Alan Martín López Chávez

Data Scientist | ML Engineer | Data Engineer | GCP | MLOps

alan.m.l.ch.1997@gmail.com
https://www.linkedin.com/in/alanmlc/
México City (open to work remotely).
+526391047301

About me.

Results-driven Engineering Physicist turned Machine Learning Engineer and Data Scientist, with expertise in designing scalable data solutions, deploying machine learning models, and optimizing end-to-end pipelines. Proven track record in banking, retail, and AI industries, with a passion for bridging technical innovation with business impact. Adept at leading cross-functional teams, automating workflows, and transforming data into actionable insights.

Professional Experience

MLOps Engineer
Gentera (Banking)
03/2024 – present. CDMX. Hybrid.

- Designed and deployed end-to-end MLOps pipelines on GCP for fraud detection models, leveraging Kubeflow, Vertex AI Pipelines, and Airflow to automate model training, validation, and deployment.
- Implemented a dynamic model versioning system in Airflow DAGs to automatically track, evaluate, and promote the latest production-ready models, reducing manual oversight by 50%.
- Built a retraining pipeline using Vertex AI that triggers based on performance drift and model performance (MAE), or scheduled intervals, improving model accuracy by 15% over static deployments.
- Developed a Pub/Sub-driven Cloud Run service to handle real-time model update notifications and operational alerts, enabling faster response to critical issues.
- Engineered CI/CD workflows with Jenkins to automate testing and deployment of ML artifacts, reducing release cycles from months to just days.
- Containerized training/scoring environments with Docker and orchestrated them via Vertex AI Pipelines, ensuring reproducibility across dev/qas/prod, using Apache Airflow to schedule the process using a trigger that detects changes in the data source on BigQuery.
- Implemented unit/integration tests (pytest) for Airflow DAGs and Vertex Pipelines, catching 90% of errors pre-production for Vertex Pipelines, and 100% from Apache Airflow.
- Collaborated with risk analysts to replace manual Excel-based workflows with an automated Python ETL pipeline, reducing reporting delays from weekly to just an hour.
- Translated business requirements into a configurable forecasting system (Python + BigQuery + Apache Airflow) that allows non-technical teams to adjust parameters dynamically via a Excel into a Cloud Source Repository.
- Established ML metadata tracking (Vertex ML Metadata) for audit trails, improving compliance with banking regulations.
- Conducted performance benchmarking of fraud models across customer segments, identifying and resolving latency bottlenecks in scoring pipelines.

Data scientist.
ECOSAT S.A DE C.V. (Retail/IoT)
05/2023 - 03/2024. Chihuahua, Chih.

- Established the foundation for the data science team as the company's first data scientist and implemented a company-wide data-driven strategy.
- Developed and deployed ML models, including:
    - An anomaly detector with one-class SVM, reducing false positives by 70% and optimizing alerting tools using Apache Spark, using Random Search Cross Validation.
    - An end-to-end pipeline using random forests to predict customer churn, integrating a RESTful API with Flask for push notifications that improved app retention.

- Designed solutions to reduce IoT device repair times from two months to two days by leveraging data insights.
- Restructured databases and data warehouses using Azure Data Factory, MS SQL Server, and MySQL to improve system performance.
- Automated ETL pipelines with Python, Dagster, and dbt, streamlining data workflows and improving efficiency.
- Created Power BI dashboards for customer-facing web apps, providing actionable insights that drove strategic decisions.

Web Scraping Parsing Engineer & Data QA Team Lead
Scale AI. (AI Industry)
08/2021 - 05/2023. Remote.

- Promoted for 3x productivity: Built scraping scripts (Node.js, Puppeteer) for 30+ websites/week.
- Led QA teams to ensure >99% data accuracy for AI training datasets.
- Automated testing processes for reducing validation time by 60%.
- Automated testing processes to ensure efficient and accurate web scraping, adhering to best coding practices.
- Built large-scale datasets for machine learning pipelines while maintaining version control using Git and GitHub.
- Collaborated with cross-functional teams to improve AI pipeline efficiency and ensure data integrity throughout the process.

Assistant university professor.
Universidad Autónoma de Chihuahua.
2019 - 2020. Chihuahua, Chihuahua.

- Designed curricula for calculus, algebra, and physics for engineering programs.

Education

Bachelor's Degree in Engineering Physics.
Universidad Autónoma de Chihuahua.
2015 - 2019. Chihuahua, Chihuahua.

- Computational physics with C, Matlab, and Python.
- Descriptive and inferential statistics.
- Advanced mathematics, including tensorial algebra, partial differential equations, and vector calculus.

DeepLearning.ai.
2023. Coursera.
Deep Learning Certificate.

https://www.coursera.org/account/accomplishments/specialization/MTGHUP5D2KJU

- Neural Networks and Deep Learning
- Structuring Machine Learning Projects

Machine Learning Certificate.

https://www.coursera.org/account/accomplishments/specialization/A9UH3X9UVD9R

- Supervised Machine Learning
- Unsupervised Learning, Recommenders, Reinforcement Learning

Languages

- **Spanish** | Native
- **English** | Advanced (C1)

Technical skills

- Cloud/MLOps: GCP, Vertex AI, Kubeflow, Docker, CI/CD, Apache Airflow
- Data Science: Python, Scikit-learn, SVM, Random Forest, Deep learning Algorithms
- Data Engineering: SQL, BigQuery, Azure Data Factory, dbt, Dagster, Apache Spark
- Tools: Git, Power BI, Flask, Puppeteer, Selenium

Competences

- Outstanding analytical and logical skills.
- Way with words and assertive communication.
- Highly skilled in problem-solving and decision-making.
- Highly experienced as a team leader.
- Great creativity and imagination to solve problems at the right time.