Mabel Villalba Jiménez

Data Engineer & Team Lead

Scalable Data Solutions, Machine Learning, and Agile Leadership

mabelvj@gmail.com

LinkedIn GitHub Website StackOverflow

skills

Cloud Computing: AWS (Lambda, Glue, S3, CloudFormation, Kinesis), Microsoft Azure

Agile & Project Management: Agile, Scrum, Team Leadership, Stakeholder Engagement, Roadmap planning

Leadership: Team Management, Mentoring, Career Development, Performance Review

Programming: Python, R, SQL, Unit Testing

Data Engineering: ETL Pipelines, Pandas, PySpark, MongoDB

AI & Machine Learning: Scikit-learn, TensorFlow, Deep Learning, Time Series Analysis

Software Development: Docker, Docker Compose, CI/CD

experience

Data Engineer **SEAT:CODE** ***(April 2021 - Present)***

Designed and maintained cloud-based **data infrastructures**, supporting large-scale analytics.

Developed and deployed real-time **ETL workflows** using AWS Lambda, Kinesis, Glue, and S3.

Optimized **SQL performance** and data modeling, fine-tuning queries for efficiency.

Established **data validation processes** and governance frameworks. Integrated **CI/CD pipelines** to automate deployment and testing of data workflows.",

Enhanced system **monitoring** and reliability by implementing logging and alerting solutions.

Implemented **predictive maintenance solutions** using machine learning models and developed AWS Quicksight dashboards por visualization, providing **real-time insights**.

Team Lead **SEAT:CODE** ***(May 2023 - November 2024)***

Led a **multidisciplinary team** of data engineers, data scientists, facilitating cross-team discussions and agreements.

Owned the full **lifecycle of PIBOD**, collaborated with the Product Owner to defining its roadmap and technical feasibility and coordinated tis **ramp-down**.

Conducted **1:1 mentoring sessions**, career planning, and guided team members using Profile Ladders, as well as executed **performance evaluations**

Coordinated the definition and tracking of team **OKRs**, along with the Business Unit's.

Participated in **hiring processes**, ensuring a diverse and talented team.

Data Science Engineer **Mática Partners** ***(December 2020 - April 2021)***

Designed and managed scalable ETL pipelines using PySpark and Docker, processing datasets with millions of records efficiently.

Deployed solutions on AWS and Azure, achieving seamless cloud integration for clients.

Automated CI/CD workflows with Jenkins, reducing deployment times by 40%.

Python Developer **Zyte** ***(April 2019 - December 2020)***

Developed advanced web scraping solutions using Scrapy, enabling high-volume data extraction for Fortune 500 clients.

Optimized ETL pipelines with Pandas and SQL, improving data processing speeds.

Integrated Celery for distributed task management, processing over 1M tasks daily.

Implemented robust monitoring and debugging systems with Sentry.

Data Scientist and Python Developer **Independent Contractor** ***(May 2017 - April 2019)***

Developed algorithmic trading systems on Quantopian for a private fund.

Created a TV attribution and ROI reporting system for Windsor AI, leveraging R and PostgreSQL for real-time insights.

Built scalable web scrapers for LISTedTECH, automating updates for large educational databases.

Mentor and Reviewer **Udacity** ***(April 2017 - May 2019)***

Mentored over 120 students in Data Science and Machine Learning.

Reviewed 750+ projects across Deep Learning and Reinforcement Learning, maintaining a 4.93/5 feedback rating.

Predoctoral Researcher **UPC** ***(October 2015 - October 2016)***

Designed Python and MATLAB simulations for optical devices, achieving a 54dB rejection on sidebands for optical networks.

Contributed to the development of wavelength shifters for next-generation optical communication systems.

education

MSc in Photonics Polytechnic University of Catalonia (UPC)

Telecommunication Engineering (BSc + MSc) University of Malaga

Data Analyst Nanodegree Udacity

Machine Learning Engineer Nanodegree Udacity

Machine Learning Course Stanford University, Coursera