

Miguel Vidal

San Antonio, TX

Phone: 863-944-5662 | E-Mail: migvidal5@gmail.com | GitHub: Migvidal5

Summary

Applied data scientist who transforms complex data into clear, practical insights. Experienced in Python and R for developing analytical models, pipelines, and visualizations that reveal trends and drive decisions. Strong background in Bayesian methods, clustering, and machine learning, with a focus on communicating technical concepts in a way that brings teams together and supports genuinely informed, shared decision-making.

Experience

Senior Data Scientist, US Soccer Federation

October 2024 – Current

- Currently building out a distributed Expected Possession Value (EPV) model with SkillCorner data using GNNs and CNNs within Databricks.
- Played a key role in the development of federation-wide post-match and opposition reports built with Matplotlib, integrating tactical, physical, and tracking data.
- Applied Bayesian modeling techniques to generate probability mass functions (PMFs) of expected goals (xG) and post-shot xG, quantifying match outcome probabilities (win/draw/loss) and uncertainty.
- Engineered asynchronous data pipelines connecting SkillCorner and StatsBomb APIs, automating ingestion and synchronization of tracking and event data.
- Conducted assist zone research utilizing heat maps and k-means clustering to identify spatial patterns in chance creation and refine attacking principles.
- Implemented CI/CD workflows (pytest, GitHub Actions) for testing and deployment of analytics pipelines, ensuring reliability and reproducibility.
- Served as a technical liaison to performance, coaching, and analytics staff, delivering actionable insights from physical, tactical, and wellness datasets.
- Collaborated with engineering teams to maintain version control and monitor system performance, ensuring consistency across analytics tools.

WNT Sport Scientist, US Soccer Federation

June 2022 – October 2024

- Revamped sport science processes to enhance training planning and roster selection.
- Developed Shiny apps to analyze and visualize physical and wellness data, implementing algorithms like acute:chronic workload and cluster analysis.
- Designed a "Morning Screen" process to capture, analyze, and mitigate athletes' responses efficiently.
- Contributed to the team's success in winning the **2024 Olympic Gold Medal**, **2023/2024 SheBelieves Cup**, and **2023 CONCACAF W Championship** by optimizing physical performance strategies and data-driven decision-making.

Lead Data Analyst, Guardian Premier Solutions

March 2022 – June 2022

- Pulled reports from databases and created R scripts to streamline data wrangling processes.
- Developed a Shiny app summarizing physiological data, improving internal reporting capabilities.
- Mentored team members by teaching data science concepts, including data cleaning, visualization, and introductory machine learning.

Software Developer/Data Analyst, Guardian Premier Solutions

August 2021 – March 2022

- Created a Human Performance dashboard for company-wide use, flagging individuals based on physiological data.
- Conducted analysis of new technologies using Python and its associated libraries.
- Built predictive models using random forest and Bayesian probability to forecast candidate success rates.

Human Performance Technician, Guardian Premier Solutions

June 2020 – August 2021

- Monitored physiological responses to training, including live monitoring and sleep tracking for Air Force trainees.
- Generated reports in Smartabase highlighting critical red flags for stakeholders.
- Organized educational discussions and trained new technicians to ensure team success.

Education

Concordia University Chicago, River Forest, IL

Master of Science, Applied Exercise Science (Sport Performance and Training) - December 2018

Huntingdon College, Montgomery, AL

Bachelor of Arts, Exercise Science - May 2016

Publications

(Co-Author) Oliva-Lozano, J., et al. Tips for Transitioning to and From International Soccer Camps.

(Co-Author) Oliva-Lozano, J., et al. Predicting the Match Outcome in the 2023 FIFA Women's World Cup and Analysis of Influential Features.

Certifications

- Deep Learning and AI in Sport, Obtained November 2025
- Essential Causal Inference Techniques for Data Science, obtained September 2024
- Fusion Sport Smartabase Certificate, obtained September 2021
- IBM Data Science Certificate, obtained Jan 2021
- US Soccer Federation D LICENSE, obtained Feb 2019
- NASM Performance Enhancing Specialist, obtained Sept 2018