

# AMBIGAPATHI V

Mettur Salem - Tamil Nadu

📞 9488936650    ✉ [ambigapathikavin2@gmail.com](mailto:ambigapathikavin2@gmail.com)    [🌐 linkedin.com/in/ambigapathi-v](https://www.linkedin.com/in/ambigapathi-v)    [🐙 github.com/Ambigapathi-V](https://github.com/Ambigapathi-V)  
🌐 <https://ambigapathi-v.github.io/portfolio/>

## Professional Summary

Results-driven **Data Scientist and Machine Learning Engineer** with expertise in **NLP**, **Generative AI**, and **predictive modeling**. Proficient in **Python**, **TensorFlow**, and **SQL**, with hands-on experience in building and deploying scalable machine learning solutions. Skilled in **large language models (LLMs)**, **data pipelines**, and **data visualization**. Passionate about solving complex problems and driving data-driven decision-making.

## Technical Skills

**Programming Languages:** Python, SQL

**Machine Learning & AI:** Machine Learning, Deep Learning, NLP, Generative AI (GenAI), LLMs, Statistical Analysis

**Frameworks & Libraries:** TensorFlow, Keras, PyTorch, Scikit-Learn, Transformers, NLTK, SpaCy, Matplotlib, Seaborn

**Tools:** GitHub, Docker, MLflow, Visual Studio Code, Jupyter Notebook, DVC, Dagshub, PowerBI

**Data Analysis:** Feature Engineering, Web Scraping, Data Pipelines, Exploratory Data Analysis (EDA)

**Soft Skills:** Collaboration, Problem-Solving, Analytical Thinking

## Experience

### Kitwe News Aggregator (Omdena)

Sep 2024 – Dec 2024

*ML Engineer*

*Remote*

- Developed a fake news detection tool using **BERT**, achieving **98% accuracy**.
- Processed **100,000+ news articles** and deployed a real-time classifier using **Streamlit** for practical applications.
- Optimized caching techniques to reduce false positives by **30%**, improving system reliability.

### Analyzing Healthcare Accessibility - Sudan (Omdena)

Sep 2024 – Present

*Data Scientist*

*Remote*

- Developed a **disease forecasting model** using **historical**, **real-time**, and **environmental data** to predict outbreaks in Sudan.
- Conducted **data collection**, **preprocessing**, and **exploratory data analysis** to identify disease patterns and healthcare gaps.
- Built and deployed a **machine learning model** and a **user-friendly web application** to provide real-time insights for healthcare professionals and decision-makers.

## Projects

### Harmful & Offensive Word Prediction | *Python, NLP* — GitHub | Live App

Jan 2024

- Built a harmful language classifier with **90% precision**, improving content moderation efficiency.
- Utilized **Spacy** and **NLTK** for optimized feature extraction, reducing processing time by **40%**.
- Deployed on a web platform, cutting **false positives by 30%** to enhance accuracy.

### Credit Risk Assessment Tool | *Lauki Finance, Streamlit* — GitHub | Live App

Mar 2024

- Designed a credit risk analysis system using **logistic regression** and **decision trees**, delivering reliable and explainable predictions.
- Processed **50,000+ loan records** and employed **SMOTE** to address class imbalance, improving model reliability.
- Built a **Streamlit-based dashboard** for real-time loan risk evaluation, enabling actionable insights for financial analysts.

### Tomato Disease Classification | *TensorFlow, Python, Deep Learning* — GitHub | Live App

Sep 2024

- Developed a CNN-based framework using **TensorFlow** to classify tomato plant diseases from **13,000+ images**, ensuring early detection with **90% accuracy**.
- Enhanced model robustness through **data augmentation** and **hyperparameter tuning**, improving performance under varied conditions.
- Delivered a **user-friendly prediction tool** to support farmers in managing crop diseases effectively.

Leadership

<b>Open-Source Contributor</b>	<b>Jan 2024 – Present</b>
<i>Omdena</i>	<i>Remote</i>
<ul style="list-style-type: none"><li>• Collaborated with a global community to solve real-world problems through AI and machine learning projects.</li><li>• Contributed to multiple <b>Omdena projects</b> including <b>Kitwe News Aggregator</b> and <b>Healthcare Accessibility in Sudan</b>, utilizing <b>NLP</b>, <b>predictive modeling</b>, and <b>data engineering</b>.</li><li>• Worked on improving project models, optimizing data pipelines, and deploying machine learning solutions that provide impactful results in healthcare and news aggregation.</li></ul>	

Education

<b>Annamalai University</b>	<b>May 2018 – May 2022</b>
<i>Bachelors in Agriculture</i>	<i>Chidambaram, Tamil Nadu</i>

Certifications

- Complete Data Science, Machine Learning, DL, NLP Bootcamp - Krish Naik (Udemy, 2024)
- Master Machine Learning for Data Science - CodeBasics (Jan 2024)
- Complete MLOps Bootcamp - (Udemy, 2024)