PRAVEEN SUBRAMANI

Chennai, Tamil Nadu • +919566166528 • Praveen.subramani88@gmail.com • linkedin.com/in/Praveensubramani88

Results-driven **Data Scientist** with a **Master's in Big Data Analytics** and 10+ years of experience in **software development and automation**. Skilled in **AI, LLMs, and machine learning** to drive **data-informed decision-making**. Proven ability to **deliver measurable business impact** through scalable, end-to-end data solutions that combine **technical rigor with strategic insight**.

SKILLS

- **Programming:** Python, SQL, Scikit-learn, TensorFlow, LangChain, NumPy, Pandas, Matplotlib, Seaborn

- **AI & ML:** LLM, RAG, Predictive Modeling, Deep Learning, NLP, Data Imputation

- **Data Engineering:** Hadoop, Spark, Hive, HDFS, Azure Data Factory, REST APIs, Data Pipelines

- **Cloud & Visualization:** Microsoft Azure, CI/CD, Docker, Tableau, Streamlit, Statistical Analysis, A/B Testing

- **Certifications:** Certified Python Programmer, Tableau Specialist, Deep Learning Essentials

WORK EXPERIENCE

Data Scientist | LYNEports   06/2025 - 10/2025
Tools: Python, LangChain, RAG, Azure, Supabase, Streamlit • Internship   **Chennai, India**

- **Designed** AI pipeline converting natural-language aeronautical inputs (TLOF, FATO) into structured, regulation-compliant JSON outputs.

- **Integrated** Retrieval-Augmented Generation (RAG) to validate aviation standards, enable regulation-aware compliance automation.

- **Implemented** Large Language Models (LLMs) for schema generation, cutting manual effort by ~60% and improving data accuracy.

Data Scientist Apprenticeship | Enfocus   01/2024 - 05/2024
Tools: Python, PySpark, SHAP, CI/CD, Logistic Regression   **Ghent, Belgium**

- Developed **AI/ML pipelines** using **Python** and **Logistic Regression**, achieving **0.82 AUC** and **€10.1K annual cost savings** by classifying 2,991+ **accounts** into **20+ industries**.

- Engineered **396 features** from **136M+ logs** using **PySpark**, improving sales efficiency by **30%** through automated data imputation.

Specialist | Ford Global Technology   03/2019 - 04/2023
Tools: Python, SQL, Big Data, HDFS, Alteryx, Workflow Automation • Full-time   **Chennai, India**

- Developed a **warranty cost prediction model** that improved **forecasting accuracy by 20%**, **saving €100K annually** through data-driven warranty budgeting decisions.

- Designed **Cuscor**, a workflow automation platform, decreasing operation delays by **35%** and accelerating request resolution.

Team Lead | Bank of New York Mellon   02/2014 - 03/2019
Tools: Adobe Coldfusion, SQL • Full-time   **Chennai, India**

- Implemented a **Crisis Management System** using **Agile**, ensuring business continuity for **1,000+ teams** and saving **€200K/year**.

- Optimized SQL queries to automated **data recovery** workflows, cutting downtime by 40% and **operational costs** by **€150K**.

Software Engineer | MitrahSoft Software Solutions   01/2013 - 02/2014
Tools: Adobe Coldfusion, SQL • Full-time   **Kovilpatti, Tamil Nadu**

- Developed a **web application** for managing **10,000+ assets**, boosting productivity by 30%.

- Leveraged **agile** methodologies and effective client management to enhance operational efficiency.

EDUCATION

Master's Degree in Big Data Analytics for Business | IÉSEG School of Management
Descriptive & Predictive Analytics, Machine Learning, Deep Learning, NLP, Recommendation Systems   Lille, France • 09/2023 - 05/2025