

Praveen Subramani

📍 Lille, France ✉ Praveen.Subramani@ieseg.fr ☎ +33 7 44 70 03 13 📧 in/PraveenSubramani88 🌐 github.com/PraveenSubramani88

SUMMARY

Data Scientist (Master's in Big Data Analytics) and **Full-Stack Developer** (5+ years) specializing in end-to-end **machine learning** systems and cloud-native solutions. Proven expertise in delivering **30%+ efficiency gains** and **€100K+ annual savings** through **predictive** modeling, **AI automation**, and scalable **MLOps** pipelines. Seeking a **Data Scientist Internship/Job** to leverage technical rigor and **innovation** in solving **high-impact business challenges**.

SKILLS

- **Programming:** Python (TensorFlow, PySpark, Pandas), R, SQL
- **Cloud/MLOps:** AWS, Azure, Docker, CI/CD, Kubernetes
- **ML Modeling:** NLP (BERT, Hugging Face), Time Series (ARIMA), SHAP/LIME
- **Data Tools:** Tableau, Power BI, PostgreSQL, Alteryx
- **Certifications:** PCEP, Tableau Specialist, Deep Learning Specialization (DeepLearning.AI)
- **Languages:** English (C1 - Professional), French (A2 - Elementary)

EXPERIENCE

Data Scientist Intern (Master's Project) | Enfocus | Ghent, Belgium | January 2024 - May 2024

- **Engineered** 396 **features** from 136M+ logs using **PySpark**, improving sales efficiency by **30%** via automated data imputation.
- **Incorporated** SHAP values to identify high-value **customer** drivers, increasing **sales** targeting accuracy by **22%**.
- **Designed** CI/CD-enabled **ML pipeline** (Python, **Logistic Regression**) to classify 2,991+ accounts into **20+ industries**, attaining 0.82 AUC and **€10.1K annual savings**.

Data Analytics Specialist | Ford Global Technology | Chennai, India | March 2019 - April 2023

- **Designed** Cuscor, a custom **workflow automation** tool, cutting operational delays by **35%** and **accelerating** request resolution.
- **Automated warranty claim analysis** for 30TB/month data using **Alteryx**, optimizing manual effort by 80% and saving **€100K/year** for 2,500+ engineers.

Tech Lead | Bank of New York Mellon | Chennai, India | February 2014 - March 2019

- **Collaborated** with **DevOps** and **QA teams** to deploy a Crisis Management System using **Agile** workflows, ensuring **business continuity** for 1,000+ teams and saving **€200K/year**.
- Automated **data recovery workflows**, cutting downtime by **40%** and **operational costs** by **€150K**.

PROJECTS

Health-Centric Recipe Recommender | Python, NLP, Azure | IÉSEG | December 2023 - January 2024

- Developed **User-Driven Collaborative Filtering** with **95.1% F1-score**, reduce unhealthy recipe suggestions by **40%** for 10K+ users.
- Integrated NLP-based preprocessing and calorie/sugar constraints, securing **99.6% recall** in ingredient substitutions.

EU Tourism Spend Forecasting | R, ARIMA, ETS | IÉSEG | February 2024 - March 2024

- Forecasted post-COVID tourism spend using **ARIMA/ETS models** (**MAE: 0.259**), quantifying **€40M pandemic impact** for EU planning.
- Optimized data distribution using **Box-Cox transformation**, STL decomposition, reducing RMSE by **50%** hyperparameter tuning.

NLP-Based Fake News Detection | Python, TF-IDF, Flask | IÉSEG | March 2024 - April 2024

- Trained SVM classifier with TF-IDF/NLP preprocessing, delivering **94% accuracy** and **92% precision** on 2,297 balanced samples.
- Benchmarked against Logistic Regression/Naive Bayes, prioritizing **F1-score** for misinformation detection in social media pipelines.

Telecom Churn Prediction | Python, SHAP, LIME | IÉSEG | April 2024 - May 2024

- Predicted churn risks with **73% accuracy** using logistic regression; identified outages as a top driver (25% risk increase).
- Identified 30% redundant features using Variance Inflation Factor analysis, enhancing model interpretability with SHAP/LIME.

EDUCATION

Master in Big Data Analytics for Business | IÉSEG School of Management | Lille, France

- Natural Language Processing (BERT, LDA), Deep Learning (TensorFlow), Cloud Computing (AWS, Azure), Time Series Forecasting