

Hemanth Kumar Kothamasu

Phone: +91 7993190439 | Email Id: kothamasuhemanth85@gmail.com | LinkedIn: <https://www.linkedin.com/hemanth>

Professional Summary

Data Scientist with 4 + years of experience in the IT industry, specializing in Python, SQL, and Tableau to solve complex problems and automate tasks. Skilled in developing deep learning and machine learning models, image processing. Continuously expanding knowledge in Generative AI to deliver innovative solutions

Skills

- **Statistical Analysis:** EDA, Probability Distribution, Descriptive statistics, Hypothesis Testing, Outlier Detection, Inferential Statistics
- **Databases & Tools:** MySQL, MongoDB, ChromaDB, AstraDB, FAISS, Power BI, Tableau.
- **Machine Learning:** Classification/Regression Algorithm (Scikit-Learn): Linear, Logistic, Decision Trees, Random Forest, SVM, KNN, XGBoost, Clustering: (K-means, DBSCAN)
- **Deep Learning & Frameworks:** CNN, ANN, RNN, Tensorflow, Pytorch, keras, LSTM
- **Computer Vision:** OpenCV, shapely, YOLO, SAM, UNET, Deeplabv3+, Image Processing
- **Natural Language Processing (NLP) & Model Fine-Tuning:** Tokenization, Stemming, POS tagging, Lemmatization, NER, TF-IDF, BOW, Word2Vec, Transformers, BERT, RAG, Prompt Engineering, LangChain, Semantic Search, Hugging Face
- **Cloud:** AWS, S3 Browser
- **Miscellaneous:** Visual Studio, QGIS, PuTTY, GitHub

Experience: -

Data Scientist | AgriBridge

01 - May 2024 - Present

- Worked on satellite image analysis for agriculture using deep learning and computer vision.
- Handled tasks like road and tree detection, crop boundary mapping, and image processing using models like DeepLabV3+, U-Net, and YOLO.
- Used libraries like OpenCV, PyTorch, Tensorflow, GeoPandas, and Shapely for spatial data handling and visualization
- Collaborated with research and field teams to validate results and improve model accuracy by 25%.
- Enhanced the accuracy and efficiency of agricultural predictions by leveraging advanced AI models.

Data Scientist | Everest Fleet

21- August 2023- 17- April 2024

- Automated data pipelines to monitor car rentals, usage, and downtime.
- Built LSTM models to forecast demand, improving planning accuracy.
- Used Python, SQL, and pandas for scalable data analysis, delivering dashboards that reduced reporting time by 40%

Data Analyst | Tata Consultancy Services (TCS)

07-January 2021– 18- August 2023

- Created interactive Tableau & Power BI dashboards, leading to a 40% reduction in manual reporting time.
- Built SQL queries in MySQL to pull and organize data from databases.
- Used charts like bar graphs, scatter plots, and line graphs to show trends and help teams make decisions.
- Worked with business teams to design easy-to-use dashboards with filters and real-time data.

Projects

Road Detection

- Developed an end-to-end pipeline for road detection using Segmentation Models PyTorch (SMP).
- Achieved 94% pixel-wise segmentation accuracy with DeepLabV3+ and a ResNet50 encoder.
- Improved model accuracy by 30% through strategic image tiling and multi-angle rotation techniques
- Extracted accurate road polygons using OpenCV contours and converted them to GPS coordinates. Converted pixel-based road predictions into real-world GPS coordinates.
- Enabled identification of non-crop zones and rural connectivity gaps to support village planning and infrastructure development.

Tree Detection

- Developed a Tree detection pipeline for horticulture satellite images at zoom level 18, optimizing accuracy through extensive trials achieving an accuracy of 92%
- Used U-Net for detecting small trees and YOLOv8-Seg for large mature trees, optimizing model performance based on tree size.
- Converted detected tree pixels to geographic coordinates for GIS-based visualization and analysis.
- Enabled monitoring of plantation density, assessment of tree health, and informed decision-making for precision horticulture.

Crop Type Detection

- Used NDVI indices values to classify crop types such as paddy, maize, and wheat.
- Developed a sequential deep learning model using RNN to predict crop types based on NDVI time-series data.
- Implemented LSTM networks, achieving an accuracy of 83% in crop classification.
- Applied smoothing techniques to preprocess raw NDVI time-series data, improving data quality for classification.
- Handled class imbalance using techniques like oversampling and weighted loss functions to ensure the model gives equal importance to all crop types.

AgriBot: AI-Powered Knowledge Assistant for Smart Farming

- Collected and processed remedy-related agricultural documents, precomputing their embeddings and indexing them in **Pinecone**, a scalable vector database for efficient semantic retrieval.
- Farmers submit **text-based queries** describing crop issues, and relevant remedies and disease information are retrieved from Pinecone.
- Retrieved context (disease details, remedies, preventive measures) is passed to an **LLM via LangChain** to generate context-aware, farmer-friendly responses.
- Containerized the application using **Docker** and deployed it on **AWS using Amazon ECS with Fargate** and an **Application Load Balancer (ALB)**, ensuring scalability, security, and high availability.

Education

- B.Tech(SRKREngineering College) - 2020

Certifications

- Introduction to R Software(conducted by IIT)
- SQL (Intermediate) HackerRank
- StackRoute Certified Data Scientist