

# Navaraja Mannepalli

Data Science & Analytics MSc Student  
Paris, France | +33 758595027 | nava-raja.mannepalli@epita.fr  
LinkedIn: [linkedin.com/in/navaraja-mannepalli](https://www.linkedin.com/in/navaraja-mannepalli) | GitHub: [github.com/navaraja20](https://github.com/navaraja20)

## PROFESSIONAL SUMMARY

MSc Data Science & Analytics student at EPITA with hands-on experience in machine learning, predictive modeling, and data analysis. Built end-to-end ML solutions during internships and academic projects. Delivered performance insights at HCL on Google Ads optimization. Proficient in Python, scikit-learn, and data visualization. Seeking data science internships in Paris/Europe.

## EDUCATION

### Master of Science in Data Science & Analytics

EPITA School of Engineering and Computer Science, Paris, France (Sep 2024 – Jun 2026 expected)

Relevant Coursework: Machine Learning, Deep Learning, Big Data Analytics, Statistical Modeling, Data Mining, AI Fundamentals

### Bachelor of Computer Applications

KBN College, India (2020 – 2024)

## PROFESSIONAL EXPERIENCE

### Process Associate – Google Ads Optimization, HCL Technologies, Hyderabad, India (Jan 2023 – Jun 2023)

- Analyzed 50,000+ daily Google Ads campaign metrics using SQL and Excel to identify underperforming keywords and ad groups.
- Automated weekly performance reports using Python scripts, reducing manual reporting time by 70%.
- Collaborated with campaign managers to optimize bid strategies, improving ROI by 18% across 12 client accounts.
- Conducted A/B testing on ad copy and landing pages, resulting in 12% increase in click-through rates.

### Amazon Intern, Online, India (Jun 2022 – Aug 2022)

- Developed customer segmentation model using K-means clustering on transaction data of 10,000+ users.
- Built interactive dashboards using Tableau to visualize customer lifetime value and churn patterns.
- Presented findings to senior management, leading to targeted retention campaigns.

### Machine Learning Intern, iNeuron Intelligence Pvt LTD, India (Jan 2022 – Mar 2022)

- Created fraud detection prototype using logistic regression and random forest on synthetic transaction data.
- Achieved 92% accuracy in identifying fraudulent patterns through feature engineering and cross-validation.
- Documented model performance and limitations in technical report for knowledge transfer.

## PROJECTS

### Self-Correcting RARE – [https://github.com/navaraja20/Self\\_correcting\\_RARE](https://github.com/navaraja20/Self_correcting_RARE)

- Developed and integrated a self-correcting biomedical question-answering system using retrieval-augmented generation (Llama-3.1-8B), SciBERT-based error detection, and T5-based answer correction pipelines, enabling high accuracy and reliability for medical queries.
- Implemented real-time error detection and correction workflows, including confidence scoring, correction validation, and audit logging in a HIPAA-compliant Streamlit web application.
- Collaborated on the end-to-end deployment of advanced analytics dashboards, batch processing features, and medical specialty filters, enhancing usability and compliance with healthcare data standards.

### Employee Performance Prediction System – [github.com/navaraja20/Employee-performance-prediction](https://github.com/navaraja20/Employee-performance-prediction)

- Built regression model to predict employee productivity scores using 6 months of HR data.
- Engineered 15+ features including attendance patterns, project completion rates, and feedback scores.
- Deployed model with 86% R-squared accuracy using scikit-learn and pandas.

### Credit Card Defaulter Prediction – [github.com/navaraja20/Credit\\_Card\\_Defaulter\\_main](https://github.com/navaraja20/Credit_Card_Defaulter_main)

- Developed classification pipeline to predict payment defaults with 89% AUC-ROC.
- Applied SMOTE for class imbalance and feature selection using mutual information.
- Visualized risk profiles using seaborn and matplotlib for stakeholder presentation.

### Data Science Portfolio – [github.com/navaraja20/dsp\\_navaraja\\_mannepalli](https://github.com/navaraja20/dsp_navaraja_mannepalli)

- Comprehensive collection of data cleaning, EDA, and ML workflows in Jupyter notebooks.
- Demonstrated reproducible research practices with Git version control.

## TECHNICAL SKILLS

Programming: Python (pandas, numpy, scikit-learn), SQL, R (basic)

Machine Learning: Regression, Classification, Clustering, Ensemble Methods, Feature Engineering  
Data Visualization: Matplotlib, Seaborn, Tableau, Power BI (basic)

Tools: Git, GitHub, Jupyter Notebook, Google Colab, VS Code  
Databases: MySQL, PostgreSQL (basic)