

B AKHIL KUMAR

Hyderabad, IN | +918977665957 | akhilkumarb1997@gmail.com | [Linkedin](#) | [Github](#) | [Portfolio](#)

SUMMARY

Data Scientist with 3 years of experience, creating cutting-edge solutions. Specialized in developing multi-language chatbots, Intelligent Video Analytics, and document parsing applications. Proficient in crafting explainable AI models and deploying end-to-end ML applications with Docker, Kubernetes, and MLflow. Skilled in leveraging OCR technology for efficient data extraction and implementing state-of-the-art and research based machine learning models for CRM and fraud detection. Strong expertise in CI/CD pipelines and a deep understanding of data science frameworks such as Pytorch, Scikit-learn, TensorFlow. United cross-functional teams and clearly translated technical concepts to business stakeholders, enabling data-driven decision making and driving the successful delivery of multiple AI initiatives.

EDUCATION

Mahaveer Institute of Science & Technology

Bachelor of Technology

Major in Electronics and Communication Engineering

Hyderabad, IN

September 2014 - October 2018

Siddhartha Junior College

Math, Physics, Chemistry

Hyderabad, IN

July 2012 - July 2014

EXPERIENCE

Pranathi Software Services

Data Scientist

Hyderabad, IN

September 2023 – Present

- Implemented a **multi-language chatbot** web application using Python, Flask, and JavaScript, enabling users to interact with the bot through both text and speech inputs.
- Designed and actualized “**Intelligent Video Analytics for Efficiency Tracking**” that can be used for Work efficiency tracking, Shopping Mall Management, Manufacturing and Warehousing etc. using Computer Vision technology.
- Engineered a cutting-edge **Document Parsing solution** utilizing advanced OCR technology to accurately extract and process government ID information (PAN and Aadhar), resulting in a 40% reduction in manual data entry errors and saving 100+ hours monthly.
- Spearheaded the creation and implementation of LLM models for Video Summarization, ASR, and NER, resulting in a 25% improvement in accuracy rates and a 30% reduction in processing time.
- Built question-answering bots using **Retrieval-Augmented Generation** frameworks, efficiently utilizing vector stores and language models to provide accurate responses in less than 1 second 95% of the time.
- Designed an **Email Automation system** that utilizes Mail API's and LLM's to efficiently understand the context of incoming mails and respond to them appropriately without human intervention, handling over 1,000 customer emails per day.

Randstad

Client: Institute for Development and Research in Banking Technology (IDRBT)

Data Scientist & Research Associate

Hyderabad, IN

September 2021 – August 2023

- Crafted **8 KNIME** workflows using **LIME, SHAP**, and Counterfactual methods to explain black box models for Analytical Customer Relationship Management in **Banking and Finance**
- Conducted research on **Variational Autoencoder** which led to development of **Chaotic Variational Autoencoder** based **One Class Classifier** for Insurance **Fraud Detection**.
- **Guided over 10 practical lab sessions** for banking professionals visiting the IDRBT for various programs.
- **Acquired, Cleaned, and Preprocessed over 15 datasets** using **machine learning** methods to gain insights.
- Collaborated with a leadership team of 4 to identify relevant questions and determine the best methods of collection, and research to increase knowledge and provide valuable contributions.
- Completed more than 20 data models and conducted analysis to produce reports that showcase the outcomes and valuable insights obtained.
- Conducted training for a group of 9 individuals from various banks on the comprehensive implementation of AI/ML models.
- Deployed 3 **End to End Machine Learning** applications using **Docker, Kubernetes and MLflow**
- Facilitated the "Advance AI/ML for Banks" program for a group of 6 individuals from diverse banking institutions.

GlobalLogic

Associate Analyst

Hyderabad, IN

April 2021 – September 2021

- Worked on **Google Waymo project** to tag images and videos for self-driving cars using 4 variations of cloud compute tools.
- **Presented 20+ findings and insights** to senior management to establish best practices and guide analysis into action and results.

DevOps Enabler&Co

DevOps Engineer Intern

Bengaluru, IN

January 2020 – June 2020

- **Designed and built 2 prototypes for CI/CD** for a website using **GCP, Git, Docker, Kubernetes, Jenkins** and documented project design for reference and future use cases.

PROJECTS

Intelligent Video Analytics for Efficiency Tracking: (Pranathiss)

- Provides facility to stream live or recorded video streams within a customizable web interface.
- Performs real-time analysis of video frames using computer vision algorithms to detect people and track occupancy.
- Occupancy statistics like total count, density, wait times, and flows are quantified and updated dynamically.
- Visualizes insightful occupancy data via customizable dashboards tailored for business needs.
- Enables configuration of alerts and notifications based on occupancy thresholds for desired automation.
- Analytics empower informed decision making around resource allocation, layout optimizations etc.

Email Automation System with LLM for Contextual Response Generation: (PranathiSS)

- Implements authentication setup and client configuration to access chosen mail API. Credentials like API keys, tokens are specified as per API requirements.

- Emails are fetched from the inbox or other folders by invoking relevant methods of the client library.
- New email messages are composed by understanding the context of incoming messages and utilizing the vector databases to fetch relevant information about that context and using an LLM to generate a response that suits appropriately.

Explainable AI for Analytical Customer Relationship Management in Banking and Finance: (IDRBT)

- 8 KNIME processes were developed to learn a black box model using a generalized linear model and to use LIME/SHAP to describe the model's choices.
- Technologies/Algorithms used: **LIME, SHAP, KNIME Analytics Platform, Generalized Linear Model, RandomForest**

Chaotic Variational Auto Encoder based One Class Classifier for Insurance Fraud Detection: (IDRBT)

- Here, we employed the logistic chaotic map to generate random noise in the latent space. The effectiveness of **C-VAE** is demonstrated on the health insurance fraud and auto insurance datasets. We considered vanilla VariationalAutoEncoder (VAE) as the baseline.
- It is observed that **C-VAE outperformed VAE** in both datasets. C-VAE achieved a classification rate of 77.9% and 87.25% in health and automobile insurance datasets respectively.
- Further, the t-test conducted at 1% level of significance and 18 degrees of freedom infers that C-VAE is statistically significant than the VAE.

SKILLS

Data Science Frameworks and Libraries: Scikit-learn, NumPy, SciPy, Plotly, Pandas, Matplotlib, Tensorflow, Keras, Pytorch, OpenCV, Huggingface, H2O, MLflow, DVC, Evidently AI

Programming languages and Big Data: Python, R, Hadoop, Apache Spark, Yarn Cluster Setup

Web Development: HTML/CSS, Flask, FastAPI, Gradio

Cloud and Databases: GCP, Linux, MySQL

Machine Learning/Deep Learning Techniques: Linear Regression, Logistic Regression, Support Vector Machine, Ensemble Trees, RandomForest, A/B Testing, Data Mining, Predictive Modeling/Analytics, Statistical Analysis, Anomaly/Outlier Detection, Customer Segmentation Analysis, Artificial Neural Networks(ANN), Convolution Neural Networks(CNN), Recurrent Neural Networks(RNN), Long Short term Memory(LSTM), Variational Autoencoder(VAE), Generative Adversarial Networks(GAN)

NLP Techniques: BoW, Word2Vec, Doc2Vec, TF-IDF, BERT, NLTK, Spacy, Sentimental Analysis

CI/CD: GIT, Jenkins, Docker, Kubernetes, Github, Gitlab

Tools and IDE: KNIME Visual Analytics, Jupyter Notebook, Google Colab, Visual Studio Code

CERTIFICATIONS

- Diploma in Data Science by HSAI

ACCOMPLISHMENTS

- Got awarded 3 times as "**Alpha Player**" in **Globallogic** in the period of 6 months for being a top and accurate contributor for the project.