Müsel Emmanuel Tabares Pardo

✉ muselemmanuel@hotmail.com　📞 (+52) 663 104 6329　⬡ musel25

🌐 musel.dev　📍 Compiègne, France　MEXICAN | DOB: 05/08/2002

About

Seeking an opportunity for a Master's student internship at Mistral. Strong desire to show what I am capable of. Passionate about learning and solving problems. Obsessed with LLMs and mathematics.

Experience

9/2023 – 01/2024	**Machine Learning Engineer**	Moodme	United States

Designed and deployed real time facial emotion recognition systems using deep learning computer vision techniques. Implemented a Graph Auto-Encoder (GAE) to learn latent representations of facial landmarks for a Unity 3D Face SDK, enabling identity preserving expression synthesis and emotion classification.

02/2023 – 08/2023	**Data Analyst**	CEMEX	Monterrey, Mexico

Developed predictive and statistical forecasting models to assess user behavior trends and adoption metrics for CEMEXGO. Use of Snowflake and SQL for data retrieval; Built interactive Power BI dashboards to support Product Management decisions.

Education

09/2024 – Current	**Erasmus Mundus Master in Advanced Machine Learning and Optimisation Systems (EMSSE)**

Universitat Politècnica de Catalunya | Barcelona, Spain (09/2024–01/2025)

Universiteti Politeknik i Tiranës | Tirana, Albania (03/2025–06/2025)

Université de technologie de Compiègne | Compiègne, France (05/2025–Current)

08/2023 – 12/2024	**Specialization in Advanced Artificial Intelligence Applied to Data Science**

Monterrey Institute of Technology and Higher Education | Guadalajara, Mexico

08/2021 – 06/2024	**B.S. in Computer Science and Technology**

Monterrey Institute of Technology and Higher Education | Monterrey, Mexico

Projects

Instruction Tuning (SFT + QLoRA) — End-to-end instruction-tuning workflow with prompt templating, tokenization, and QLoRA adapters; reproducible configs, logged runs, checkpoints, and artifacts; containerized runner. ⬡

Preference Optimization (DPO/ORPO) — Post-training with pairwise preference data; trained/evaluated DPO/ORPO adapters, added guardrails/refusals, and compared pre/post behavior on task-specific evals. ⬡

LLM Evaluation Harness & Leaderboard — Reproducible offline evals across reasoning/knowledge/safety with deterministic prompts/seeds; compact leaderboard and model cards; contamination checks and error analysis. ⬡

High Performance Serving(vLLM+ TensorRT) — Production-style serving optimized for throughput/latency with quantization and observability; streaming tokens, autoscaling, tracing, and A/B backend switching behind a clean API. ⬡

Multi-Node Fine-Tuning on Slurm — Distributed fine-tuning on HPC with checkpointing, preemption safety, NCCL tuning, shard-aware dataloading, and artifact sync; "zero to cluster" guide. ⬡

Telemetry RAG YANG/Openconfig	🔖	End-to-end RAG indexing lab telemetry docs (Markdown/YANG/JSON) into Qdrant; context-only prompting via Ollama/OpenAI; reproducible Docker setup and CLI for index/ask. ⍟
Medical LV Segmentation	🔖	U-Net for left-ventricle segmentation on Stanford EchoNet-Dynamic; compared pixel-wise masks vs. landmark detection; used Dice score and morphology post-processing. ⍟
ML for Activity Recognition	🔖	Evaluated 18 ML algorithms with grid search on accelerometer data to classify activity levels in older adults; validated via CV and significance tests. ⍟
Vision Hydrological Forecasting	🔖	Deep CNN pipeline predicting river stage from imagery; EfficientNet transfer learning and multi-modal inputs (visual + hydrological) for accuracy gains. ⍟
LLM + RAG for Metaheuristics	🔖	Automated framework using RAG + LLMs to generate and refine metaheuristics for black-box optimisation; integrated Qwen2.5-coder, ChromaDB, and Optuna. ⍟
ML Obesity Estimation	🔖	End-to-end pipeline classification on UCI obesity dataset; DR (PCA, t-SNE), mutual information, 31 models; best: XGBoost (96%). ⍟
Leukocyte CNN + LIME	🔖	CNN to classify leukocytes (cancerous vs non-cancerous) with LIME explanations to highlight decision-driving regions.
Fine-tuned GPT-2 Chatbot	🔖	Fine-tuned GPT-2 on WhatsApp chats; regex parsing to HF Datasets; trained and deployed interactive chatbot pipeline. ⍟
NLP Web App	🔖	Real-time text emotion classifier with logging and analytics; built using Streamlit and standard NLP preprocessing. ⍟
Advanced Statistical Learning	🔖	From-scratch Monte Carlo, Bootstrap, PCA, multiple regression (Ridge, Lasso), and classic classifiers; annotated theory notes applied to real datasets. ⍟
NLP Project Collection	🔖	Curated models: DistilBERT sentiment (PyTorch), Levenshtein spell checker, and Keras RNN text generator with temperature control. ⍟

Skills

LLM Training & Alignment	🔖	PyTorch; Hugging Face Transformers/PEFT/TRL; tokenization; SFT, LoRA/QLoRA, DPO; multilingual (EN/FR).
Serving & MLOps	🔖	vLLM, TGI; FastAPI; llama.cpp/GGUF; Git, Linux, Docker, Kubernetes; experiment tracking (W&B/MLflow); CI/CD; PyTorch Profiler.
Data, RAG & Evaluation	🔖	Dataset curation (cleaning/dedup/PII); vector search (FAISS/Qdrant) & rerankers; prompt/retrieval pipelines; lm-evaluation-harness; A/B evals; safety/robustness.

Languages

Mother Tongue	🔖	Spanish
English	🔖	Listening C2 | Reading C2 | Writing C1 | Spoken Production C1 | Spoken Interaction C1
French	🔖	Listening B2 | Reading B2 | Writing B1 | Spoken Production B1 | Spoken Interaction B1
German	🔖	Listening A2 | Reading A2 | Writing A1 | Spoken Production A1 | Spoken Interaction A1
Albanian	🔖	Listening A2 | Reading A2 | Writing A1 | Spoken Production A1 | Spoken Interaction A1

Volunteering

08/2018 – 07/2020   **Leader, ProV**   Mexico
Coordinated a social aid organisation to support vulnerable communities.