Ignacio Javier Mermet | **in** javier-mermet | **⌗** CrossNox

📍 Colegiales CP1426, Buenos Aires
📞 (+549) 1123478174
✉ ijmermet@gmail.com
✉ imermet@fi.uba.ar
Last updated: 2024-08-29

I'd defne myself as a full stack data developer. While my career started as a backend developer, I quickly turned into Data Science. Throughout my career I've contributed as a Data Engineer, MLOps/ML Engineer and just overall, a data hacker. My education orientation is on distributed systems, with attention to the intricacies of low level and scalable architectures. I care about solving problems now while setting the ground for the future success of the projects I'm involved in.

I'm a curious learner, a nonconformist always looking to improve myself and my toolset. I bring over 8 years of experience both in consulting/services and product companies, as well as freelancing. Having been a TA at the top three universities in Argentina helped me improve a great deal my soft skills and group management capabilities.

TECHNOLOGY SKILLS

This is a non exhaustive list of technologies I've worked with and feel comfortable using.

Programming Languages: Mastery of Python
Operations Technologies: git, aws, azure, gcp, apache airflow, terraform
Database Technologies: PostgreSQL, MySQL, Redis, Hive, PrestoDB, Pinecone, pgvector
Backend Technologies: FastAPI, Flask
Data Science Stack: Pandas, scikit-learn, Keras, gensim, matplotlib, pyplot, seaborn, PySpark, pytorch
Software Engineering: OOP, agile methodologies, design patterns

EMPLOYMENT

Sr. AI Engineer
August 2024 - **Currently**
HeyAtlas

At Atlas, I'm currently working on forecasting problems using time series and econometrics techniques. On the side, I'm building a data warehouse to integrate data from several sources that will allow us to gain insights into the full behavior of the userbase on the platform.

Machine Learning Engineering, Optimization Scientist
2020 - **Currently**
Freelancer

I've developed several data projects end-to-end.

I'm currently working on an agentic RAG system to help an investments bootcamps scale to over 10K students per opening, providing reliable answers to the most common questions and aiding tutors with the volume of messages in the forums.

Recently, a RAG system for largest spanish talking medical channel on YouTube to create an AI that replies based on their videos.

Created the fixture for the LNB/LAB/LFA basketball leagues in Argentina for the 2023/2024 season. This is a particular instance of the Traveling Tournament Problem,

regarded as one of the, if not the, most difficult problems in combinatorial optimization.

Developed end-to-end data pipelines to predict employee issues at Lara AI prior to joining.

Developed a custom model using drone and satellite imagery to optimize crops in the northern region of Argentina.

TA - Data Engineering
January 2024 - August 2024

Universidad Torcuato Di Tella

The course covers extensively relational databases and big data technologies, using both SQL and NoSQL databases, as well as distributed data processing. My role is to give lectures, prepare materials, guide and help students throughout the semester.

Sr. Machine Learning Engineer
May 2023 - July 2024

Lara AI, Buenos Aires

At Lara, I cover the grounds for everything Machine Learning. I've trained and put into production tabular models for employee churn prediciton, RAG using GPT models on Azure and Pinecone indexes. I also take care of our Data Engineering tasks, making data easily available with robust pipelines.

TA - I200 - Algoritmos y Estructuras de Datos
December 2022 - December 2023

Universidad de San Andrés

The course focuses is on low level programming of data structures using C. We also cover topics such as memory management, time/memory complexity of algorithms, graphs. I was in charge of one of the practical slots of the course, making focus in applying the theory students learn.

TA - 75.06 Organización de Datos
September 2020 - March 2022

Universidad de Buenos Aires

This is an introductory course to Data Science and Machine Learning. I actively created material, maintained the infrastructure for the course, lectured classes, prepared notebooks for reference and helped with grading of practical assignments and mid terms.

Freelance - Machine Learning Engineer
March 2020 - April 2024

Sports Analytics World

Help building the market of sports analytics in the FIBA world. I was been part of the team fetching and consolidating several data sources, in a scalable, automated pipeline. I've helped with several ad-hoc reports and analysis for teams and coaches.

Data Scientist - Machine Learning Engineer
November 2019 - May 2023

Mutt Data

Several consultancy projects for adtech, fintech, marketplaces and other bussinesses. Development of automatic data pipelines, from data ingestion and processing to machine learning models training, deploy and monitoring. Usage of IaC on AWS for scalable resource management.

Worked as Lead Dev for over a year, dealing with customer's needs, prioritizing tasks and making architecture and technological decisions for the team.

Some side projects include Mutt Data's blog development and maintainance, ocasional talks to distribute knowledge across teams, management of Gitlab runners on AWS

with Terraform and leading the development of muttlib, which serves as a playground for onboarders to learn our stack and methodologies. Worked on the revamp of the code challenge for interviews and took interviews for a few months, to evaluate candidates' potential. Sole developer of `muttplotlib`, a wrapper on top of `matplotlib` branded for Mutt Data.

Spearheaded the Research Squad, focusing in serving state of the art generative models through the best MLOps practices.

Data Scientist
October 2018 - July 2019

Jampp, Buenos Aires
Worked on a behavioural lookalike model to segment audiences across a market, built to be highly scalable and run on demand. RTB open market fraud analysis.

Data Developer
October 2017 - October 2018

7puentes, Buenos Aires
Python/Ruby developer for Big Data and Data Science projects.

Improved the rate of empty queries by about 60% on an e-commerce app using word embeddings to enrich it's keywords. Built multi-label text classifiers using neural networks, improving previous bayesian models' performance by 5x. Worked on customer's data analysis in order to predict churn probability and behavior patterns. Acquired knowledge on data mining and web data extraction.

Full Stack Software Developer
June 2016 - October 2017

C&S, Buenos Aires
.Net developer. Worked in teams developing backoffice webapps and highly-scalable, concurrent, RESTful web APIs. Collaborated on internal frameworks to simplify the process of building webapps.

Ad Honorem TA - Algorithms & Programming II - Wachenchauzer Course
2016
FIUBA, Buenos Aires
The course covers topics such as time and memory complexity, data structures and programming techniques using C. Helped assigned students through their assignments and gave some practice classes.

Ad Honorem TA - Algorithms & Programming I - Méndez Course   2016
FIUBA, Buenos Aires
Theory and practice lessons. The course is oriented towards teaching freshmen students about control structures, how to work with files and basic programming techniques.

EDUCATION   Software Engineering, **Universidad de Buenos Aires**
Expected graduation date: late **2024** (final project on computer vision for sports analytics ongoing)

AI for Medicine, **Coursera**
Certificate

2011, Pre-International Math Olympiads
2009, XV May Math Olympiads Honorable Mention

LANGUAGES   Spanish, native level

English, professional level
FCE - B2 exam passed with merits

OTHER ACTIVITIES

Collaborating Student
LABDAC, FIUBA, Buenos Aires
Co-gave a talk on word embeddings and unsupervised translation. Helped develop and manage several data science related research projects.

TALKS

Airflow para Data Scientists
PyConAr, 2019
Introduction to airflow for ETFL workflows.

Behavioral Lookalike Model for App Advertising
Nubank Data Science & Machine Learning (São Paulo, Brasil), 2019
Building a Behavioural Lookalike Model: from PoC to prod.

PERSONAL PROJECTS

BicisBA
Add predictive capabilities to Buenos Aires' bikes system. Let users find the optimal bikes station near their location to ensure they get a bike to travel the city.

YouConfigMe
Python library to manage configuration in a pythonic fashion.

PySpark exercises
Repository of solved PySpark exercises for FIUBA course 75.06 Data Organization. Using jupyter notebooks, the objective is to show different strategies and concepts associated with the Map-Reduce paradigm.

fsqrt
Numerical method to calculate an IEEE-754 float number's square root. The method uses a magic constant derived based on their bit representation.

m2r2
m2r actively maintained fork. Markdown to rst converter, useful for sphinx. Over 120k downloads per month as per this writing.

INTERESTS

natural language processing, data science, network analysis, education, optimization, machine learning