Praveen Subramani

· Praveen.Subramani@ieseg.fr · +33744700313 · in/PraveenSubramani88 · github.com/PraveenSubramani88

SUMMARY

Master's student in Big Data Analytics, specializing in predictive modeling, machine learning, and NLP. Delivers impactful data-driven solutions that reduce costs, improve efficiency, and drive business growth through actionable insights.

SKILLS

Programming & Data Tools: Python, R, SQL, Hadoop, Apache Spark, Azure Databricks, Alteryx, Git, Excel.

Machine Learning & NLP: Scikit-learn, TensorFlow, Keras, PySpark, BERT, LDA, Sentiment Analysis.

Data Visualization: Tableau, Power BI, Streamlit, Seaborn.

EXPERIENCE

Data Scientist | Enfocus | Ghent, Belgium | January 2024 - May 2024

- Architected and implemented a **real-time customer segmentation pipeline** using PySpark and Scikit-learn, reducing data processing time by 40% and enhancing decision-making speed.
- Developed a robust customer segmentation model with **Logistic Regression** and **Gradient Boost Trees**, achieving a **Test AUC of 0.75** across **136 million user logs**.
- Engineered **450 features** and automated segmentation for **7,210 users**, improving scalability by **30%** and reducing manual effort.
- Applied **SHAP analysis** to enhance model interpretability, leading to a **15% increase** in marketing **ROI**.
- Delivered actionable insights that **reduced marketing costs** by 21% and improved lead conversions by 18%.

Specialist | Ford Global Technology & Business Center | Chennai, India | March 2019 - April 2023

- Built a **real-time warranty application** using Hadoop and Alteryx, slashing claim processing time by **30%** and enhancing root cause analysis.
- Designed a warranty dashboard that improved data visualization and reduced **customer complaints by 20%**.
- Optimized workflows and managed client interactions, increasing **productivity by 40%**.

PROJECTS

Data Science Projects | IESEG | September 2023 - May 2024

- **NLP-Based Fake News Classification:** Constructed an NLP model using **TensorFlow**, **BERT**, and **LDA** for fake news classification, attaining an AUC of 0.91. Enhanced content moderation by minimizing misinformation spread.
- **Supply Chain Optimization:** Established a predictive model using **PySpark** and **Apache Spark** to refine supply chain strategies, securing an **AUC of 0.89**. Streamlined product recommendations and operations.
- **Personalized Recipe Recommendation Engine:** Designed a recommendation engine using Python with **collaborative filtering** and **content-based techniques**, obtaining an **NDCG score of 0.85**. Improved user engagement through personalized recipe suggestions.
- **Accident Severity and Sentiment Analysis:** Built **CNN** and **LSTM models** using **TensorFlow** and **Keras**, achieving **79% accuracy** in accident severity prediction and **84% accuracy** in sentiment analysis. Improved safety and real-world applications.
- **U.S. Part-Time Job Trends Forecasting:** Analyzed U.S. part-time job trends (1994–2024) using R with **ARIMA**, **ETS**, and **STL models**, identifying seasonal and trend-based influences. Provided actionable insights for policymakers.

EDUCATION

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

- Predictive Modeling, Data Visualization, Cloud Computing, Time Series Analysis, Deep Learning, NLP, Machine Learning.

Bachelor of Technology in Information Technology | Dhanalakshmi College | Chennai, India

CERTIFICATIONS

Tableau Desktop Specialist | Tableau | 2023

PCEP – Certified Entry-Level Python Programmer | Python Institute | 2023

Deep Learning Specialization | DeepLearning.AI | 2024