Abhinav Chowdary

📧 abhinav.mallampati@kanerika.com

📱 None

# Summary

With 2+ years of experience in building AI-based applications, I am passionate about leveraging Large Language Models (LLMs) and data engineering to create impactful real-world solutions. My expertise includes working on a wide range of prediction and classification ML models and image processing tasks using deep learning frameworks like YOLO. Through these projects, I have honed skills in developing end-to-end machine learning pipelines, deploying scalable models, and integrating cutting-edge AI into practical systems.

# Skills

* Python, RAG, LangChain, SQL, C/C++

# Education

* B.Tech in Computer Science – IIIT HYDERABAD (Jul 2019 - May 2023)

# Experience

* Karl – The Statistician (Report AI) – None, None (Sep 2024 - Present)
* • Designed and implemented a decision-making flow leveraging LLM function-calling capabilities to identify user intent across three key functionalities: chart generation, query answering, and data analysis.
* • Chart Generation: Utilized the LIDA architecture to interpret user queries, derive key performance indicators (KPIs), and generate step-by-step code implementations, ensuring precise and hallucination-free chart creation and editing.
* • Query Answering: Enhanced LLM contextual understanding by preprocessing file metadata (e.g., column names, numerical ranges, unique object values, and counts) to improve accuracy in question answering.
* • Data Analysis: Developed a three-fold analysis framework using Pandas to retrieve relevant data subsets, compare correlated columns, and derive actionable insights for user-defined scenarios.
* • Streamlined processes for data visualization, query generation, and root-cause analysis to cater to diverse business and analytical needs.
* FlipIndex – None, None (May 2024 - Jun 2024)
* • Engineered a dynamic configuration system to securely connect to databases and extract schema details, including all DDL statements across multiple schemas.
* • Developed a mechanism to integrate database metadata (DDL statements), question-SQL pairs, and natural language explanations into a vector database for efficient embedding and retrieval.
* • Implemented context generation for large language models (LLMs) by retrieving relevant embeddings from vector databases to improve SQL query generation accuracy.
* • Facilitated database-agnostic interaction and automated query generation, significantly streamlining the process of database QA and EDA.
* Resume Intelligence & Standardization – None, None (Feb 2024 - Apr 2024)
* • Built Resume Standardization module for format conversion, data redaction, and content enhancement.
* • Utilized RAG and vector databases for efficient resume retrieval and matching.
* • Applied prompt engineering with OpenAI and Anthropic Claude LLMs for flexible model integration.
* • Implemented LLM conversation to evaluate and maintain resume quality, format and ensure professionalism.
* • Used ReRanker models to refine candidate ranking based on job requirements.
* Inventory Optimization – None, None (Dec 2023 - Jan 2024)
* • Developed a forecasting framework to accurately predict inventory requirements by store and product, ensuring optimal stock levels and reducing overstock or shortages.
* • Integrated a large language model (LLM) interface to enable natural language queries for inventory forecasts, allowing seamless interaction and accessibility for non-technical users.
* • Engineered precise prompt strategies to format user inputs and outputs effectively, ensuring clear, actionable results presented in tabular form.
* • Streamlined real-time predictions by combining NLP, prompt engineering, and forecasting methodologies for enhanced decision-making.

# Certifications

* Microsoft Certified: Fabric Data Engineer Associate (DP-700) – None (None)
* Microsoft Certified: Fabric Analytics Engineer Associate (DP-600) – None (None)
* Microsoft certified Azure Data Engineer Associate (DP-203) – None (None)