

João Pedro da Cunha

Machine Learning Engineer

Location: Cottbus, Germany | **Phone:** +55 61981341392 | **Email:** joaopedrocdias@hotmail.com

Linkedin: [linkedin.com/in/jpcunhadas](https://www.linkedin.com/in/jpcunhadas) | **Portfolio:** github.com/jpcunhadas

Professional Summary

Machine Learning Engineer with 3+ years of experience developing and deploying machine learning systems, from experimentation through production monitoring and maintenance. Experience spans classification, regression, anomaly detection, and NLP, with a focus on translating models into reliable production systems.

Experienced in building data pipelines, APIs, and containerized services, as well as working with cloud infrastructure, CI/CD, and MLOps practices. Focused on reproducible and maintainable machine learning systems, with a growing interest in LLM applications.

SKILLS

- Machine Learning: PyTorch, TensorFlow, scikit-learn, NLP, Deep Learning
- ML Engineering & MLOps: MLflow, DVC, Kedro, Airflow, FastAPI, Docker, CI/CD
- Data & Cloud: SQL, Spark, AWS, GCP
- Programming: Python

Professional Experience

Machine Learning Engineer

Pernambuco State Health Secretariat — Remote (Contractor)

Oct 2024 – Jan 2026

- Developed NLP pipelines using Kedro, DVC, and MLflow to automate the classification of medical reports, reducing manual data entry and ensuring full traceability for auditing purposes.
- Automated the entire OCR pipeline (OpenCV, Tesseract, PaddleOCR), reducing processing time from 20 days to 7 hours while integrating outputs into MongoDB and PostgreSQL.
- Designed a modular and reusable OCR architecture parameterized via CLI/YAML, enabling scalable processing of multiple document and image formats.
- Built APIs using FastAPI, containerized with Docker and integrated with MinIO as a data lake and artifact storage layer to serve models and data efficiently.
- Implemented CI/CD with Jenkins, automated testing with Pytest, and code quality checks with Ruff and Black, standardizing validation and deployment workflows.
- Standardized the model lifecycle with MLflow (experiment tracking and remote model registry), reducing experiment rework and accelerating model performance comparison.
- Conducted exploratory data analysis and A/B testing (Mann-Whitney, t-test, Chi-square) to validate models and support statistically grounded decisions.

Data Scientist

Fácil Espaider — Blumenau, Santa Catarina, Brazil (Remote)

May 2023 – Sep 2024

- Developed machine learning models using scikit-learn, Pandas, and NumPy, contributing to an 18% increase in customer satisfaction in service operations.
- Implemented NLP pipelines (NLTK, spaCy, fastText, Word2Vec) for text classification and analysis, automating insight extraction and reducing manual analysis effort.
- Integrated machine learning models into legacy C# systems, expanding functionality and increasing the value of existing products.
- Conducted A/B testing to compare the performance of existing models with new solutions, supporting data-driven decisions based on robust metrics.
- Collaborated directly with clients to gather requirements, propose improvements to existing products, and identify new opportunities for applying machine learning.

Researcher & Data Scientist (Research Contract)

Ford Motor Company — Bahia, Brazil (Remote)

May 2023 – Apr 2024

- Analyzed large-scale datasets using PySpark and MLlib, applying inferential statistics (distributions, hypothesis testing, multivariate analysis) to generate insights on electric vehicle (EV) performance.
- Worked in a Google Cloud Platform environment (BigQuery, Vertex AI), handling large-scale data and provisioning cloud resources to accelerate AI experimentation.
- Designed and conducted research involving regression, classification, and clustering using EV telemetry data.
- Collaborated with business, engineering, and sales teams to develop features and align technical solutions with strategic objectives.
- Communicated complex analytical results clearly to non-technical stakeholders, ensuring understanding and practical application of insights.
- Interacted weekly with global leadership, presenting project progress in English and coordinating efforts with international teams.

LANGUAGES

- **English:** C1 — IELTS 7.5.
- **Spanish:** B1
- **Portuguese** — Native

EDUCATION

Postgraduate Program in Machine Learning Engineering — FIAP, 2025

B.Sc. in Computer Engineering — UniCEUB, 2021