



PEDRO SATORRE-MULET

Machine Learning & AI Engineer · Specialty: Football/Sports AI & Analytics

Palma de Mallorca / Madrid, Spain · +34 630 72 61 84 · pesatomu.r.m@gmail.com · [LinkedIn](#)

Available immediately · Freelance / contract · Full-time / Part-time
Remote or hybrid within/outside the EU · Open to travel and/or relocate within/outside the EU

PROFILE

ML & AI Engineer specialized in football (and other sports) AI & Analytics, with 6 years of hands-on Python and ML practice, and an MSc in Data Science from Universität Potsdam (Germany), in collaboration with Hasso Plattner Institut (HPI).

At **EY** I single-handedly **founded, designed, and successfully built a Sports AI & Analytics business line** from scratch (with a special focus on football), **under the direct supervision and sponsorship** of a department-linked share-holding **Partner**, in parallel to my formal duties. He supported carrying out this initiative after I proactively proposed it, shortly after joining EY, pitching the potential of this idea and business opportunity to the entire department. I led a small team of 3 technical people, shipping end-to-end ML, Deep Learning, and Generative AI/Agentic AI solutions mainly on: Databricks, LangChain/LangGraph/LangSmith, and GCP. Collaborated cross-functionally with technical and non-technical client stakeholders.

During my [MSc. Data Science thesis](#) I developed a cutting-edge, state-of-the-art **research project** under the direct supervision from [Dr. Gabriel Anzer](#) (Head of Football Data Analytics at **RB Leipzig**) - successfully demonstrated statistically that Heterogeneous Graph Transformers (HGTs) classify, estimate, and calibrate better football shots & their respective goal probability (i.e. xG = Expected Goals), compared to conventional xG models; this work was selected by [David Sumpter](#) as **1 of 4 finalists at Twelve Football's "Pitch to the Pros #3"**.

Available for freelance and contract engagements as an **AI Engineer, Agentic AI Engineer, GenAI Engineer, ML Engineer, DL Engineer** or **AI-focused Data Scientist**.

KEY ENTITIES / ENTERPRISES

EY (Ernst & Young)
Global Professional Services

Twelve Football
Football Data & AI Company

RB Leipzig
German 1. Bundesliga Football Club

Universität Potsdam
w/ Hasso Plattner Institut (HPI)
(Top-tier #1 German Institution for CS & IT)

Lancaster University
UK Top-10 University
UK Top-7 in Physics & Astrophysics

CORE EXPERTISE

- Machine Learning & Deep Learning** : Graph Neural Networks (GNNs, Heterogeneous Graph Transformers – HGTs) · supervised, semi-supervised & unsupervised learning · feature engineering · model calibration & evaluation · AutoML · experiment tracking · MLOps
- Generative & Agentic AI** : LLM agents & tool use · context engineering · prompt engineering · RAG · agent tracing & evaluation · agentic conversational analytics applications · LangChain · LangGraph · LangSmith
- Data/AI Cloud Platforms** : Databricks (Mosaic AI, Agent Bricks, Genie Agents, MLflow 3.x, PySpark & distributed data processing and analysis) · Google Cloud Platform – GCP (Gemini Enterprise Agent Platform/Vertex AI, BigQuery, Dataform) · Azure ML · Power BI · Streamlit
- Football/Sports AI & Analytics** : Event & tracking data · (Spearman) Pitch Control · Expected Possession Value (EPV) modelling · EPV-Added modelling · Expected Goals (xG) modelling · shot and set-piece analyses · player scouting & pairing analyses · match performance analytics · agentic conversational analytics application (CBs & CB-Pairings AI Analyst agent + Streamlit)

TECHNOLOGY STACK (years of experience)

Prog. Languages	Python (6) · PySpark (2) · SQL (2) · R / R-Studio (1)
DA, BI, ML, DL & AI	NumPy (6) · pandas (6) · Matplotlib (6) · Seaborn (6) · scikit-learn (6) · Keras (6) · PyTorch (4) · PyTorch Geometric – PyG (1) · Spark MLlib (2) · AutoML (2) · MLflow (2) · LangChain / LangGraph / LangSmith (1.5)
Platforms & Cloud	Databricks – Mosaic AI (2) · Databricks – Agent Bricks (1) · Databricks – Genie Agents (1) · Databricks Apps (1) · GCP – Agent Platform/Vertex AI (1.5), GCP – BigQuery + Dataform (0.5) · GCP – Google Colab (4) · Azure ML Studio (0.5)
Tooling & Delivery	Jupyter Lab / Notebook (6) · Visual Studio Code (4) · GitHub (4) · Git (4) · Power BI (1.5) · Streamlit (1)
Project Management	Jira (0.75)



Agentic AI Engineering – Football/Soccer CBs & CB-Pairings AI Analyst Agent

Twelve Football · Remote

Feb 2026 – May 2026 · 3 mos | 1st Edition of “AI Context Engineering for Football” | Agentic AI Engineering

- **Selected** onto [Twelve Football's 1st edition of their “AI Context Engineering for Football” course](#) - a very limited cohort of **45 participants worldwide**, hand-picked by [David Sumpter](#) himself plus his committee- and delivered the accompanying Agentic AI Engineering project as sole developer.
- Designed, developed, and delivered a [conversational AI Analyst agent for Centre-Backs \(CBs\) & CB-pairings evaluation & scouting](#), grounding an LLM in structured football data through AI context engineering and tool-use techniques.
- Released as a fully equipped & governed Streamlit front-end chat application, answering natural-language scouting questions with analyst-grade player analyses, comparisons, pairing recommendations, and more...

TECH STACK Python · LLM tool use & context engineering · RAG · LangChain / LangGraph / LangSmith · scikit-learn · NumPy · pandas · Matplotlib · Seaborn · football event data · Streamlit · UI/UX · GitHub · Git · VS Code



Technology Consultant – AI, Data Science & Analytics

EY (Ernst & Young) · Madrid, Spain

Oct 2024 – Dec 2025 · 1 yr 2 mos | Full-time | ML/AI Engineer

- Designed and built a [new Sports AI & Analytics business line](#) from scratch – single-handedly and as a parallel initiative alongside daily delivery – after proactively pitching the idea & business opportunity to the entire department, securing sponsorship and direct supervision from a department-linked share-holding Partner.
- Delivered end-to-end Machine Learning (ML) and Deep Learning (DL) solutions for clients: data preparation and feature engineering, model training and evaluation, experiment tracking, MLOps, deployment and monitoring.
- Developed Generative (GenAI) & Agentic AI solutions for clients: LLM-based knowledge assistants, retrieval-augmented (RAGs), and tool-using agents.
- Led a team of three technical specialists: task planning, technical guidance, code review, and delivery quality.
- Built decision-ready dashboards and visual analytics: translating model output into recommendations for business stakeholders.
- Ran cross-functional collaboration with technical and non-technical client teams, from requirements workshops through to hand-over.

TECH STACK Python · PySpark · SQL · Databricks (Mosaic AI, Spark MLlib, MLflow, AutoML, Agent Bricks, Genie Agents, Databricks Apps) · GCP (Gemini Enterprise Agent Platform/Vertex AI, BigQuery, Dataform) · Azure ML Studio · LangChain / LangGraph / LangSmith · scikit-learn · PyTorch · Keras · NumPy · pandas · Matplotlib · Seaborn · Power BI · Streamlit · UI/UX · GitHub · Git · VS Code · Jira



Machine Learning & Deep Learning Researcher – Football AI & Data Analytics

RB Leipzig · German 1. Bundesliga Football Club · Potsdam / Leipzig, Germany

Jun 2023 – Jan 2024 · 7 mos | Supervised by [Dr. Gabriel Anzer](#) – Head of Football Data Analytics, RB Leipzig

- Built a Graph Neural Network pipeline over state-of-the-art event & tracking data, representing every shot as a heterogeneous graph of all players, the ball, as well as the instantaneous spatial & physical context.
- Trained and benchmarked [Heterogeneous Graph Transformers \(HGTs\) for shot classification and goal-probability \(xG\) estimation](#), successfully proving statistically significant gains in shot outcome classification, and their respective goal-probability estimation & calibration over traditional xG models.
- Written [recommendation letter from Dr. Gabriel Anzer](#) is available.
- Selected by [David Sumpter](#) (and his committee) as [1 of 4 finalists for Twelve Football's “Pitch to the Pros #3”](#).

TECH STACK Python · PyTorch · PyTorch Geometric (PyG) · Heterogeneous Graph Transformers (HGTs) · scikit-learn · NumPy · pandas · Matplotlib · Seaborn · football event & tracking data · Git · VS Code

(Databricks-like) Football Data, Analytics & AI Platform

Independent football AI architecture & platform project | Jun 2026 – Present | Developer & ML/AI Engineer

Designing and engineering a [Databricks-inspired, full-stack Football Data, Analytics & AI \(end-to-end\) Platform](#) combining provider-adapter ingestion, RBAC & audited sharing, DuckDB governed data/model catalogues across 40 sections with capability- & data-gated features. Supports Streamlit/FastAPI interfaces, optional MCP and a Next.js/WebSocket collaboration client. Implemented capabilities: football modelling, ML (+ MLOps) & RL experimentation, capability-gated DL & computer-vision (CV) prototypes, event-derived pitch-value & graph-based analytics, RAGs/GraphRAGs, AI agents, EvalOps, observability & monitorization, and more... In progress: tactical simulation, multimodal fusion, interactive Ops Canvas, a governed ball-impersonated copilot assistant christened "Tiki" (from the 'tiki-taka' term), and more on their way...

TECH STACK Python · TypeScript · FastAPI · Streamlit · Next.js · React · Tailwind CSS · UI/UX · DuckDB · pandas · NumPy · scikit-learn · LightGBM · PyTorch · PyTorch Geometric (PyG) · LangGraph · FAISS · Plotly · OpenCV · Ultralytics YOLO · Stable-Baselines3 · Gymnasium · SHAP · statsmodels · PuLP · Kùzu · OpenAI API · Gemini API · MCP (optional) · REST/OpenAPI · WebSockets · Pytest · Git · GitHub · VS Code

CBs & CB-Pairings AI Analyst (Agent)

Twelve Football | Feb 2026 – May 2026 (ongoing into the present) · 3 mos | Agentic AI Engineer

I got [selected](#) onto [Twelve Football's 1st edition of the course \(AI\) "Context Engineering for Football"](#) (a very limited cohort of 45 people worldwide was hand-picked by [David Sumpter](#) himself + his committee) - implemented a hands-on project as part of the course, to develop an AI Analyst agent; I implemented a conversational [CBs & CB-Pairings AI Analyst agent](#): grounding an LLM in structured football event data through context engineering, RAG, and tool-use techniques. It is then [delivered as a Streamlit chat front-end app](#), answering natural-language football scouting and analyst questions with professional-analyst-grade player comparisons & pairing recommendations. AI Analyst Agents for other positions on the pitch are on their way...

TECH STACK Python · LLM tool use & context engineering · RAG · LangChain / LangGraph / LangSmith · scikit-learn · NumPy · pandas · Matplotlib · Seaborn · football event data · Streamlit · UI/UX · GitHub · Git · VS Code

New Sports AI & Analytics Business Line

EY (Ernst & Young) | 2025 · 9 mos | ML/AI Engineer

As a side-project, independent from my daily duties, I single-handedly designed and built successfully from scratch a new business line for my department on [Sports AI & Analytics](#), with special focus on football/soccer, under the sponsorship and direct supervision of a department-linked share-holding Partner. He supported carrying out this initiative after I proactively proposed it, shortly after joining EY, pitching the potential of this idea and business opportunity to the entire department.

Shot Classification & Goal-Probability (xG) Estimation Using Graph Neural Networks (HGTs)

MSc Data Science thesis | Jun 2023 – Jan 2024 · 7 mos | Researcher & ML/DL Engineer

State-of-the-art and cutting-edge [MSc. Data Science thesis' research project](#) – modelling football shots as heterogeneous graphs of all players, the ball, as well as the instantaneous spatial & physical context, and then applying Heterogeneous Graph Transformers (HGTs) to classify shot outcomes and estimate their respective goal probability (i.e. xG = Expected Goals); Successfully demonstrated statistically that HGTs significantly improve the shot outcome classification, and their respective goal-probability estimation & calibration over traditional xG models. This thesis was directly supervised under [Dr. Gabriel Anzer](#) (Head of Football Data Analytics at RB Leipzig).

In addition, this project was selected by [David Sumpter](#) as [1 of 4 finalists for Twelve Football's "Pitch to the Pros #3"](#).

TECH STACK Python · PyTorch · PyTorch Geometric (PyG) · Heterogeneous Graph Transformers (HGTs) · scikit-learn · NumPy · pandas · Matplotlib · Seaborn · football event & tracking data · Git · VS Code

Expected Possession Value (EPV)-Added surfaces & contours – static & animated

Independent football Data Science & modelling project (parallel to MSc Data Science thesis) | 2023 · 1 mo | Developer & Analyst

[Implemented EPV-Added surface + contour overlays](#), both in static (.png) and in animated (.mp4) video formats, from football spatio-temporal event & tracking data to visualize which areas of the pitch should the ball be displaced to, at every instant in time by the team in possession, to maximize their xG (Expected Goals) within that same passage play.

TECH STACK Python · NumPy · pandas · Matplotlib (animation) · football event & tracking data · Git · VS Code

Spearman Pitch Control surfaces – static & animated

Independent football Data Science & modelling project (parallel to MSc Data Science thesis) | 2023 · 1 mo | Developer & Analyst

[Implemented Spearman's Pitch Control surface overlays](#), both in static (.png) and in animated (.mp4) video formats, from football spatio-temporal event & tracking data to visualize the areas of the pitch that are under control by each of the teams and which areas would be under dispute for the ball.

TECH STACK Python · NumPy · pandas · Matplotlib (animation) · football event & tracking data · Git · VS Code

EDUCATION



MSc. Data Science

Universität Potsdam (w/ Hasso Plattner Institut – HPI) · Potsdam, Germany

Oct 2020 – Mar 2024 | Grade/Score: (German) 1.9 ≈ (U.S. GPA) 3.2 / 4.0



- Delivered in cooperation with the **Hasso Plattner Institute (HPI)** – Top-tier #1 German Institution for CS & IT.
- **Main Courses** : Advanced Machine Learning · Deep Learning · Data Science · Applied ML in Digital Health · Statistical Data Analysis · Business Data Analytics & BI · Bayesian Inference & Data Assimilation · Data- & Knowledge-Base Systems · Biostatistics & Epidemiological Data Analysis
- **Thesis** : “Shot Classification & Goal Probability Estimation Using Graph Neural Networks” (specifically, Heterogeneous Graph Transformers – HGTs) – directly supervised by [Dr. Gabriel Anzer](#) : Head of Football Data Analytics at RB Leipzig.
- Grade/Score : (German) 1.2 ≈ (U.S. GPA) 3.8 / 4.0



BSc. (Hons) Physics, Astrophysics & Cosmology

Lancaster University · Lancaster, England, United Kingdom (UK)

Oct 2017 – Jun 2020 | UK Top-10 University | UK Top-7 in Physics, Astrophysics & Cosmology

Grade/Score: (English/UK) 2:1 Upper-Second Class ≈ (U.S. GPA) 3.4 / 4.0

- Quantitative knowledge in mathematical and physical modelling, statistics, and scientific computing – the building blocks for later work in ML, DL, and spatio-temporal football/sports models & visualizations:
 - **Mathematics** : Advanced Calculus · Adv. Linear Algebra · Adv. Statistics · Numerical & Complex Methods
 - **Physics** : Advanced Quantum Mechanics · Advanced Astrophysics & Cosmology · Special & General Relativity · Particle Physics · Nuclear & Atomic Physics · Solid State Physics · Thermodynamics · Electromagnetism

CERTIFICATIONS & PROFESSIONAL TRAINING

- | | |
|---|----------|
| ▪ Agent Evaluation on Databricks (and MLflow 3.x) – Accreditation Badge · Databricks | Aug 2026 |
| ▪ Building RAG Agents (Knowledge Assistants) with Agent Bricks – Accreditation Badge · Databricks | Aug 2026 |
| ▪ Introduction to LangSmith – Certificate · LangChain | Jul 2026 |
| ▪ AI Context Engineering for Football – Certificate · Twelve Football | May 2026 |
| ▪ Get Started with AI Agents on Databricks – Accreditation Badge · Databricks | Feb 2026 |
| ▪ Introduction to LangGraph – Certificate · LangChain | Oct 2025 |
| ▪ AI Agent Fundamentals – Accreditation Badge · Databricks | Oct 2025 |
| ▪ Generative AI Engineering – Certificates · Databricks | Oct 2025 |
| ▪ Generative AI Fundamentals – Accreditation Badge · Databricks | Sep 2025 |
| ▪ Machine Learning Engineering – Certificates · Databricks | Jul 2025 |
| ▪ GCP Machine Learning: Vertex AI Gemini Generative AI – Certificate · Udemy | Jul 2025 |

PUBLICATION

- **Facial Landmarks Detection: A Brief Chronological Survey & Practical Implementation.** Bätz, A., De Miguel Palacio, A., Mishra, G., Rodriguez Llorca, I., Raihan Pranti, M., Satorre Mulet, P. & Lippert, C. (2021). [DOI: 10.13140/RG.2.2.36199.98722](#)

LANGUAGES (proficiency)

Spanish (Native) · Catalan (Native) · English (Bilingual) · Greek (Intermediate) · German (Basic) · Italian (Elementary) · Chinese (Elementary)

AVAILABILITY, ENGAGEMENT & REFERENCES

Availability	Immediately, full capacity – freelance / contract, remote or hybrid engagements within/outside the EU, open to occasional travel, also open to relocation; based in Palma de Mallorca / Madrid, Spain
Target roles	AI Engineer / Agentic AI Engineer · GenAI Engineer · ML Engineer · DL Engineer · AI-focused Data Scientist · Football/Sports AI-focused roles
Typical engagements	AI/Agentic AI/GenAI/ML solutions & platform work on Databricks, GCP, LangChain/LangGraph/LangSmith/Deep Agents, AWS, Azure · xG, (Spearman) Pitch-Control, and EPV(-Added) modelling on event & tracking (sports) data
References	Recommendation letter from Dr. Gabriel Anzer – Head of Football Data Analytics, RB Leipzig; its respective formally signed version may be provided upon request . Further EY references upon request.