DEVALLA JYOTHI SWAROOP

Data Engineering and Data Science
Work-Ex: January 2018 – Present

Email: jyothiswaroop4270@gmail.com
Phone: +917660066099

RESUME OBJECTIVE

When the world is moving with data, Here I am working on it,developing AI applications end-to-end, leveraging Machine Learning, Big Data Analytics and Cloud technologies to provide stake holders with crucial information and earn my place in this fast changing world. Delivering high standard output with strong sense of responsibility given me an edge. Looking forward to connect with you and share our Data Journey

SKILLS

Data Science	**:** Machine Learning, Deep learning, Natural Language, Computer Vision
Mlops	**:** Mlflow,Dvc,Azure sdk
Frameworks	**:** Scikit-Learn, Pytorch, TensorFlow 2/Keras, Nltk, Spacy
Big Data	**:** Hadoop,Hive, Presto, Hive, Airflow
Spark	**:** Spark,Sparksql,Sparkmllib,Spark Nlp, Spark Streaming
Streaming	**:** Kafka
Visualization	**:** Tableau, PowerBi
Etl	**:** Alteryx, Tableau prep
DevOps	**:** Git, Azure Repos, Azure Pipelines, Docker, Kubernetes, Jenkins
Query engines	**:** Presto, Apache druid
Sql	**:** Postgres, Postgres TDE, MySQL, Oracle;
NoSql	**:** Mongodb,Hbase,Cassandra
Time Series db	**:** Influx db
Graph DB	**:** Neo4j
Programming	**:** Python
Api	**:** Flask, FastApi, Streamlit
Azure	**:** VMs,Data Factory, DataLake Gen2, Databricks, Luis Framework, Azure ml, Synapse
AWS	**:** EC2, S3, EMR, Glue, Athena, Redshift, Lambda, Kinesis, Kinesis firehorse
Monitoring	**:** Prothemus, Grafana
Elk stack	**:** Elasticsearch, Logstash, Kibana

EXPERIENCE

Tata Consultancy Services India **Jan 2018-Present**

DATA SCIENCE AND DATA ENGINEERING (Full stack)

- The app features modern connected KPIs such as driver score, vehicle health, fuel efficiency score, Common origins, common destinations, routes, break points based on telematics data.
- Built Machine Learning models to power KPIs such as K-Prototype Clustering for generating Driver's personas, Different models for churn prediction, Association Rule Learning for subscription recommendation on spark MLlib to handle big data.
- Designed and Built Data Pipelines for data ingestion from Data Lakes using presto, for data Cleaning, transformations and mining through Spark jobs, and for storing analytics output in RDMS which powers the AI platform.
- The app is expected to be used by owners for vehicle maintenance and to improve driving skills, by OEMs for customer relationship management, by fleet managers for predictive maintenance and for training needs and by insurance companies to estimate vehicle insurance based on vehicle health.
- The application functions on hybrid platforms of both batch and streaming applications where Azure iot acts as pub-sub model and thereby received by spark streaming applications to perform Operations and followed by batch processes
- Used google and azure maps API for reverse geocoding purposes for poi data, Which thereby used for context based advertisements on to different requirements quoted
- Containerized all the applications using dockers and spinning off kubernetes cluster
- Orchestarized and scheduled the application pipeline with airflow using DAGs.
- Created flask API to all the applications and exposing the ports for mobile app and for other departments to carry forward operations
- Created application for data anonymization adhering to GDPR.
- Been a main part of data model, Data architecture and project discussions.

NATURAL LANGUAGE PROCESSING

- The app demonstrates the complaints produced at National Highway Traffic Safety Administration is Responsible for keeping people safe on America's roadways and helping reduce the complaints Generated on vendors at the forum
- Developed Named Entity Recognition with the custom dataset generated by scraping from the archives Of the twitter data, using the Distil-Bert model have created the taxonomy for the manufacturing domain
- Established the custom Relationship model using the Transformer model with pytorch framework making components of taxonomy build intra-relationship, therefore making the vendor have a quick glance and reducing the timeframe to know about the complaint case and reaching the customer more quicker.
- Building the knowledge graph topping to relationship model thereby giving a quick real time glance Of the customers, complaints, taxonomy there by helping to created ml models to history data created by it
- Build a multi-class classification model on Roberta with pytorch, for classifying the tweets on to different Categories again making the vendor to know more about the customer's opinions thereby which reducing Gap between vendor and customer.

DATA SCIENCE

- Built ML and DL models according to client's requirements on HR analysis, Sales analysis, Maintenance, Public relations and operations.
- Built chat bots using Azure cognitive services according to the client's requirement.
- Creating custom intent and entity recognition and routing the unknown intents and entities to cosmos db and making it to retrain On the model developed by luis framework
- Have used python and flask(micro web framework) as an API.
- Have used Luis framework and Qn/A maker for chatbot.
- Have used different transformers for building the document summarization.
- Have used azure elastic search to decrease the latency.

ADVANCE ANALYTICS (BI)

- We have built dashboards in tableau and power bi from different sources ofdata.
- Have used all the advanced analytics such as prescriptive and predictiveforms.
- Have integrated the tableau and powerbi using pycaret for the advanced analytics usingpython.
- Different forecasting analyses have been done to give the client a better reach on thedata.
- Have used R-programming for the statistical purpose for all the feature engineeringconcepts.
- Have used different custom visualizations on tableau and powerbi.
- Published reports in tableau servers and powerbi-apps.
- For Data preprocessing we have used tableau-prep and python for all the transformationactivities.

EDUCATION

SCHOOLING:

SAINIK SCHOOL KORUKONDA(2015-2012)

B.TECH:

J.N.T.U. KAKINADA (2013-2017)