J. Venkateswararao

Data Scientist

jammula.venkatesh10@gmail.com

+91-8790818108

Guntur, Andhra Pradesh

SUMMARY

Dynamic and results-driven Data Scientist with over 5+ years of experience specializing in Generative AI, Large Language Models (LLMs) and good knowledge on Agentic AI

Demonstrates expertise in the complete data science life cycle, including data gathering, wrangling, extraction, and visualization of both structured and unstructured data.

Proficient in developing and deploying sophisticated machine learning models, including supervised algorithms for classification and regression, as well as unsupervised techniques for clustering and association.

Strong analytical skills combined with a solid understanding of data mining and deep learning methodologies, ensuring effective training and evaluation of models to drive actionable insights and innovative solutions.

Passionate about leveraging data to solve complex problems and enhance decision-making processes.

KEY SKILLS

Generative AI | Agentic AI | LLM Models | Machine Learning | Deep Learning | NLP | Python

Transformers | Scikit-Learn | TensorFlow | Pytorch | Pandas, NumPy | Langchain | LlamaIndex | RAG

Git | Aws Sagemaker | Databricks | OpenAI, Huggingface | Docker, kubernetes

API: Django, FastAPI, Flask | DB: MongoDB, PostgressSQL, VectorDB | IDE: VsCode, Colab, Jupyter, PyCharm

PROFESSIONAL EXPERIENCE

Data Scientist

Technomold IT Solutions Pvt Ltd

April 2020 Present / Hyderabad, India

Domain # : Generative AI (Question and Answer Generating System)

The main objective of this project is to develop an advanced Question and Answer (Q&A) system that leverages **Generative AI** and **Retrieval-Augmented Generation (RAG)** techniques using **LLM**.

This system aims to enhance the accuracy and relevance of answers by combining generative models with a robust retrieval mechanism.

Responsibilities

Knowledge Management is a question and answering system where we are using LLMs to build the communication system for the customer to get related answers from the system.

The dataset should be Prepared from pdf files by preprocessing the data and need to prepare Instruction dataset like structure.

Split documents into manageable chunks for embedding generation and indexing document into a vector database Implementing RAG via LangChain, integrating a retriever (vector similarity search) with a generative LLM for context-aware responses.

Fine-tune generative models and to improve response quality continuously assess model performance using metrics like accuracy, relevance, and user satisfaction.

Domain # : Content Writing (AI Agent)

An AI-powered content writing helper can significantly streamline the process of creating high-quality content.

Automating repetitive writing tasks such as grammar checks, sentence structuring, and topic suggestions helps writers focus on creativity and idea development.

This AI agent can also generate draft content based on input topics, optimize keyword usage for SEO, and suggest style improvements, making it an essential tool for writers, marketers, and bloggers.

Responsibilities

An AI-powered content writing helper can significantly streamline the process of creating high-quality content.
Automating repetitive writing tasks such as grammar checks, sentence structuring, and topic suggestions helps writers focus on creativity and idea development.
This AI agent can also generate draft content based on input topics, optimize keyword usage for SEO, and suggest style improvements, making it an essential tool for writers, marketers, and bloggers.

Domain # : Object Detection

The main aim of this project is to build a system that detects objects from the image or a stream of images given to the system in the form of previously recorded video or the real time input from the camera. Bounding boxes will be drawn around the objects that are being detected by the system. The system will also classify the object to the classes the object belongs. Python Programming and a Machine Learning Technique named YOLO (You Only Look Once) algorithm using Convolutional Neural Network is used for the Object Detection.

Responsibilities

Gather relevant datasets and pre-process the data, including image resizing, normalization, and augmentation to enhance model performance.

Implement the YOLO algorithm using Python and appropriate libraries like PyTorch, OpenCV. Train the YOLO model on the prepared dataset, adjusting hyper-parameters to optimize performance. Fine-tune the model using transfer learning techniques if applicable. Evaluate model performance using metrics such as accuracy, precision, recall, and F1 score on validation datasets. Optimize the model for speed and accuracy, ensuring it meets real-time processing requirements.

Test the system in various environments to validate its robustness and accuracy, Identify and resolve any issues related to object detection. Develop the real-time inference pipeline to process video streams or images from a camera. Implement bounding box drawing and object classification in the output frames.

Deploy the system in the target environment and establish a maintenance plan for regular updates, performance monitoring, and troubleshooting and collaborate with team members, including other software engineers, and domain experts, to ensure project alignment.

Domain # : Speech Recognition (Automatic Speech Recognition)

The project aims to create an advanced speech recognition system using deep learning techniques to enhance the accuracy of transcribing spoken language into text. Automatic Speech Recognition (ASR) is vital for applications like virtual assistants and accessibility tools, but current systems face challenges with diverse accents, background noise, and specialized vocabulary. To improve ASR performance, the project will focus on three key areas: enhancing accent recognition, increasing noise robustness to filter out background sounds, and improving contextual understanding to better handle domain-specific language.

Responsibilities

Data Collection and Preprocessing: Assemble a diverse dataset that includes various accents, noise conditions, and domain-specific terminologies. Preprocess the data for effective training of deep learning models.

Model Selection and Training: Exploring various deep learning architectures, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformer-based models (e.g., BERT, Wav2Vec), to identify the most suitable approach for speech recognition tasks.

Evaluation Metrics: Establish a set of evaluation metrics such as Word Error Rate (WER), Accuracy, and Processing Latency to assess the model's performance comprehensively.

Implementation and Testing: Develop a functional prototype of the speech recognition system and conduct extensive testing across different environments and user groups to validate its effectiveness.

Domain # : Anomaly Detection (Image Classification))

In India and globally, rail transportation is favoured for its comfort and affordability, with India boasting one of the largest rail networks. Maintaining rail tracks is essential, involving the identification and correction of various defects. This research aims to automate defect identification on tracks using Deep Learning techniques, focusing on seven types of defects. The study trained Convolutional Neural Network (CNN), Xception, and MobileNet models on a dataset of images, revealing that the MobileNet model achieved the highest accuracy on the test data.

Responsibilities

Dataset Compilation: Gather a comprehensive dataset of images representing the four types of rail track defects from various sources, including existing databases and field data and label the collected images with appropriate defect categories to create a well-annotated dataset for training and evaluation purposes. Implemented data augmentation techniques (such as rotation, scaling, and flipping)

Model Development: Research and select suitable deep learning architectures for defect detection, focusing on Convolutional Neural Networks (CNN), Xception, and MobileNet.

Train models on an annotated dataset to enable effective identification and classification of rail track defects. Establish evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrix to assess model performance. Optimize model settings through hyper-parameter tuning techniques like grid search or random search, and compare the performance of the different models to identify the most effective.

Model Deployment: Collaborate with software engineering teams to integrate the trained models into existing rail maintenance systems for real-time defect detection and designed a user-friendly interface for maintenance personnel to visualize defect classifications and related information effectively.

Model Performance Monitoring: Implement a monitoring system to track model performance in production, continuously assessing accuracy and adjusting for changes in data distributions. Scheduled regular updates to the dataset and retrain models as needed to maintain performance effectiveness.

Domain # : Sentiment Analysis (Review Management)

In the competitive restaurant industry, understanding customer feedback is crucial for enhancing service quality and satisfaction. Online reviews on platforms offer valuable insights but can be time-consuming and subjective to analyse manually. This project aims to develop a sentiment analysis model to automatically classify restaurant reviews as positive, negative, or neutral, helping restaurant owners and managers derive actionable insights and make informed decisions to improve customer experience and operational efficiency.

Responsibilities

Gathering reviews clean and pre-processe to remove noise, such as HTML tags and irrelevant information, resulting in a usable dataset. Finally, the reviews are annotated with sentiment labels (positive, negative, neutral) based on predefined Criteria or expert input to create a training dataset.

Feature Engineering: Utilize NLP techniques to extract meaningful features from the reviews, such as n-grams, sentiment scores, and word embeddings (e.g., Word2Vec, Glove).

Model Development and evaluation: The model development process involves researching and selecting suitable machine learning or natural language processing algorithms (such as logistic regression, support vector machines, recurrent neural networks, or transformer models like BERT) for sentiment classification. The chosen model(s) are trained using a labelled dataset to accurately classify restaurant review sentiments. Evaluation metrics, including accuracy, precision, recall, and F1-score, are defined to assess performance, and cross-validation techniques are implemented to ensure the model generalizes well to unseen reviews. Additionally, hyperparameter tuning methods, such as grid search or random search, are used to optimize model parameters and enhance classification performance.

Implementation and Integration: Collaborate with IT teams to integrate the sentiment analysis model into existing restaurant management systems or dashboards for real-time review analysis or design a user-friendly interface for restaurant owners and managers to easily access sentiment analysis results and insights.

Model Updates: Regularly update the sentiment analysis model with new data to improve accuracy and adapt to changing customer sentiments and trends.

Domain # : Mining (Iron ore Quality and Quantity Analysis)

This project aims to develop two machine learning models:
Silica Concentration Prediction: To estimate the percentage of silica in iron ore, reducing the current detection time and improving steel production efficiency by minimizing impurities.
Iron Ore Quantity Forecasting: To predict the amount of iron ore extracted, optimizing inventory management, production planning, and operational efficiency in the steel industry.
The models will analyse historical, geological, and operational data to provide accurate and timely predictions, enabling better decision-making, reducing costs, and enhancing supply chain reliability.

Classification Responsibilities

Data Collection and Preprocessing: Gathering and preparing relevant historical data for classification tasks.
Feature Selection: Identifying key features that influence silica classification (e.g., mineral composition, particle size).
Model Development: Building classification models to predict silica levels (low, medium, high).
Model Evaluation: Assessing performance using metrics like accuracy, precision, recall, and F1-score.
Hyper-parameter Tuning: Optimizing model parameters to improve accuracy and reduce errors.
Deployment and Monitoring: Integrating the model into production and continuously monitoring its performance.
Documentation and Reporting: Providing detailed reports on the process, results, and insights for stakeholders.

Regression Responsibilities

Data Collection and Preprocessing: Gather and preprocess historical extraction data, geological data, and operational factors relevant to predicting silica concentration and iron ore quantity.
Feature Engineering: Create and refine features to capture relationships between input variables and target outputs for better regression model performance.
Model Development: Develop regression models to predict silica concentration percentages and iron ore extraction quantities.
Model Evaluation: Assess model accuracy using metrics like Mean Absolute Error, Mean Squared Error, and R-squared.
Hyper-parameter Tuning: Fine-tune model parameters to improve prediction accuracy and reduce forecasting errors.
Documentation and Reporting: Document the process, results, and insights, providing comprehensive reports to stakeholders on the model's capabilities and operational benefits.

Domain # : Banking (Credit card fraud Detection)

The objective of this project is to develop a credit card fraud detection system using the dataset provided by the client. The dataset contains many numbers of credit card transactions, with both fraudulent and non-fraudulent cases. The goal is to build a machine learning model that can accurately identify fraudulent transactions to help customers of the bank from the risk and minimize losses due to fraudulent activities. It is important that credit card companies can recognize fraudulent credit card transactions so that customers are not charged for items that they did not purchase

Responsibilities

Involved in entire data science and machine learning life cycle process like data gathering, data augmentation, visualizing the data points
Encoding the data and setting the dataset for training and testing
Building the model with the dataset using a classification algorithms like logistic regression and random forest classifier to find weather the transactions are fraudulent or non-fraudulent.
Validating the model by applying the metrics like f1 score for finding the accuracy, if the performance is not good need to improve by doing hyper-parameter tuning on the model to achieve best predictions.

EDUCATION

JNTU University
B-Tech(IT)

06/2006 – 05/2009
Hyderabad, India