PEDRO JULIUS LEMOS DOS SANTOS

AI Engineer

Machine Learning, Deep Learning, Data Engineering and Analytics

Fortaleza, Ceará, Brazil | +55 (85) 98553-6293 | pdrjuliusls@gmail.com

pedrojulius.tech | linkedin.com/in/pedro-julius | github.com/pjuliusl

Summary

Applied AI and data professional working as a Data Analyst for the Civil Police of Ceará, Brazil. I build production pipelines, analytical models and integrations between AI applications and data tools, using Python, SQL, Airflow and dbt, plus MCP, LoRA fine-tuning and model evaluation with MLflow. My work covers data quality and traceability, automated testing and human review of AI answers.

Technical skills

AI and MLOps: PyTorch, Transformers, PEFT/LoRA, MLflow, LiteLLM, MCP, fine-tuning, model evaluation and calibration.

Data engineering: Python, SQL, Apache Airflow, dbt, PostgreSQL, ClickHouse, ETL/ELT pipelines, data validation and quality.

Analytics and BI: Apache Superset, Power BI, data modeling, operational metrics and results analysis.

Development and delivery: FastAPI, Pydantic, SQLAlchemy, Alembic, pytest, Git, Docker, Jenkins and training Jobs on Kubernetes/K3s.

Secure integrations: REST APIs, OAuth2/OIDC, JWT, access control, CSRF protection and Redis/Sentinel sessions.

Experience

IT Department, Civil Police of Ceará (PCCE)

Data Analyst | contractor | Jul 2025 - present

- Design and maintain Airflow DAGs and dbt transformations on SQL, PostgreSQL and ClickHouse, investigating failures and improving production pipelines.
- Fixed an incremental-capture gap that left changes to old records out of the target database, and confirmed the result in the final database.
- Built extraction of metrics from semi-structured PDFs in Python/Airflow, with validation, traceability and a monthly comparison of two sources in Superset dashboards.
- Build MCP integrations for Superset and tooling for training, evaluating and human review of AI models.

iREDE (Institute for Management, Technology Networks and Energy)

Project Assistant | Oct 2024 - Jun 2025

- Tracked deliveries and wrote technical documentation for projects.

Android Development Teaching Assistant | Oct 2023 - Dec 2024

- Prepared hands-on classes and Kotlin learning tracks, and mentored beginners building Android apps.

Maria Bezerra Quevedo Municipal School

IT Support Intern | Aug 2022 - Feb 2023

- Supported the computer lab: software installation, equipment maintenance, Chromebook operation and reports.

Applied AI and Machine Learning projects

MCP integration with Apache Superset | PCCE

- Delivered the base of a Python MCP Host/Client that connects an analytics assistant to Superset, with authentication, sessions and tool catalog validation.
- Contributed to the Superset MCP runtime: JWT/JWKS validation, scopes and restriction to authorized tools, with tests for authentication and integration contracts.
- Built the in-dashboard chat bootstrap with CSRF protection and access checks. Isolated sessions and nonces per user, session and tab, using Redis with expiry and atomic operations.
- Added Redis Sentinel support and container secret loading. Set up Jenkins builds, smoke tests and image publishing to Harbor.

Fine-tuning and evaluation of decision models | PCCE

- Built a LoRA fine-tuning pipeline for structured four-class classification with PyTorch, Transformers and PEFT on a pretrained language model.
- Prepared training, validation, calibration and final test sets grouped by record, so no record leaks across splits and the evaluation stays independent.
- Set up GPU training Jobs on Kubernetes/K3s with pinned dependency images, checkpoints, resume and controlled interruption.
- Integrated MLflow to log parameters, metrics and version references, while keeping record content out of logs and experiment tracking.
- Implemented evaluation by macro-F1, balanced accuracy and per-class precision and recall, plus checkpoint selection rules that cap severe errors and support calibration.
- Built weight export with training/serving parity checks, release identification and candidate preparation for evaluation before production.

Human review platform for AI answers | PCCE

- Built supervised review apps with FastAPI, SQLAlchemy, PostgreSQL and Alembic: work queue, drafts, written reasons and decision history.
- Implemented blind review comparing model and analyst answers, with separate access roles and review flows protected by authentication and CSRF.
- Built dashboards for agreement, severe errors, coverage and confidence, plus operational metrics for volume, failures and latency.
- Added review import, results export and automated tests, and used simulated data to validate the flows before going live.

Chat Infinito, a fork of UniiChat | Personal project

- Adapted UniiChat, Victor Taelin's open source chat that keeps the whole conversation in a tree of summaries, to run on Claude and ChatGPT subscriptions through the official command-line clients instead of API keys.
- Implemented batched summaries through Codex, model fallback when a plan hits its limit, exact search over the history and a rule that keeps links and numbers in summaries, with end-to-end tests.
- Added live subscription usage, isolation between integrations and a systemd service for daily use. Stack: TypeScript, Bun, MCP.

Data and automation projects

Power BI analytics dashboard | Portfolio

- Built a Power BI dashboard on PostgreSQL with the Northwind sample database, analyzing sales, revenue by product, stock and average ticket.
- Organized the metrics for exploring results and comparing products in a sample business scenario.

Education

B.Eng. in Telecommunications Engineering

Federal Institute of Ceará (IFCE) | *Jul 2023 - expected Jul 2028*

Undergraduate degree in progress, focused on telecommunications systems.

Technical High School Diploma in Computing

Federal Institute of Ceará (IFCE) | *Jul 2018 - Jul 2022* | *Completed*

Technical computing program integrated with high school.

Additional training

Introduction to Android Development

Fundamentals of Android mobile app development.

Starting a Front-end Developer Career

Fundamentals of web interface development.

Languages

Portuguese: Native.

English: EF SET 82/100, C2 Proficient.