

## Professional Summary

Motivated and results-driven professional with a strong passion for AI, Machine Learning, and Data Engineering, combining practical industry expertise with research-driven innovation. I bring 3 years of industry experience as a Data Engineer, delivering ETL/ELT pipelines, data platforms, and predictive models across healthcare, retail, banking, and logistics domains, along with 2 years of AI/ML research experience as a Research Assistant, building Generative AI applications, RAG pipelines, predictive modeling, and multi-modal AI solutions.

## Skills

Category	Tools & Technologies
Programming & Data Science	Python, FastAPI, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, SQL, PL/SQL, MongoDB
Natural Language Processing (NLP)	Transformers (Hugging Face), spaCy, NLTK, TextBlob
AI & Machine Learning	LLaMA, LangChain, LangGraph, RAG (Retrieval-Augmented Generation), Multi-modal AI
Data & ETL/Backend	Google Colab, Jupyter Notebooks, Vector DBs (Chroma, FAISS), ETL & Backend Development (Extraction, Transformation, Query Optimization), Canvas LMS integration
Database & Enterprise Tools	PL/SQL Development, Stored Procedures, Triggers, Oracle APEX, Oracle Forms & Reports, Siebel CRM, Enterprise Application Support & Enhancement
Development & Collaboration	Git, VS Code, Agile Collaboration (Slack, Trello), API Design & Testing, Pytest

## Education

- **Bachelor of Technology in Electronics and Computer Engineering** | *Kalasalingam Academy of Research and Education, Tamil Nadu, India* | 2017 – 2021
- **Master of Science (MS) in Computer Science** | University of North Texas, Denton, TX | 2023 – 2025

## Experience

### Graduate Research Assistant – Predictive Modeling & AI

- Executed large-scale **data cleaning, feature scaling, and model training** to improve predictive accuracy and stability; developed **prompt variations and sentiment classification pipelines** to analyze ChatGPT's response polarity across domains.
- Built **structured RAG pipelines** for metadata creation and vector-based retrieval, enabling real-time, context-aware AI responses for academic and business knowledge retention.

- Applied **multi-modal AI techniques** combining textual and visual inputs for summarization, and designed/deployed **agentic AI workflows** for IT automation and knowledge management.

#### Associate Consultant | Wipro | Client: METRO AG | Jun 2021 – Aug 2023

- **Designed and optimized SQL/PLSQL workflows** (queries, triggers, stored procedures, functions) to support high-volume retail transactions and enterprise reporting.
- **Built and managed ETL/ELT pipelines**, ingesting data from Oracle, FTP, and APIs into centralized platforms.
- **Automated backend workflows** using Python and Shell scripting with data validation and anomaly detection for reliable reporting.
- **Performed database performance tuning** (Oracle 12c/19c) and optimized queries, reducing reporting latency by 30%.
- **Implemented containerization and CI/CD** with Docker, Kubernetes (AKS), and Jenkins for scalable, automated deployments.
- **Integrated monitoring and logging** using CloudWatch and Grafana, creating dashboards for pipeline health, drift detection, and SLA compliance.
- **Supported ML POCs and data visualization**, preparing datasets, feature engineering, and building KPI dashboards in Power BI/Tableau for business insights.

#### Software Developer | Intern & Software Developer | March 2020 – June 2021

- **Analyzed COVID-19 epidemiological and socio-economic datasets**, providing insights on infection, recovery, and trends by correlating population density, mobility, healthcare infrastructure, and policy interventions.
- **Designed and implemented ETL/ELT pipelines** using SQL, PySpark, and Databricks to transform raw data into curated datasets for large-scale analytics, with automated workflow scheduling via Control-M and Airflow.
- **Developed predictive ML models** (regression, decision trees, random forest, XGBoost) for outbreak forecasting and patient risk analysis, supporting healthcare strategies including pharmacy analytics.
- **Built interactive dashboards and ensured data quality**, using Tableau and Power BI to deliver actionable insights, validate data accuracy, detect anomalies, and collaborate with cross-functional teams for scalable data solutions.

### Projects

- **AI Autopilot for IT Request Automation:** Built a FastAPI microservice with LLM agents to automate IT tasks and generate structured reports. **Tech:** FastAPI, LangGraph, LLM Agent Engineering, API Design
- **Multi-Modal Summarization for Scientific Papers:** Created context-aware summaries by integrating textual and visual data for research insights. **Tech:** Python, RAG, LangChain, Vector Storage, Parsing
- **Smart Tutor AI (STA) – Intelligent Tutoring System:** Developed a RAG-based AI tutor improving student performance and evaluating ChatGPT sentiment consistency. **Tech:** Python, Scikit-learn, MLflow, Pandas, NumPy, NLTK, TextBlob, Matplotlib, Seaborn
- **Brain Tumor Detection using CNN & U-Net:** Built a CNN/U-Net pipeline to segment and classify brain tumors from MRI images accurately. **Tech:** Python, TensorFlow/Keras, OpenCV, NumPy, Matplotlib, CNN, U-Net