KRISHNA ANVITH V

Data Scientist | ML Intern

+1(551) 331 4666 | vattikutianvith@gmail.com | LinkedIn | GitHub | Anvith Portfolio

PROFESSIONAL SUMMARY

Data Scientist with hands-on experience developing end-to-end ML systems, explainability tools, and robust evaluation pipelines. Proficient in Python, TensorFlow, and Scikit-learn, with applied expertise in Deep Learning (CNN, LSTM), model interpretability using SHAP, and building human-in-the-loop pipelines for ML evaluation. Developed and deployed tools for understanding ML model failures across domains such as audio classification, environmental health, and sensor-based activity recognition. Passionate about responsible AI, red teaming, and building safe, explainable generative systems. Seeking to contribute to mission-critical safety efforts in LLMs, diffusion models, and AI reliability at scale.

SKILLS

Languages: Python, SQL, C, Java, HTML, CSS
ML/AI: TensorFlow, Keras, PyTorch, Scikit-learn, NLP, Deep Learning
Data Analysis: Power BI, Tableau, Jupyter, Anaconda, Google Colab
DevOps/Cloud: Apache Spark, VS Code, Flask, MySQL (AWS – learning)
Big Data: Proficient in handling and processing large datasets.

TECHNICAL EXPERIENCE

Data Science Intern | UpGrad **Jan 2024 - May 2024**

- Gained Built automated ML pipelines on real-world datasets using Scikit-learn and TensorFlow; focused on explainability and failure diagnostics.
- Developed a Human Activity Recognition system (97.53% accuracy) using LSTM and CNNs; conducted misclassification analysis using confusion matrices and performance metrics.
- Reduced training time by 20% and processing time by 30% through pipeline optimization.
- Contributed to robust model evaluation using domain-specific metrics and interpretability techniques.

Computer Vision & IoT Summer Intern | TekWissen Software **May 2023 - Aug 2023**

- Designed and deployed a plant disease detection system using TensorFlow and OpenCV.
- Created annotated image datasets and augmented them to simulate adversarial and noisy conditions.
- Evaluated model failure cases using precision, recall, and F1 metrics; documented challenges in low-power edge inference.
- Prototyped deployment using Raspberry Pi + TensorFlow Lite, aligning with real-world red-teaming setups.

TECHNICAL PROJECTS

Explain My Model – SHAP Dashboard (Deployed Web App) **Jan 2025 – May 2025**

- Developed an interpretable ML dashboard for understanding model predictions via SHAP summary and waterfall plots.
- Integrated preprocessing, outlier removal, and target column selection to support tabular CSV input.
- Enabled failure analysis and misclassification insights by visualizing SHAP feature contributions per prediction.
- Deployed publicly on Streamlit and version-controlled on GitHub **explain-my-model.streamlit.app**

EHR & Predictive Analytics for Environmental Health **Jan 2025 – May 2025**

- Integrated 25,000+ rows of EHR, pollution, and health indicator data using Spark and Pandas.
- Predicted city-level Environmental Health Index using Gradient Boosting and Random Forest ($R^2 = 0.83$).
- Designed risk visualizations and geospatial plots in Tableau to aid health policy decision-making.
- Delivered a reproducible ML pipeline ready for weekly updates and real-time deployment.

UrbanSound8K: Environmental Sound Classification **Jan 2025 – May 2025**

- Processed 8,732 audio clips into Mel-spectrograms and MFCCs, enhanced preprocessing pipeline for noise-rich data.
- Developed CNN and DNN models achieving up to 76% accuracy on 10 urban sound categories.
- Boosted model generalization by 18% using augmentation (noise injection, time-shifting).
- Evaluated predictions with ROC-AUC and confusion matrix; identified misclassification patterns.

Bank Management System — Flask & MySQL Web App **Jan 2025 – May 2025**

- Developed a secure full-stack banking dashboard using Flask, MySQL, and HTML/CSS to manage customer accounts, loans, and transactions.
- Designed and optimized complex **MySQL queries** (JOINs, filters, COUNT, LIMIT) for customer analytics, account summaries, and loan KPIs.Implemented admin authentication with password hashing and session-based access control.
- Built features like search by account number, new customer registration, and interactive dashboards.
- Tools: Python, Flask, MySQL, SQL, HTML/CSS, Bootstrap, Jinja2

EDUCATION

Master of Science in Data Science **Aug 2024- Present**
New Jersey Institute of Technology, Newark, New Jersey, US
Bachelor's in computer science engineering Specialization Data Science (AI and ML) **Aug 2020-May 2024**
Lovely Professional University, Punjab, India

CETIFICATIONS & LEARNING

Neural Networks and Deep Learning – Coursera **Expected July 2025**
Google Data Analytics Professional Certificate – Coursera **Jan 2025 – March 2025**
Advanced SQL for Data Scientists – DataCamp **March 2025 – May 2025**

AWARDS & RECOGNITIONS

National Service Scheme (NSS), Volunteer **Aug 2022 - May 2024**

- Received official recognition from the Ministry of Youth Affairs and Sports, Government of India, for outstanding contributions to social service initiatives.
- Completed 470+ hours of community service, actively contributing to social welfare and awareness campaigns.
- Led and coordinated multiple community-driven projects, including blood donation camps, sanitation campaigns, and educational workshops, benefiting 500+ individuals.