

Juan Manuel Zapata

Santiago, CL | juanmazm9@gmail.com | 9 4015 2597 | linkedin.com/in/juanma-zapatam

github.com/JZapataMA | Portafolio

Professional Summary

Professional with studies in Data Science Engineering, possessing excellent communication skills and a proactive attitude, greatly benefiting both individual and team work. Specialized in the development of **Machine Learning** and **Deep Learning** models, as well as in the processing and analysis of **big data** with a focus on scalability. I have a strong understanding of machine learning theory, advanced knowledge in **statistics**, and extensive experience in implementing solutions based on LLM models and data processing techniques. My career spans from algorithm exploration and optimization to generating actionable insights to solve complex problems in various contexts.

Skills

- **Advanced** English proficiency.
- Development of machine learning and deep learning models using **TensorFlow**, **Scikit-learn**, and **PyTorch**.
- **Advanced** proficiency in **Python**, **R**, and libraries such as **Pandas** and **NumPy** for data analysis and manipulation.
- **PySpark** intermediate knowledge
- Google BigQuery
- Solid knowledge of **Bayesian statistics** and **data structures**.
- Design of recommender systems (**RecSys**) and customized algorithms for large datasets.
- Management of relational databases with **PostgreSQL** for design, querying, and optimization.
- Competence in low-level programming with **C/C++** and graph-based analysis.

Education

| | |
|---|-------------|
| Pontificia Universidad Católica de Chile , Bachelor's Degree in Data Science Engineering | 2021 – 2024 |
|---|-------------|

Experience

| | |
|---|---------------------|
| AI Analyst , Gather Consultores – Santiago, CL | Oct 2023 – Mar 2025 |
|---|---------------------|

- Responsible for developing integrated applications that use AI models for classification, detection, recommendation, and generation tasks, creating tools that optimize resources and streamline daily processes.
- Participation in various events and training sessions, including the **IBM TechxChange 2024**, held in Las Vegas.

Projects

Identification and Recognition of Informal Settlements

- Use of low-quality satellite images to identify informal settlements.
- Classification of census blocks through clustering to determine sector ratings and potential expansions using neural networks with **Pytorch** and **Scikit-learn**.

User Pair Analysis and Search

- Processing of a database of tweets related to the 2020 Chilean new constitution project.
- Development of a search algorithm to identify and classify users with similar posts using **BigQuery** and **Local Sensitive Hashing** techniques.

Identification of Human Resource Duplication

- Design of an algorithmic scheme to identify duplicate projects in a public institution.
- Optimization of evaluation criteria in co-financing processes, improving the accuracy and efficiency of application analysis with **Pandas** and **Numpy**.

Song Recommendation System Based on Lyrics

- Processing and vectorization of string variables using **transformers** to generate song representations.
- Algorithmic development for an item prediction system based on a given list using **Pytorch**, **Scikit-learn**, and models available on **Hugging Face**.