

SIVAKUMAR A

AI & ML Engineer

Erode, Tamilnadu | sivakumarbe.lingam@gmail.com | +91 9677753584 | linkedin.com/in/sivakumar-lingam | github.com/sivakumarbe | sivakumar-a-portfolio.netlify.app

PROFILE SUMMARY

AI Machine Learning Engineer with strong project-based experience in Generative AI, NLP, and deep learning, specializing in LLM-powered applications, RAG pipelines, and time-series forecasting. Hands-on in building end-to-end AI solutions including data preprocessing, feature engineering, model development, evaluation, and API deployment using FastAPI. Proficient in Python, TensorFlow, PyTorch, Transformers, embeddings, and vector databases (FAISS) with a strong focus on scalable, production-ready systems.

TECHNICAL SKILLS

- **Languages:** Python
- **Machine Learning:** Supervised & Unsupervised Learning, XGBoost, Scikit-learn, Feature Engineering
- **Deep Learning:** TensorFlow, PyTorch, Keras, CNN, LSTM, Attention Models, Transformers
- **NLP & Computer Vision:** Hugging Face Transformers, BERT, GPT, SpaCy, OpenCV, YOLO, ResNet, VGG
- **Data Science:** Pandas, NumPy, Matplotlib, Seaborn, Plotly, EDA, Statistical Analysis
- **Deployment & Tools:** FastAPI, REST APIs, Git, Docker, Jupyter, VS Code
- **Generative AI:** LLMs, GPT, RAG, Prompt Engineering, Embeddings, Transformers, Tokenization, FAISS, Vector Databases, Semantic Search, FastAPI-based LLM APIs

KEY PROJECTS

AI Career Intelligence Assistant *RAG, LLM, FAISS, spaCy-based NLP, FastAPI, Sentence Embeddings, Redis*

- Designed and implemented an end-to-end RAG pipeline for resume-job description matching using embeddings and FAISS, improving semantic relevance of AI responses by 60–70% compared to keyword-based approaches.
- Built a deterministic skill extraction and matching engine to compute accurate skill-match percentages, eliminating LLM-hallucinated scores and ensuring 100% consistency in metric-based outputs.
- Developed a session-aware, multi-user backend with FastAPI, Redis, and X-Session-ID headers, enabling concurrent document ingestion and analysis without cross-user data leakage.
- Optimized LLM usage via Mock/Real execution modes and controlled prompting, reducing development-time API costs by 80% while maintaining production-ready behavior.
- Delivered an interactive AI product using Streamlit, integrating chat-based RAG queries, skill gap analysis, resume rewriting, and ATS keyword extraction, reducing user task completion time by 40%.

Advanced Time Series Forecasting using Neural Networks *TensorFlow, LSTM, Attention*

- Built LSTM and Attention-based deep learning models for long-horizon time-series forecasting.
- Performed advanced sequence generation, normalization, and feature engineering to prepare temporal data.
- Implemented end-to-end forecasting pipelines including data preprocessing, windowing, scaling, and multi-step prediction, enabling accurate forecasting on high-frequency sequential data.
- Achieved 20–30% lower RMSE compared to traditional statistical baselines (ARIMA / moving averages).
- Evaluated models using MAE, RMSE, and MAPE, creating data visualizations to communicate trends.

Credit Risk Prediction - Machine Learning Pipeline

Scikit-learn, SHAP, LIME

- Built a full ML classification pipeline to predict loan default risk, supporting data-driven credit decisions.
- Conducted comprehensive EDA, preprocessing, feature selection, and SMOTE-based class imbalance handling, improving minority-class recall by 25%.
- Integrated SHAP and LIME explainability, enabling transparent, regulator-ready model interpretation for high-risk predictions.
- Reduced false-negative default predictions by 20%, directly improving risk mitigation outcomes.

ADDITIONAL PROJECTS

Electricity Demand Forecasting using XGBoost

- Developed a time-series regression model to forecast electricity demand from historical consumption data.
- Engineered time-based features (hour, day, week) and lag features to capture temporal patterns.
- Achieved 15–25% improvement in RMSE over baseline regression models.
- Enabled accurate short-term demand forecasting, supporting grid load-balancing decisions
- Trained an XGBoost Regressor to model non-linear demand patterns, achieving high regression accuracy.

Smart Grid Anomaly Detection using Autoencoders

- Built an unsupervised deep learning model to detect abnormal electricity consumption patterns in smart grid data.
- Detected abnormal usage with high precision using reconstruction error thresholds, reducing false alarms by 30%.
- Applied threshold-based anomaly detection to classify abnormal usage patterns.

WORK EXPERIENCE

Service Engineer — Technical support

Aug 2015 – Jan 2025

- Delivered installation, maintenance, and troubleshooting support for technical systems across multiple client sites, resolving 2,000+ service tickets over 10 years.
- Performed root cause analysis (RCA) to diagnose recurring failures, reducing repeat incident rates by 25–30% through corrective and preventive actions.
- Provided on-site and remote technical support, achieving 95% first-time resolution rate and minimizing system downtime by 40%
- Improved system uptime to 99% by proactively identifying potential failures and executing preventive maintenance schedules.
- Acted as the primary technical interface for customers, increasing customer satisfaction scores by 15–20% through timely and reliable service delivery.

CERTIFICATE & EDUCATION

Cultus Education & Technology	2025
Data Science & Artificial Intelligence	
Outskills	2025
Generative AI Mastermind	
B.E. - Electrical and Electronics Engineering	2015
Nandha Engineering College	
Higher Secondary	2010
R.C.M.H.S. School, Erode	