Ambigapathi V

Data Scientist

Mettur Salem - Tamil Nadu

9488936650 ambigapathikavin2@gmail.com linkedin.com/in/ambigapathi-v github.com/Ambigapathi-V

https://ambigapathi-v.github.io/portfolio/

Professional Summary

Aspiring Data Scientist with a strong foundation in machine learning, statistical analysis, and data engineering. Proficient in Python, SQL, and TensorFlow, with hands-on experience in building predictive models and data visualization. Strong analytical skills and a passion for leveraging data-driven insights to drive business decisions and innovation.

Education

Annamalai University **May. 2018 – May 2022**

Bachelor of Agriculture *Chidambaram, Tamil Nadu*

Relevant Coursework

- Python for Data Science
- Statistical Machine Learning(ML)
- Deep Learning
- Natural Language Processing (NLP)
- Machine Learning
- Data Visualization Techniques
- Big Data Analytics
- Data Structures and Algorithms

Technical Skills

Programming Languages: *Python*
Machine Learning Tools: *TensorFlow, Keras, Scikit-Learn, NLTK, Spacy, Transformers*
Data Visualization: *Matplotlib, Seaborn, Plotly, Power-BI*
Development Tools: *GitHub, MLflow, Docker, Visual Studio Code, Jupyter Notebook, DVC, Dagshub*
Data Preprocessing: *Feature Engineering, SMOTE,EDA*
Soft Skills: *Team Collaboration, Problem-Solving, Critical Thinking*

Projects

Harmful & Offensive Word Predictiont | *Python, NLP, Text Classification — 90% AccuracyGitHub* **November 2024**

- Developed a Python-based model to predict harmful and offensive words in text, achieving **90% accuracy** in identifying inappropriate content.
- Implemented text classification using **Spacy and NLTK for feature extraction and pre-processing**, enhancing prediction accuracy. using **Spacy** and **NLTK**, improving response time by **40%**.
- Deployed the model on a web platform, reducing false positives by **30%** and improving content moderation efficiency.

Credit Risk Model Development | *Lauki Finance, Streamlit — 92% Accuracy — GitHub* **August 2024**

- Led the creation of a credit risk model using **logistic regression and decision trees**, categorizing loan applications as Poor, Average, Good, or Excellent, enhancing risk assessment accuracy and model explainability.
- **Collaborated** with cross-functional teams to ensure smooth integration with the financial system, enhancing operational efficiency.

Tomato Disease Classification | *TensorFlow, Python, Deep Learning — 90% Accuracy — GitHub* **September 2024**

- Led a team to develop a **CNN model** with **TensorFlow**, classifying tomato diseases from a dataset of 13,000 images, achieving 90% accuracy to help farmers identify diseases early and minimize crop loss.
- Utilized knowledge from my degree in Agriculture to understand the implications of plant health, enhancing the model's relevance to real-world agricultural challenges.
- Implemented **data augmentation** techniques to enhance model robustness and improved prediction accuracy through hyperparameter tuning.

Customer Churn Prediction | *Deep Learning, Streamlit — 85% Accuracy — GitHub* **September 2024**

- Designed and implemented a deep learning model using an **Artificial Neural Network (ANN)** to predict customer churn. Analyzed customer behavior patterns and utilized **Keras** for model development, achieving an accuracy of 85%, enabling strategic retention efforts.
- Led **data preprocessing and feature engineering** to optimize the model's accuracy.

Certifications

Complete Data Science, Machine Learning, DL,NLP Bootcamp- - Krish Naik *Udemy (2024)*

Master Machine Learning for Data Science *CodeBasics (January 2024)*

Complete MLOps Bootcamp - Krish Naik *Udemy (2024)*