTful API and email notification system using AWS CodeDeploy and AWS CloudFormation, enabling streamlined and automated software deployment.
- Monitored and maintained the infrastructure using AWS CloudWatch, implementing real-time visibility into system performance and preventive incident management. Additionally, implemented error handling and logging mechanisms to quickly identify and resolve issues.
- Enforced **128-bit** Secure Socket Layer (SSL) to encrypt connections between the application and Amazon RDS DB

Social Media Analytics – *GCP, Apache Airflow, API, NLP, Docker, Data Visualization*   Jan 2022 – Apr 2022

- Developed an ingestion pipeline for processing streaming data from Twitter using Apache Beam, hosted on Google Cloud Dataflow, and storing the transformed data in BigQuery.
- Utilized Apache Airflow to orchestrate the pipeline for every **1** Hr. schedule using Google Cloud Composer. Store data consists of more than **27 M** tweets with a **0.2%** of error rate.
- Performed sentiment analysis on the tweets gathered and built an Intelligence system for all the tweets, leveraged AWS Lambda function to utilize NLP models as ML as a service on a serverless framework containerized the whole package with Docker. Resulting in real-time sentiment analysis of tweets with high performance and scalability