Praveen Raj S

LinkedIn • Github

PROFESSIONAL SUMMARY

Data professional with experience in analytics, machine learning, and end-to-end data pipeline development. Skilled in Python, SQL, Excel, and Power BI, with expertise in predictive modeling, anomaly detection, and data visualization. Experienced in transforming complex datasets into actionable insights and scalable solutions that drive data-informed business decisions.

EDUCATION

Integrated M.Sc (Data Science) — 2019 - 2024
Thiagarajar College of Engineering — Grade: 9.26/10.0

Intermediate (Class XII) — 2017 - 2018
Don Bosco Matriculation Higher Secondary School, Thanjavur, Tamilnadu. — Grade: 83.75%

Matriculation (Class X) — 2015 - 2016
Don Bosco Matriculation Higher Secondary School, Thanjavur, Tamilnadu. — Grade: 94.6%

EXPERIENCE

AI/ML Engineer — June 2025 - Present
Tata Elxsi — *Chennai, Tamilnadu*

- Designed and deployed end-to-end ML pipelines (EDA, feature engineering, model training, deployment), reducing manual processing time by **30%**.
- Developed predictive models (regression, classification) to enable data-driven decisions, improving model accuracy by **20%+**.
- Built automated **collective and contextual anomaly detection services** to identify abnormal patterns in time-series and operational data, improving early issue detection and system reliability.
- Orchestrated scalable ML workflows using **Kubeflow**, reducing model retraining effort by **40%**.
- Collaborated with cross-functional teams to translate business requirements into deployable analytical solutions and performance dashboards.

  Tools: Python, SQL, Pandas, Scikit-learn, TensorFlow/PyTorch, Kubeflow.

INTERNSHIPS

Trainee — January 2024 - June 2024
Tata Elxsi — *Chennai, Tamilnadu*

- Employee Management Application Development: The objective of the project is to develop a web application for managing employees based on their project status.
- The Manager will be able to visualize the employees' project in various forms such as grid view, tile view graph view, and the manager can able to log on and modify the team information.
- Tools used**:** ReactJS, Python, Flask, MongoDB

Trainee — July 2022 - November 2022
Bonbloc Technologies Private Limited — *Chennai, Tamilnadu*

- Master Data Management for IoT: The objective of the project is to develop a web application that helps in edge data management.
- The client will be able to visualize the edge data in various forms such as grid view, tile view graph view, and the user can now able to log and modify the setup details and sensor details regarding temperature, humidity, and air pressure.
- Tools used: Angular, Typescript

PROJECTS

Chinook Music Store Analysis (Github) — July 2025

- Analyzed customer behavior by geography, spending, and purchase frequency, segmenting over 5,000 customers and identifying top segments contributing to 40% of total revenue.

- Identified top-selling tracks, albums, genres, and artists by revenue and volume, highlighting items generating 35% of overall sales.
- Evaluated regional trends to guide targeted marketing and product localization, improving engagement in high-performing regions by 20%.
- Assessed customer retention and churn patterns, providing insights to reduce churn by 15%.
- Discovered product affinities to support cross-selling and recommendation strategies, increasing potential upsell opportunities by 25%.
- Delivered actionable, data-driven recommendations for business growth and customer engagement, supporting strategies projected to increase revenue by 10%.
- Tools used: MySQL (Data Analysis), MS Excel (Visualization).

Zomato Restaurants Analysis (Github) May 2025

- Analyzed restaurant data from 15 countries using Excel to uncover insights on market competition, pricing, service models (delivery/booking), and customer preferences.
- Identified countries and cities with low competition and high growth potential, supporting strategic decisions for new restaurant openings.
- Studied market trends in pricing, customer ratings, online delivery, and table booking preferences, improving understanding of customer behavior by 30%.
- Examined the impact of price levels and cuisine variety on customer ratings, highlighting factors influencing 25% higher ratings.
- Delivered data-driven recommendations for new restaurant locations and service optimization, contributing to projected 10–15% growth in market expansion.
- Excel tools used: Pivot Tables, String Functions, Logical Functions, Lookup Functions, and Visualization.

Hyperspectral Vomitoxin Prediction (Github) March 2025

- The objective of the project is to process hyperspectral imaging data, perform dimensionality reduction, and develop a machine learning model to predict mycotoxin levels in corn samples.
- Dimensionality reduction via PCA significantly improved model training efficiency.
- Random Forest outperformed CNN and XgBoost, achieving the best prediction accuracy.
- Deep learning models, while effective required extensive tuning for optimal performancce.
- Tools used : numpy, pandas, scikit-learn, matplotlib, seaborn, xgboost, tensorflow, torch, streamlit

Food Delivery Cost and Profitability Analysis (Github) January 2025

- The aim of the analysis is to investigate all the costs related to food delivery, including indirect costs such as restaurant commission fees and customer discounts, as well as the direct costs such as delivery fees and packing.
- The purpose of the analysis is to shed light on how profitable the food delivery service as per order by contrasting these expenses with the revenue
- Tools used : Python, Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn.

CERTIFICATIONS

Foundations of Sports Analytics: Data, Representation, and Models in Sports University of Michigan January 2022

- In this course, I have learnt variety of techniques that can be used to represent sports data and how to extract narratives based on these analytical techniques

Deep Learning Fundamentals with Keras IBM January 2022

- In this course, I have learnt about some of the inspiring and motivating applications of deep learning and learnt about artificial neural networks and how they learn and optimize their parameters for a given task or application, and learnt about deep learning models such as convolutional neural networks, recurrent neural networks and autoencoders. Finally, I have learnt how to use the Keras library to build different deep learning models.

SKILLS

Computer Languages: SQL, Python, Machine Learning, Python, SQL
Data Tools: Power BI, NumPy, Power BI
Software Packages: Excel, Pandas, Matplotlib, Excel, MySQL, MongoDB
Soft Skills: Communication Skills
Others: Spreadsheet, LLM, API testing