

EDUCATION

- **National Institute of Technology Karnataka, Surathkal** 2019
Karnataka, India *B.Tech, Mining Engineering- 8.25 CGPA*
- **Govt Science College-IGNOU** 2020-2021
Karnataka, India *PG, Pg Diploma in Applied Statistics - Pursuing*

EXPERIENCE

- **Quantiphi Analytics** Bangalore,India
Data Engineer *May 2019 - May 2021*
- **Qognition - AI:**
 - A UI based training and inference platform for different NLP tasks that has features for fine tuning of model on different datasets. Google AI Platform was used for model training and deployment.
 - A generic ETL pipeline for pre-processing is developed in Apache beam to create tf-records for training the ML model for text Summarization .
 - Created Flask API endpoint for the deployment of text summarization model on Docker for online inference.
- **Operational Data lake :**
 - Humongous real time data pipeline is built using pub-sub and data flow where data is stored across Neo4j for establishing relationships, in big query and timescale db for KPI Metrics.
 - To lower the operational costs, provisioned Timescale db across 3 nodes for high availability, faster query performance and almost achieved the performance of big-query using optimization techniques.
 - Timescale db is employed for back filling operations and Calculated KPI metrics are visualized in Looker
- **Computer Vision for Safety:**
 - Computer Vision pipeline is for safety in parks, resorts, etc. This involves generating real-time safety alerts using applications of computer vision on the live camera feed.
 - Built a custom solution for classification on top of Object Detection model based on SSD Algorithm followed by a rule-based approach to describe the cause for movement of the hand during movement of the passage.
 - Developed Flask API to trigger model deployed on Kubernetes engine to infer the results on video for action(stand/sit) with an average latency of 0.3s
- **DriveBoard(Product) :**
 - One-stop platform for all the DevOps activities which include Infrastructure provisioning, security checks, template generator, Git and Jira Project creation, etc are developed using Django as back end.
 - This product minimized the time of Infra provisioning by about 1 day, Security Checks while provisioning reduced attacks and real-time alerts for project helped in minimizing costs.
- **Cyber attack Prediction on websites :**
 - Kafka is employed in building near real time streaming pipeline on web logs and pre processed data is stored in HDFS for re-training.
 - The web logs are sent to model hosted on ECS and able to classify in 4-5 seconds.
 - A batch job is built on web logs where data from HDFS is further processed and sent to model for Prediction.

PROJECTS

- **Smart Study:** (On-going)
 - A UI based topic search platform where user can get information and newspaper editorial, up on user request,summary is provided for the given topic. A feature for query resolution has been provided with Elastic search, Hugging Face transformer as back end.
 - Vocabulary and Quiz at the end helps to gain and test knowledge respectively. This Product is hosted on Google compute engine in a Docker Environment.
- **Spark Streaming based sentiment analysis:** Streaming pipeline is built on twitter real time data with the help of Kafka and Spark Streaming and real time data is sent to spark Mlib Logistic regression classification model and result are tabulated.

PROGRAMMING SKILLS

- **Languages:** Python
- **Bigdata:** Pyspark,Hive,Hbase,Kafka
- **Cloud Data:** Dataflow,Bigquery,Bigtable,Elastic search,Datastore
- **Cloud Infrastructure:** Docker,Kubernetes,Terraform
- **Machine learning:** Sklearn,Tensorflow
- **Time series database:** Timescaledb,Opentsdb
- **Backend:** Django,Flask
- **Business Intelligence:** Tableau

CERTIFICATIONS

- Google Associate Cloud Engineer
- AWS Certified Developer Associate
- Microsoft Certified Azure Fundamentals

COURSES

- Applied Machine Learning Course - **Applied roots**
- Getting Started with AWS Machine Learning -**Coursera**
- Fundamentals of bigdata - **Databricks**
- Natural Language Processing With Transformers in Python - **Udemy**
- Foundations of Data science -**Guvi**
- PySpark Essentials for Data Scientists (Big Data + Python) -**Udemy**

COMMUNITY OUTREACH

- Community Member at Data talks a Data science Community forum
- Attendee :Application of AI in Health-care
- Core member for mining tech fest part of Engineer.2018..
- Coordinator in CRRESCENDO club for recreation.