GUILLAUME ORSET-PRELET

+33 6 10 80 32 75 • GUILLAUMEORSETPREL@GMAIL.COM • LinkedIn • Portfolio

Data Scientist Consultant with commercial experience in Natural Language Programming (NLP), Software Engineering and Cloud development. Trilingual in English, French and Spanish I seek a meaningful career developing ML and AI based products for a progressive organization leveraging my general, technical and managerial strengths.

Education

CY-Tech: Programme Grande Ecole (2018-2022), *Master of Artificial Intelligence,* Cergy, Ile-de-France
Grenoble School of Management: Programme Grande Ecole (2018-2022), *Master of Management,* Grenoble, Isère
French Baccalaureate Distinction: best honors (summa cum laude)
Polytechnic University of Valencia: Academic Exchange, (January - July 2021), *Computer Science*, Valencia, Spain

Technical Work Experience

SIA Partners, *Data Scientist Consultant,* Paris, France / **L'OREAL**, *Gen AI Software Engineer*   May 2023 - now

- Implemented the first Generative AI POC at L'OREAL to allow chemists to extract information on research papers more efficiently. Created the entire back-end using a software architecture to deploy a Search Engine based on Arxiv, Google-Scholar, PubMed and a Q&A interface based on LLM (PALM) to chat with documents. Collaborated with the Head of AI at L'OREAL using LangChain to make the code modular and compatible with any LLM and Vector Store (e.g. Vector Search Google). Deployed this package for L'OREAL's AI team of 6 persons.

- Developed a new Q&A tool leveraging L'OREAL's documentary database containing over 5 million documents using the aforementioned package. This new tool allows chemists to "chat" with the articles and isolate specific information among the set of relevant documents. The LLM quotes each source and the tool displays a preprocessed summary to provide additional context.

Unify Group (gathers the digital activities of TF1, the 1st French TV channel), *Machine Learning Engineer,* Paris, France
September 2021-September 2022

- Independently fine-tuned a summarization model based on BARThez to provide abstracts of online media. Used Pytorch environment for fine-tuning and industrialized the project through a Kubeflow pipeline using Kubernetes. I obtained a ROUGE1 score of 33.68 and a BERT score of 72.32. Wrote a *Medium* article about it: https://medium.com/@guillaumeorset/.
- Implemented CI/CD for 3 projects to automate testing, documentation and releases using GitLab.
- Developed a new content-based recommender system for *Marmiton* with two data scientists to offer cooking recipes based on the ones visited by a user. Deployed the model using Google Cloud Platform (GCP).
- Conducted A/B testing for a recommender system, improving the click rate by 300%.
- Worked on churn prediction to maintain our subscribed users using LightGBM and XGBoost models.

SNCF Voyageurs, *Data Scientist (work-linked training),* Lyon, France   September 2019-September 2020

- Developed a Python passenger flow forecasting model linked to on-board counting technologies to get counts for 90% of the trains against 20% before.
- Wrote the specifications of the new data-visualization tool and mock-up realized in WEB language (HTML, JS, CSS) to help business units create the actual tool in Power BI.

Extracurricular Activities

AggregEstate, *Founder - Django Application,* Paris, France   January 2021- April 2022
- Created a MySQL database fed by web scraping scripts to get as many real estate ads as possible from various companies (JLL, CBRE, KAUFMAN & BROAD). Normalized all metrics to get a usable database and to provide insights for each request on the website like the average price per square meter in the search area.
- Created a web platform in Django managing the backend in Python and the front-end in HTML, CSS, JS. Connected the MySQL database to the web platform through Django to make all the real estate ads available on the website.
- Created an inquiry system based on GPS coordinates using Google Maps API.

Skills & Interests

Programming: Python, Git, GCP, Kubeflow, Kubernetes, Spacy, Pytorch, TensorFlow, CI/CD, Transformers models, Docker, DevOps, LangChain, LLMs
Databases: BigQuery, MySQL, Postgres, Vertex AI MatchingEngine (Vector Store)
Web Development: HTML, CSS, JavaScript, Flask, BeautifulSoup, SalesForce, Web Scraping, Selenium, Django, Streamlit.
Interests: tennis, diving, films, skiing, soccer, biking, architecture, gastronomy, traveling
Languages: native in French, fluent in English and Spanish