

Current Role :

Technical Lead Machine Learning

Skills:

| | |
|-------------|--|
| Software | Python,Pytest,FastAPI,Flask,Pandas,Numpy,Matplotlib,Spacy,Scikit-learn,TensorFlow,Keras,MachineLearning,DeepLearning(DNN,CNN,MaskRCNN,Yolov5),NLP,MLOps,Pyspark,Prefect, Neo4j, Java/J2ee, Oracle ATG, Spring, UNIX, SQL |
| Tools/Cloud | AWS(EC2,S3,ECR,CodeCommit,Sagemaker),Docker,DockerCompose,MLFlow,VSCode,Nginx,WSGIserver,JupyterNotebook,SoapUI,Postman,Eclipse,Weblogic, Websphere, Jenkins, Maven |
| Domain | Life Sciences and Healthcare, Banking and Financial Services ,Telecom |

Work Experience :

1. TCS Life Sciences AI COE (June 2020 – Present), Role : Technical Lead ML

Working as technical lead in machine learning for the Life sciences North America AI COE, managing a team of 8 associates. Work involves transforming business use cases to technical requirements in machine learning, designing and implementing solutions with python based ML frameworks using open source softwares(OSS) for the development of POC and projects. Assessing the various open sources ML models with permissible licenses to be used in products and solutions. Integration of ML models in the web application.

Experiment tracking for reproducibility. Deploying and monitoring models in production. Retraining of models.

Assisting teams responding to various RFX(Request For Information, Request For Proposal etc) from customers in the Machine Learning space with technical design.

Life Sciences Gen AI with Large Language Model :

- *Tech Stack:* Python,Pandas, HuggingFace, OSS, FastAPI, Rest API, AWS
- *Objective:* The various use cases involve generating auto narratives using data fields from pharmacovigilance safety database, generating summary of daily safety updates .
- Analyzing the requirement and selecting the open source Large Language Model.
- Data source identification and collection/creation.
- Prompt engineering by creating prompts and prompt templates.
- Generative configuration.
- Finetuning approaches for the use cases .

Comparison of marketed drugs information document :

- *Tech Stack:* Python,Pandas, Spacy ,HuggingFace, Scispacy , BERT ,Grammarly Gector ,AllenNLP ,Flask, Rest API, XGBoost, ReactJS,Nginx,Docker,AWS
- *Objective:* The objective is to compare the prescribing information documents(labels) of marketed drugs and find the deviations among the content in labels for the different use cases.
- Grammatical error detection with the OSS model from grammarly.
- Unit testing of the modules.
- MLFlow for experiment tracking.
- Custom NER model for the Adverse drug reactions and System Organ Classification.
- HuggingFace hub for the various open source nlp models.
- Rest api for the individual models endpoint.
- AWS ECR used as a container registry.

- Packaging and deployment with docker images and docker compose.
- Monitoring with evidentlyai.

Inappropriate objects in marketing documents:

- *Tech Stack:* NLP,OSS Model- Yolov5, MaskRCNN ,Linux, Python, Pandas, Numpy, seaborn,Keras, VGG Image Annotator, Flask , ReactJS,AWS,Docker,Nginx
- Objective: The objective is to assist the reviewer by identifying the presence of objects which are deemed inappropriate as per the health authorities and marketing guidelines.
- OSS Model Yolov5 was used initially for the detection of objects. Another OSS model we used was the MaskRCNN.
- Annotation of the data using VGG Image Annotator.

2. Tata Innovation Genomics Labs (September 2017 – May 2020), Role : Data Scientist

Worked on development of PoC's and deployment for the various offerings from the Tata Innovation Genomics Labs. Rule based system for supporting decision making for patients ,semantic mapping of clinical terms, mortality prediction for outpatients.

Exploratory data analysis of the Genomic data generated from the lab using python libraries such as pandas ,numpy ,matplotlib .Deep learning based natural language processing system for extraction of terms from the free text data . Image classification .

Extraction of Adverse Drug reaction(ADR) from text :

- *Tech Stack:* NLTK, Spacy, SciSpacy, Unix, Python, Pandas
- Objective: Identify and extract the ADR from the text corpus.

Relational data to graph database POC:

- *Tech Stack:* Neo4j, Cypher query, Unix,
- Objective: Identify indirect relationships present in the data by modeling the data in graphDB.

Mortality prediction with OHDSI CDM data model (Hosted by UW):

- *Tech Stack:* Python,Scikit-learn, Pandas, numpy, Keras, Unix, Jupyter
- Objective: Predict the mortality of the patient within the next 6 months after discharge from hospitals.

Digital Medical Assistant():

- *Tech Stack:* Drools, Spring ,RestAPI, Weka .
- Objective: To develop an application for implementing a rule based clinical decision support system for HCP'S for diagnosis of breast cancer. It assisted the HCP's in diagnosing the patients effectively.

3. TCS - ABN Amro Bank (June 2016 – Aug 2017), Role: Software Developer:

- *Tech Stack:* Java, Spring, RestAPI, Tibco, Python, Unix, SQL
- Worked on TIBCO Spotfire to develop, support, test and analyze the request data being obtained in the live form.
- Worked on TIBCO BPM, anomaly detection in the business process.
- Prepared visualizations for the customer use cases.
- Development and support of restful api for other modules.

4. TCS - Vodafone (March 2013- May 2016), Role: Software Developer:

- *Tech Stack:* Java, Oracle ATG, Hibernate,SQL, Oracle DB, Google Analytics
- Worked on developing solutions using Java framework for the sale of various tariff cards ,live tracking of the service requests being used in the various departments of sales and marketing in the Vodafone. Lead data collection scenario to address various queries by the users across different regions within India.
- Worked on Google and Adobe Analytics to analyze the performances.

Certifications:

- [Neural networks and Deep Learning: Coursera](#)
- [Improving Deep Neural Networks: Coursera](#)
- [Structuring Machine Learning Projects: Coursera](#)
- [Convolutional Neural Networks: Coursera](#)
- [Intro to ML in production: Coursera](#)
- NLP with Deep Learning in python: Udemy

Education: BE(Computer Science And Engineering) , CGPA: 8.11 , VTU Belgaum,Karnataka