# AKSOY CEBRAIL

06 52 64 01 15 | cebrailaksoy01@gmail.com
**Portfolio** : https://cebsmind.github.io/portfolio/
Bron 69500

Experienced statistical analyst and data scientist with **2+ years of expertise**, possessing a strong background in mathematics and a passion for AI.

## SKILLS

**Programming Languages & Tools:**

- Python (Pandas, NumPy, Matplotlib, Scikit-Learn, TensorFlow)
- SQL (MySQL, PostgreSQL)
- SAS
- Tableau
- Big Data: PySpark
- Cloud: AWS, GCP, Azure
- Excel

**Statistical Analysis & Machine Learning:**

- Advanced Statistical Analysis and Modeling (linear regression, logistic regression, ANOVA, multinomial models, etc.)
- Experience in creating and deploying Machine Learning models (supervised/unsupervised learning, classification, clustering, etc.)

**Advanced Techniques:**

- Proficient in Deep Learning methodologies
- Familiarity with Natural Language Processing (NLP) techniques

## WORK EXPERIENCE

**DATA SCIENTIST –** University Gustave Eiffel **–** Bron 69500 Sept 2021 – Present

- Worked on the research project "**Reg-Trauma2**," aiming to leverage **hospital data** from Rhône and **Police data** to **estimate** the number of accidents and injuries on national roads.
- Created Police data tables / **generated identifiers** for accidents, locations, vehicles, and users / **corrected** variables / **filtered** data / ensured data coherence.
- **Automated** the preparation (correction / adaptation of new modalities / imputation, etc.) of both data sources for statistical analysis.
- Tested and developed various **logistic models** to predict severity using the intersection of the two data sources, model **validation**, and subsequent **estimation** of severity on a national level.
- Aggregated data and developed a **Capture-Recapture** model to extract correction coefficients for estimating the total number of national injuries by profile.
- Programmed in **SAS / SQL / Python** and utilized **SAP BI** for querying Police data.

**MATHEMATICS TEACHER–** Collège FAUBERT– Villefranche sur Saône 69400 Dec 2020 - Aug 2021

- Taught **mathematics** to middle school students (4th and 3rd grade levels) and prepared students for the Brevet des collèges.
- Developed and implemented innovative pedagogical projects.

## PERSONAL PROJECTS

**DEEP LEARNING IMAGE CLASSIFICATION –** Personal Kaggle Project

- Used **Python** to develop a **Convolutional Neural Network (CNN)** architecture using **TensorFlow's** Keras API to classify images of cats or dogs.
- Preprocessed data using **ImageDataGenerator**: resizing, normalization, image transformations, etc.
- Implemented a **pre-trained model** (VGG16 based on 1 million images) to significantly enhance our model (**Transfer Learning**).
- Developed **scraping** techniques to obtain a diverse dataset essential for testing and evaluating the model's performance on new random images, using the Unsplash **API**.

**HOUSE PRICES - ADVANCED REGRESSION TECHNIQUES –**Kaggle Competition

- Created an **XGBoost** regression model (**Ensemble Learning**) to predict Boston house prices.
- Implemented a **Pipeline** for data preprocessing (encoding, imputation, normalization, etc.).
- Generated new variables (**feature engineering**) to enhance the model's performance.
- Submitted predictions on Kaggle, achieving a **top 19%** ranking.

More projects available in my portfolio HERE!

## EDUCATION

**MASTER IN APPLIED MATHEMATICS IN DATA SCIENCE**– Claude Bernard University Lyon 1 – LYON. 2020
**BACHELOR'S DEGREE IN GENERAL AND APPLIED MATHEMATICS** – Claude Bernard University Lyon 1 – LYON. 2018