


PAVITHRA KANDHASAMY SELVARAJ

+91-9500798747 ✉ pavithraselvaraj98@gmail.com  PavithraKS

Education

Coimbatore Institute of Technology

July 2016 - Sep 2020

Bachelor of Engineering in Computer Science

Coimbatore, India

Technical Skills

Languages: Python, SQL, HTML, CSS, Terraform.

Tools & Technologies: ML, DL, LLM, GCP, BigQuery, VertexAI, PostgreSQL, Pandas, Numpy, Sklearn, Looker, PySpark, Keras, Tensorflow, Mlflow, Jenkins, Tekton, Docker, Git, Excel, Flask, Postman, Splunk, Pgloader.

Certifications: Google Certified Cloud Digital Leader.

Professional Experience

Tiger Analytics

Dec 2024 - Present

Data Scientist

Chennai, India

- * Built hybrid **machine learning attributes** for a **Credit Reporting Bureau** by integrating feature engineering in **BigQuery** across **15+** large-scale census tables with **Vertex AI**-powered predictive models such as **Logistic Regression** and **XGBoost** and **PySpark** involving **30GB+** records, optimizing marketing and campaign targeting.
- * Developing a nationwide **Segmentation tool** leveraging **8000+ variables** across **15 million** census records to cluster and classify the Australian population by neighborhood characteristics, enabling hyper-localized market insights.
- * Performing comprehensive **Data analysis** post-attribute generation to compare **Data distributions** across historical and current periods, ensuring **95%** attribute stability and alignment with evolving population dynamics.

Ford Motor Private Limited

Jan 2022 - Nov 2024

Software Engineer (AI/ML)

Chennai, India

- * Designed and implemented a **Machine Learning** model to improve the effectiveness of collections efforts by determining the most optimal communication channel and timing for soliciting a response from delinquent customers.
- * Implemented **Data Preprocessing** and selected **Random Forest**, improving model accuracy from **62%** to **73%** through feature selection and hyperparameter tuning, and deployed the solution to **50%** of customers, achieving a **66%** reduction in contact frequency while maintaining collections effectiveness and enhancing customer satisfaction.
- * Migrated the application to **Google Cloud Platform**, leveraging **Cloud Run** and **Cloud Scheduler**, resulting in a **45K USD** annual cost reduction through database migration from On-Prem SQL Server to **PostgreSQL**.
- * Ensured consistent cloud deployments across environments using **Terraform** reducing cloud service provisioning time by **67%** and a **Tekton-based CI/CD** pipeline to streamline the process from source code push to production deployment.

Ford Motor Private Limited

Jul 2020 - Dec 2021

Junior Software Engineer (AI/ML)

Chennai, India

- * Developed a **Python** application to contact over **70k** past-due customers daily, utilizing a rules-based system to determine optimal contact timing and channel (phone/SMS) based on frequency, time restrictions, and state-specific regulations.
- * Stored over **10 million** customer contact records in an on-prem **SQL Server** by integrating data from **5+ systems** to identify key drivers of payment behavior which powers machine learning models for improved customer communication.
- * Designed and implemented a **CI/CD** deployment strategy for a **Flask** application using **Jenkins** and **Pivotal Cloud Foundry**, reducing deployment time to under **10 minutes**, and integrated **Splunk** for real-time monitoring, enabling **90%** of issues to be resolved within **24 hours** through proactive error detection.

Hackathon Projects

Multilanguage Video Summarization {VertexAI, Gemini-Pro, HTML, CSS, CloudBuild, Flask}

Mar 2024

- Developed a prototype for an LLM-powered video summarization model that transcribes audio, extracts key points, and generates concise, multi-lingual summaries, enhancing accessibility and user experience for lengthy video content.

Question Your Dataset using LLM {VertexAI, PaLM 2, Pandas, PySpark, Langchain, HTML, CSS, Flask}

Mar 2024

- Designed an LLM-based system enabling real-time querying with datasets. The model analyzes data, generates follow-up questions, and guides users to key insights, improving data analysis efficiency and supporting data-driven decisions.

Meeting to Article Content Generation {VertexAI, Gemini-Pro, HTML, CSS, CloudBuild, GCS, Flask}

Mar 2024

- Programmed an LLM-based article generator that transcribes meeting audio, identifies relevant points, and formats them into articles, streamlining content creation and facilitating effective sharing of meeting insights.

Accessing Knowledge Articles using LLM {Langchain, VertexAI, PaLM 2, Gradio}

Oct 2023

- Engineered a prototype of a multi-lingual LLM-powered Knowledge Article access platform, improving user experience by enabling seamless search and retrieval of information in lengthy documents across multiple languages.

LLM based Talent Discovery System {Langchain, VertexAI, PaLM 2, Gradio}

Oct 2023

- Addressed the challenges of traditional recruitment methods, including time-consuming manual processes and potential biases, and developed a prototype using Large Language Models to automate candidate screening and shortlisting.

Car Insurance Cold Calls Prediction {Python, XGBoost, Pandas, Sklearn}

Feb 2022

- Developed an XGBoost model with **89%** accuracy to predict Car Insurance purchase using cold call data. This award-winning approach identified key success factors, optimizing outreach strategies and improving customer retention.