SAI SIVARAM SAKETH VOOTLA

vootla.s@northeastern.edu | https://github.com/SakethSai31 | https://sakethsai31.github.io/
linkedin.com/in/saketh-vootla/ • Boston, MA • (+1) 716-550-3721

EDUCATION

Northeastern University Master of Science in **Information Systems** May 2023
Coursework Includes: Network Structures and Cloud Computing, Big Data Systems and Intelligence Analytics, Database Management

SRM University, India 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 Gateway, IAM, CloudWatch), **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

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
- Collaborated with Business Intelligence team to optimize report generation through ad-hoc analysis using DIS, resulting in a **80%** time savings on report creation

ACADEMIC PROJECTS

RestAPI Cloud Application – *AWS, MySQL, Flask API, CI/CD, Microservices, Serverless* Aug 2022

- Constructed a 3-tier architecture for **RESTful API** services using web technologies like Python with MySQL on AWS through AWS CloudFormation on RDS instance as the backend and tested APIs using JUnit Testing with Jest framework
- Built an **AWS CloudFormation** template that provisions various cloud resources including VPC, Subnets, EBS, EC2, RDS, S3, ELB, SES, CloudWatch, Route53, ALB, IAM Roles, SES, SNS, Lambda functions, DynamoDB, and Security Groups, resulting in a streamlined and automated infrastructure deployment process.
- Successfully created and deployed a CI/CD pipeline with GitHub actions to automate software build and deployment with **99%** uptime, using Packer and Code Deploy. Unit Tests and Integration test were in place to ensure there is no stop in process
- Configured custom domain on AWS Route 53 DNS service using Application Load Balancer and performed load-testing on the endpoints to ensure high availability and performance
- Supervised systems by monitoring and maintenance of EC2 instances using AWS CloudWatch to collect metrics, analyse logs, and monitor system load, through setting thresholds and trigger alarms. This allows for real-time visibility into system performance and preventive incident management.
- 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.
- Enforced **128-bit** Secure Socket Layer (SSL) to encrypt connections between the application and Amazon RDS DB instance.

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