Raju Gopi

+91 9353973699 rajugraj29@gmail.com www.linkedin.com/in/raju-gopi-contact

SUMMARY

ML/Gen AI Engineer| 2+ Years of Experience

Results-driven **ML/Gen AI Engineer** with **2+ years of experience** in developing AI-driven solutions for business automation. Skilled in **LangChain, OpenAI, Google Gemini**, and **Hugging Face Transformers**, with expertise in **Retrieval-Augmented Generation (RAG), code modernization, and predictive analytics**. Proven ability to integrate scalable AI models, optimize performance, and streamline workflows through intelligent automation. Passionate about delivering high-performance solutions that enhance business efficiency and innovation.

EDUCATION

RNS Institute of Technology **08/2018 – 07/2022**
*Bachelor of Engineering – **8.1** CGPA* ***Bengaluru, India***

TECHNICAL SKILLS

Programming Languages: Python (Primary), Java, SQL
Generative AI & LLMs: Paid and Open-Source LLMs (OpenAI, Google Gemini Pro, Llama2 and Hugging Face Transformers)
AI/ML Technologies: Fine-tuning with custom data, Vector Embedding, NLP, Prompt Engineering, Apache Spark (PySpark), QLoRA (Quantized Low-Rank Adapters), LoRA (Low-Rank Adaptation), RAG (Retrieval-Augmented Generation)
Frameworks & Tools: LangChain, FastAPI, GitHub, VS code, Jira, DataBricks
Deployment & Platforms: GitHub Action, AWS, Docker
Visualization & Documentation: PowerBI, Microsoft Office
Data Science & Miscellaneous: Pandas, NumPy, Scikit-learn, Traditional ML Algorithms, MLFlow, Pyspark
Frontend Development: HTML, CSS, Bootstrap

EXPERIENCE

Infosys – Client: 1

08/2022 – Present
Bengaluru, India

- **AI & NLP Model Integration**: Developed and integrated **AI models** like **Google GenAI**, **OpenAI**, and **Hugging Face Transformers** for automating business processes, including natural language understanding, task generation, and data categorization.
- **Python Backend Development:** Developed Python backend using **FastAPI** for web applications, creating scalable and high-performance APIs to support business automation and data processing workflows.
- **SQL Database:** Leveraged **SQL** for data extraction and manipulation across **relational databases**, creating optimized queries to support data analysis workflows and integration with machine learning pipelines.
- **API Integration:** Designed and implemented robust **API integrations** for data pipeline automation, connecting multiple data sources to establish seamless workflows that improved data accessibility and processing efficiency by 40%.
- **Model Optimization & Evaluation**: Fine-tuned **Machine Learning Models** like **BART** and **GPT**, using domain-specific datasets to improve model accuracy, and applied performance metrics such as **sacrebleu** to evaluate and optimize outputs for task generation and summarization.
- **Exploratory Data Analysis (EDA):** Conducted comprehensive data exploration, handling structured and unstructured datasets to identify trends, correlations, and anomalies, ensuring high-quality input for ML models.
- **Feature Engineering:** Applied feature engineering techniques to optimize data for **predictive modeling** and performance improvements.
- **RAG:** Created scalable **VectorDB integrations** for **retrieval-augmented search (RAG) applications**.
- **Version Control & Collaboration**: Utilized **Git** and **GitHub** for version control and collaboration across teams, ensuring smooth development workflows and effective codebase management.
- **Performance Tuning & Optimization**: Focused on optimizing model inference speed and reducing computation overhead, ensuring automated systems are highly efficient and capable of processing large datasets.

Client: 2

- **Predictive Model Development**: Designed and deployed a predictive model to forecast **Time on Wing (ToW)** for engines, managing the full lifecycle from data collection to visualization, providing actionable maintenance insights.
- **FMECA Analysis**: Led **Failure Modes and Effects Criticality Analysis (FMECA)** for gas turbines using **Excel Power Query** and **Power BI**. Automated data extraction with **Python**, reducing processing time by over 90%. Developed dashboards to highlight critical failure modes.
- **Design Iteration Management**: Oversaw **Design Definition Change Document (DDCD)** assessments, coordinating with internal teams and managing communications with **Airbus** on design changes.
- **Process Optimization & Stakeholder Communication**: Streamlined data workflows and optimized processes, improving efficiency and ensuring effective communication with internal teams and external clients.

ACCOMPLISHMENTS

- **Insta-Award- Client Appreciation:** Designed and developed intuitive and visually compelling Power BI dashboards to present complex, large-scale data in a clear, accessible, and actionable format, enabling stakeholders to easily interpret and make informed decisions.