RESUME

NAME: THARUN VENKATA MANIKANTA YAJJUVARAPU

Email id: yajjuvarapu.tharun@gmail.com

Mobile No: 6305520510.

Linked In Profile: https://www.linkedin.com/in/tharun-venkata-manikanta-yajjuvarapu-ab2697268.

Git Hub : https://github.com/YajjuvarapuTharun123.

My Portfolio Website: https://tharunai.in/.

SUMMARY:

- An accomplished data scientist with a strong foundation in data analysis, predictive modeling, and data driven strategy development.
- Proven track record of extracting actionable insights from complex datasets and implementing data driven solutions to solve real-world business problems.
- Proficient in programming languages such as Python and experienced in data manipulation using SQL.

SKILLS SUMMARY:

- Developed and deployed machine learning and deep learning models based on client requirement.
- Collaborated cross-functionally with to understand their data needs and provided data- driven recommendations that enhanced decision-making processes.
- Utilized data visualization tools (matplotlib, seaborn with python and Power BI) to create informative dashboards and reports for stakeholders, resulting in improved data accessibility and comprehension.
- Participated in data pipeline development and data engineering tasks to ensure data quality and availability for analysis.

TECHNICAL PROFICIENCY:

- **Programming Languages:** Python , SQL(Structured Query Language).
- **Data Analysis & Visualization:** Python(Pandas, Numpy, Matplotlib, Seaborn).
- **Machine Learning & Deep Learning:** Python(Scikit-Learn, Tensorflow, Keras).
- **Natural Language Processing & Computer Vision:** Python(NLTK, Gensim, Spacy, Open CV).
- **Gen AI frameworks & LLMs :** LangChain, Google Gemini, Ollama, Open AI, Hugging Face.
- **Web Framework:** Python(Flask).
- **Analytical Tools:** Microsoft Power BI, Excel.
- **Databases:** MYSQL, Microsoft SQL Server.

ACADEMIC QUALIFICATION:

Institution Name	Qualification	CGPA	Year of Passing
MVR College of Engineering and Technology	B. Tech(Computer Science and Engineering)	7.0	2024
Sri Chaitanya Junior College	Intermediate Board	7.4	2020
Sri Chaitanya Techno School	SSC	8.7	2018

PROJECT 1:

Title: Email Spam Detection using Natural Language Processing(NLP).

Technical Environment: Python, Pandas, Numpy, Matplotlib, Seaborn, Scikit-Learn, NLTK, Flask.

Project Description:

The Email Spam Detection project utilized the Multinomial Naive Bayes classification algorithm to identify malicious emails, determining whether an email is spam or non-spam. Key steps included:

- Collecting data from various sources and cleaning it using Natural Language Processing (NLP) techniques.
- Preprocessing text data and applying text feature extraction techniques to convert it into numerical format.
- Training the dataset using Multinomial Naive Bayes and Support Vector Machine (SVM) classification algorithms.
- Achieving a high prediction accuracy of 96% with the Multinomial Naive Bayes algorithm, which outperformed SVM.
- Deploying the model using the Flask web framework for user access.

Git Hub Repository: https://github.com/YajjuvarapuTharun123/email_spam_detection-using-NLP.

PROJECT 2:

Title: Potato Leaf Disease Classification using CNN.

Technical Environment: Python, Pandas, Numpy, Matplotlib, Seaborn, Tensorflow, Keras, Flask.

Project Description:

Developed a deep learning project focused on agriculture, specifically aimed at classifying potato leaf diseases. Key components of the project included:

- Collecting image data from various sources and applying preprocessing and data augmentation techniques to enhance model performance.
- Building a Convolutional Neural Network (CNN) architecture tailored for image classification tasks.
- Evaluating model performance, achieving an accuracy of 97%.
- Deploying the CNN model using the Flask web framework for accessibility.

Git Hub Repository: https://github.com/YajjuvarapuTharun123/potato_leaf_disease_prediction-using-CNN.

PROJECT 3:

Title: PDF Insight: RAG-Based Q&A System Using LangChain and Google Gemini.

Technical Environment: Python, LangChain, Faiss, Google Gemini, Streamlit.

Project Description:

Developed a robust Retrieval-Augmented Generation (RAG) application that utilizes LangChain and Google Gemini LLM for question answering based on custom PDF data. Key components of the project included:

- Designing a system to extract and preprocess data from PDF documents, ensuring high-quality input for the model.
- Leveraging LangChain to facilitate effective interaction with the Google Gemini LLM, enabling advanced natural language processing for question answering.
- Implementing a question-answering interface that allows users to query the PDF data, providing accurate and contextually relevant responses.
- Focusing on user experience by creating an intuitive interface for seamless interaction and efficient data retrieval.

CERTIFICATIONS:

- **Python Fundemantals** – Infosys.
- **Data Analytics fundamentals -** IBM.
- **Data Science with Python –** IBM.
- **Machine Learning with Python –** IBM.
- **Deep Learning using Tensorflow –** IBM.
- **Google Analytics Certification –** Google.
- **Building Gen AI APP 12+ Hands-on Projects with Gemini Pro –** Udemy.

INTERNSHIP 1:

Company Name: Afame Technologies, Bengaluru(Remote).

Title: Customer Churn Prediction Using Machine Learning(ML).

Duration: June 2024 – August 2024.

Role: Machine Learning Engineer Intern.

Description: This internship project involved predicting customer churn using Logistic Regression, Random Forest, and Gradient Boosting algorithms. Data preprocessing included handling missing values, encoding categorical variables, and normalizing features. Models were evaluated on Accuracy, Precision, Recall, F1Score, and ROC-AUC, with Random Forest achieving the highest accuracy. The model was deployed as a web application using Flask for real-time predictions. This project highlights the effectiveness of ensemble methods in churn prediction and offers a practical solution for enhancing customer retention strategies.

INTERNSHIP 2:

Company Name: EVOASTRA VENTURES PVT LTD, Mumbai(Remote).

Title: license_number_plate_recognition using YOLOv8(ML).

Duration: September 2024 –December2024.

Role: Data Scientist.

Description: During my internship at EVOASTRA Ventures Pvt Ltd as a Data Scientist, I worked on a project focused on license number plate recognition using YOLOv8 for detection and Optical Character Recognition (OCR) for recognition. This project involved developing an efficient system to identify and extract vehicle number plates from images. I implemented YOLOv8 to accurately detect the number plates in various conditions, ensuring high precision in diverse environments. Following detection, I utilized OCR techniques to convert the visual text into machine-readable format. The integration of these technologies aimed to enhance automated vehicle monitoring systems, showcasing the effectiveness of deep learning in real-time applications.

STRENGTHS:

- Accuracy.
- Communication Skills.
- Team Work.

HOBBIES:

- Problem Solving.
- Learning New Technologies.
- Leadership.
- Presentations.

DECLERATION:

I hereby declare that all the above mentioned is true to the best of my knowledge and belief.

Y T V Manikanta