Babacar GUEYE
CROUS
Caen / Paris, France
Tel: +33 6 64 12 73 08
Tel Whatsapp: +221 77 303 71 28
Email: gueyebaba159@gmail.com

EDUCATION

Sept 2020 – Sept 2023

Master 1 & 2 Applied Statistics and Data Science at University of Caen Basse-Normandie, Caen, France

Relevant courses: Advanced Probabilities, Statistics, Machine Learning, Optimization, Time Series, Regression Models, Data Analysis, Web Analytics, Business Intelligence, Fundamental Statistics, Operational Research, Probabilities, SQL, Data Base.

Relevant Projects & Technos:

- 6-month research project: Built algorithms to detect patterns and predict stock prices, developed trading strategies based on those algorithms (ARIMA MODELS / Time Series).
- Decent Applications using SQL for databases, Talend for integration and Data Lake models, SAS for statistics applications. Advanced Statistics tools using Rstudio (Probability laws, Statistic tests(Anova, Khi2, Student, Pearson), Supervised and Unsupervised Learning).

Sep 2015 – Jun 2020

Licence (Bachelor) Applied Maths & Computer Science at University of Caen Basse-Normandie,
Relevant courses: Economics, Data analysis, SQL, Database, Python, Discrete Mathematics, Advanced Probabilities.

Sep 2012 – Jun 2014

High School at Ecole Saint Pierre, Dakar, Senegal
High school degree in Mathematics and Physics (Baccalaureate S1) - with honors.

EXPERIENCE

Oct 2022 - April 2023

Internship at ORANGE SONATEL (BUSINESS SERVICES) as Data Scientist
Within the Marketing Department, i work on Geomarketing Monitoring and Value Management. Key Missions :

- Setting up reportings for the business KPIs
- Setting up a Machine learning (K-means) algorithm for Customer Segmentation and churn prediction
- Setting up Surveys and analysing data from Business Tours around the Geomarket

Oct 2021 – Feb 2022

School Projects :
Analyse de données AgroAlimentaire : The objective of this study is mainly to explain the productivity according to the characteristics of the fruit, highlight fruit typologies and set up predictive models.
ARIMA MODELS / Time Series : Prediction of stock market indices With the ARIMA model.
PCA (Principal Component Analysis) model in details using SAS program.
Methodologies : Descriptive and exploratory analysis, hypothesis testing, PCA, MCA, Classification, Random Forest, Regression model. ARIMA models (p,q,d)
Tools : Rstudio, Power BI, Rmarkdown, Yahoo finance, Rstudio.
Kaggle Training Competition Projets : Supervised & Unsupervised Learning using Machine Learning techniques.
Housing Price Prediction : For each Id in the test set, we predict the value of the Sale Price variable. The evaluation metric is Root-Mean-Squared-Error (RMSE) between the log of the predicted value and the log of the observed sales price. **My results led me to top 11%.**
Credit Card Fraud Detection : This dataset presents 492 frauds out of 284,807 transactions. Eval metric is the Area Under the Precision-Recall Curve (AUPRC). Confusion matrix accuracy is not meaningful for unbalanced classification. **Fraud detected with 99% accuracy.**
Methodologies : Descriptive and exploratory analysis, hypothesis testing, Random Forest, Gradient Boosting, Decision Trees, Classification.
Tools : Rstudio, Power BI, Rmarkdown, Jupiter Notebook, Python, libraries (Numpy, Pandas, Sklearn, Keras, Matplotlib).

SOFT SKILLS

Languages: French: Native, **English**: Fluent, **Wolof** (a Senegalese language): Native.
Hobbies: Football, Video Games.

ASSOCIATIVE WORK

Treasurer at AESC (Association des Etudiants Sénégalais de Caen)
Organized events to promote African culture on the campus and gatherings with Alumni students