

B. KARTHICK

Fresher Data Scientist | AI Programmer

Hosur, TN | 9489132149 | karthickappu2505@gmail.com | [LinkedIn](#) | [GitHub](#) | [ePortfolio](#)

PROFESSIONAL SUMMARY

Fresher Data Scientist and AI Programmer with a strong foundation in Bio-Medical Engineering. Specialized in Generative AI (LLMs), Retrieval-Augmented Generation (RAG), and Predictive Modeling. Experienced in building ETL pipelines, conducting Root Cause Analysis, and transforming complex datasets into actionable intelligence. Proficient in Python, SQL (joins, aggregations, subqueries), Scikit-learn, LangChain, Pandas, Power BI, and Tableau. Certified in Data Science (Great Learning PGP) and AWS (ML Foundations, GenAI). Seeking a fresher Data Scientist role to apply ML and GenAI skills to real-world problems.

TECHNICAL SKILLS

Languages:	Python, SQL — joins, CTEs, aggregations, subqueries (MySQL)
AI & GenAI:	Meta Llama-3 (LLM), Prompt Engineering, LangChain, RAG, Hugging Face Transformers
Machine Learning:	Random Forest, XGBoost, Gradient Boosting, K-Means Clustering, PCA, Scikit-learn
Data Engineering:	Pandas, NumPy, Snowflake, ETL Pipelines, REST APIs, Data Validation
Analytics & BI:	Power BI, Tableau, Matplotlib, Seaborn, EDA, Statistical Analysis
DL (Basics):	TensorFlow, Keras, PyTorch, Hugging Face Transformers
Tools:	Jupyter Notebook, Git, GitHub, Google Colab
Methodologies:	Feature Engineering, Hyperparameter Tuning, Model Evaluation, Cross-Validation, Data Validation

TECHNICAL PROJECTS

Medical Device Safety Intelligence

Generative AI | RAG | Python | LangChain | Meta Llama-3 | Snowflake | ETL

- **Challenge:** Automated ingestion of live, high-volume JSON records from the OpenFDA API to monitor medical device failure modes — 100% data integrity ensured via custom validation layer.
- **Action:** Engineered a robust ETL pipeline in Python resolving timestamp overflows and schema inconsistencies.
- **AI Integration:** Architected a Retrieval-Augmented Generation (RAG) system using LangChain and Meta Llama-3-8B to interpret engineering failures and generate clinical risk summaries.
- **Impact:** Translated complex engineering failures into actionable clinical summaries, streamlining safety monitoring processes.

Optimum Insurance Cost Forecasting

Predictive & Prescriptive Analytics | Python | Gradient Boosting | K-Means | PCA | Scikit-learn

- **Modeling:** Built a Gradient Boosting model achieving $R^2 = 0.958$ for high-accuracy insurance cost forecasting.
- **Analytics:** Applied K-Means Clustering and PCA to analyze pricing insights and improve transparency across risk segments.
- **Strategy:** Delivered prescriptive recommendations for risk management and targeted health programs.

Predictive Maintenance & Defect Detection

Supervised ML | Python | Random Forest | XGBoost | Scikit-learn | Feature Engineering

- **Development:** Designed Random Forest and XGBoost models achieving 94% training accuracy for predictive maintenance.
- **Specialization:** Developed a specialized model for Laser Melting Deposition achieving 81% defect detection accuracy.
- **Results:** Engineered sensor data features leading to a 15% reduction in manufacturing defects.

WORK EXPERIENCE

Engineer – QA & RA | Panacea Medical Technology Pvt Ltd

Jan 2023 – Dec 2023

- Conducted structured data validation and performance analysis on medical device datasets, achieving 98% accuracy in verification tests.
- Applied data-driven root cause analysis and statistical techniques to detect defect patterns, reducing product failures by 60%.
- Developed and presented quantitative reports and dashboards to cross-functional teams to drive quality improvements.
- Leveraged analytical problem-solving to align product quality with industry-standard data practices.

EDUCATION

Artificial Intelligence Programmer | TNSkills Vetri Nichayam | Feb 2026

PGP in Data Science & Business Analytics | Great Learning | Sep 2025

Certificate: mygreatlearning.com/certificate/JCGDXIIC

B.E. in Bio-Medical Engineering | Adhiyamaan College of Engineering | 9.1 CGPA

CERTIFICATIONS & LANGUAGES

- PGP Data Science — Great Learning | Power BI — Simplilearn | Tableau — Simplilearn | AWS Educate (ML Foundations, GenAI, Databases)
- English (Fluent) | Tamil (Native) | Kannada (Basic)