Sitraka FORLER

 Sitraka17 | Sitraka | sitraka.fr | sitrakaforler@gmail.com | +33 6 31 27 08 50

DATA SCIENTIST

Through my years of professional experiences I was able to progress in SQL and NoSQL(to retrieve data), Python (in order to exploit them) then Dataiku and PowerBI (to share insights). To realize various projects, from simple dashbording to Machine Learning for an Research and development project in NLP. All projects were made in IT teams using GitLab (CI/CD).

This CV can also be seen on GitHub. For details, click here.

WORK EXPERIENCE

Experienced Consultant in Data and Digital Transformation @Initio   Sept 2022 - present

Writing specifications and improving processes Automating a workflow to save staff time Design and development of test automation scripts at each level of the test pyramid .

Automating front-end and back-end tests. Creating and developing the test automation framework, test data. Development and implementation of processes and approaches for effective test automation. Integrating automated tests into CI/CD pipelines.

Integrating automated tests with test case and defect management tools. Integrate manual and automated testing into quality assurance practices to follow compliance department instructions.

Communicate with project teams on strategy improvements, share experiences, best practices and lessons learned.

Tools used: Python, PowerBI, AWS Redshift

Data Scientist @ VirtualExpo   Fev 2021 - Sept 2022

I fine-tuned a BERT algorithm in order to summarize products descriptions for the B2B platform of VirtualExpo. I was also able to create dashboards thanks to Qlik and Dataiku in order to follow up the sales and Web traffic. Finally, I had the opportunity to work on Churn prediction model. Carry out audits on various projects with the quality team. Identify, analyse and refocus anomalies, deviations and non-conformities, then propose preventive and corrective actions to resolve them. Set up indicators and monitoring dashboards.

Intervene on Quality indicators when they are inconsistent. Carry out Quality reports.

Participate in the drafting of procedure adjustments (internal and outsourcing). Monitoring of outsourcing production.

Monitoring and analysis of KPI's for management.

Monitoring and processing of deviations (online files without products, status conflicts, etc.)

NLP via HuggingFace to classify product texts and generate qualitative summaries facilitating the customer journey while increasing SEO. Prediction of the best distributors. (TensorFlow)

Implementation of Apache Spark within the DATA team to improve the performance of the division.

Tools used: Python, HuggingFace, BERT, MongoDB, Qlik, Dataiku, AWS EMR

Data Analyst @ STMicroelectronics, Internship   June 2020 - Nov. 2020

- Implementation of two dashboards on the performance indicators of the production line via TIBCO-Spotfire. Integration of the data and visualization of the latter processed (R, .net and html).
- Implementation of the automation of the sending of service documents to the Industrial Engineers and to the Managers of the Fab.

Tools used: TIBCO Spotfire, R, .net, C and SQL

Junior Data Analyst @ JDA , Alternance   Oct. 2018 - June 2020

Data Cleaning of databases for different clients. Dashboard creation via RShiny for CorsicaFerries and AltaEthics. Also proposed ML predictions with simple methods thus as: OLS, Multilinear Regression,Logistic Regression and Decision Tree.
Tools used: Python, Excel, VBA

EDUCATION

2018-2020 Master of Science in **Quantitative Economics**   (GPA: 3.7)
Major in Big Data, minor in Quantitative Finance in partenership with Centrale Marseille link.

2016-2018 Bachelor of Arts at **Economics and Management**
Major in Quantitative Economics (Econometrics)

2014-2016 PACES **Faculty of Medicine of Marseille**
Biology, Biophysics,Fluid dynamics, BioStatistics

SKILLS

Soft Skills	Pragmatic, Enthusiastic, Team player
Technical Skills	Machine Learning, Econometrics, Python

CERTIFICATIONS

Machine Learning	The most famous course proposed by Coursera and Standford. Learning about queuing to gradient descent. link.
Advanced Statistics for Engineers	Descriptive and inferential Statistics. Econometrics with cross-sectional data, time-series data, pooled cross-sectional data and longitudinal data. link.
Dataiku Core Designer	Dataiku is the world's leading platform for Everyday AI, systemizing the use of data for exceptional business results. link.
Major Engineering Project Performance	Understand the key phases of a project life cycle. Writing documentations and directives. Implementing Cost Benefits Analysis method. link.
Digital Marketing Fundamentals	SEO and SEM basics tought by Google link.
Dataiku Core Designer	Dataiku is the world's leading platform for Everyday AI, systemizing the use of data for exceptional business results. link.

Click here to see all certifications.