

**LANKIREDDY SITA RAMI REDDY**  
8008375144; [sitaramireddy1994@gmail.com](mailto:sitaramireddy1994@gmail.com); Hyderabad

---

## PROFILE SUMMARY

---

- Working as Senior Machine Learning Engineer in a product based company with more than 6.5 years experience with strong engineering skills, focusing on realtime retrainable decision agents across industries using Microservice Architecture, container orchestration, computer vision, Mlops, large data handling, NLP, image processing and image text extraction, while exceeding expectations of client and providing deliverables in time.

## ACADEMIC CREDENTIALS

---

- **Msc in Data Science** | LJMU & UpGrad | 63% Oct 2021- Jun 2022
- **PG Diploma in Data Analytics** | IIIT-Bangalore & UpGrad | CGPA 3.35 / 4 May 2017- July 2018 A comprehensive 11-month program taught by Industry experts and IIITB faculty; 7 case studies & projects; 400+ hours of academic learning & 30+ hours of industry mentoring  
Consistently scored >80<sup>th</sup> percentile on all projects for data management, statistics and predictive analytics courses
- **B. tech, Electronics and Communication engineering** | CGPA 9.3/10 2016  
k L University, Vijayawada
- **Intermediate** | State Board- Andhra Pradesh with 96% 2012
- **10th** | State Board- Andhra Pradesh with 90.5% 2010

## Msc Thesis:

---

**Light weight U-Net model for Retinal Blood Vessel Segmentation:** This study includes about various image segmentation methods, importance of retinal vessel segmentation, how deep convolutional neural networks are used for vessel segmentation. Implemented the U-net variant for vessel segmentation by experimenting on 3 different loss functions with 2 databases(CHASE and DRIVE), and achieved state of art results with 0.9675 accuracy, 0.769 sensitivity and 0.9865 specificity.

## PG DIPLOMA CASE STUDIES

---

**HR Analytics Model:** Predicted employee likely to attrition with 85% accuracy by analysing 4400+ employees' data with logistic regression model.

**Car price prediction:** Predicted price of the car with 88% accuracy by analysing 25+ variables of each car with linear regression model.

**Uber supply demand gap:** Predicted the root cause for supply demand gap of uber cars and suggested ideas to improve business.

**NYC Parking Data Analysis:** Analysed the 6 GB of Car parking data using SparkR and find out the Insights that helps for better parking.

**Market Mix Modelling:** Analysed the E-Commerce data of different product categories with different advertisement channels by building MMM and suggested the best investment channels for the company to get high ROI.

## WORK EXPERIENCE

---

**Senior Machine Learning Engineer, Exponential AI, Hyderabad**

March 2019-present

- Trained and Deployed **hugging face question answering pipeline** for one of the client.
- Build **NLP microservice** to tag the extracted elements to domain object. Used **Spacy, Keybert** and **zero shot text classification** to extract entities from the text. Also implemented **ontology** service using **Elasticsearch**. Implemented **Negation** concept using Spacy.
- Involved in development of **ML model versioning**, to track the experiments and to ease the deployment of ML models for the solution engineers.
- Trained and Deployed **Faster-RCNN** model for handwriting detection and table detection.
- Implemented **histogram method** to detect rows and columns on non-line table.
- Developed and Deployed **GAN** model to draw column lines for non line table.
- Designed and developed versioning of ML custom functions for platform, by using **git, DVC** and **Minio**.
- Worked with OCR-free **donut** model for information extraction from the document images.
- Designed and developed **assisted extraction** feature for the platform, to retrain the ML models as data changes.

- Designed and developed **feedback loops for ML models**, based on the human review given by the business user.
- Designed and developed **benchmark service**, to get the accuracy, precision and recall for all the elements of the benchmark dataset.
- Designed and developed **data augmentation** tool to generate the **synthetic data** for training the models, and integrated this to platform's annotation tool.
- Trained and deployed the **LayoutLMv2** model as a custom function for one of the client, to extract information from the document.
- Resolved **memory leakage issues** in tensorflow models and increased the performance of platform.
- Did **Memory profiling** for complete document extraction pipeline and improved the performance.
- Migrated tensorflow code and models from version 1.14 to version 2.9.
- Experimented **Tesseract OCR** extraction with different PSM modes and verified the hocr accuracy.
- Worked on **Drug sales data** to find the duplicate and triplicate rebates, there by to reduce the loss for drug manufacturers.
- Worked on **health insurance claims data** to score the providers based on the past data for different ICD and CPT codes.
- Worked on text document images, to extract elements like Headings, Paragraphs, Header Footer, Key-Value pairs, Tables, Bullet Lists and OMR by using **Machine learning, Deep Learning, Image processing** and **NLP** techniques.
- Improved **Heading** accuracy from 50% to 85%, by implementing **XGBoost** model.
- Trained and Deployed **Paragraph** Model with 73% accuracy by using **XGBoost** model.
- Trained and Deployed **Header Footer detection** Model with 70% precision and 65% recall.
- Implemented Bullet lists detection by using generic **template matching** method, and grouped the bullets using **Hierarchical clustering**.
- Implemented service tasks for all the models developed and integrated in the platform.
- Implemented Algorithm to **detect Bold word's** in an text document.
- Implemented **Import Export feature** for solutions of our platform.

**Systems Engineer, Infosys Limited, Hyderabad**

May 2016-Feb 2019

- Applied **Explainable** models by using **LIME** and **SHAP** for classification Model.
- Developed MNIST image classification model with CNN, by pre training with Autoencoder Model.
- Developed the text summarizing model using python and **NLTK**.
- Developed loan defaulter classification model with 78% accuracy.
- Implemented Explainable AI for **Neural Machine Translation** using **Layer Wise Relevant Propagation**.
- Developed Basic GAN to generate Digits.
- Categorize the tickets based on the description of tickets data, which can be later used for exploratory analysis.
- Analyze the trend in the tickets by using **volume analysis**.
- Improved data quality by using **Hierarchical clustering** for automate testing.

## KEY SKILLS

- **Machine Learning: Data cleaning**, EDA, Model selection, Model training, ML-ops, Model Deployment, Data Versioning, Data cleaning, Model versioning.
- **Databases:** MySQL , MongoDB
- **Big Data Technologies:** Hive, Spark
- **Deep learning:** ANN, CNN, RNN, F-RCNN, GAN, U-Net
- **Software Engineering Skills:** Microservice Architecture, Debugging in container pods, Elasticsearch, Python with Clean architecture, Caching, Memory profiling, Improving performance with respect to time and memory, Git, Docker, kubernetes.
- **Other technical skills:** AWS, Tensorflow, RabbitMQ, Minio, Redis, MongoDB.
- **Certifications:** 1. Deep learning certification by IITM.  
2. Introduction to NLP by Udemy  
3. NLP with Deep learning certification by Udemy.  
4. Deep learning for Visual computing certification by IITKGP