

# Diego Garcia

816-377-1273 | [diegofgarcia2003@gmail.com](mailto:diegofgarcia2003@gmail.com) | [linkedin.com/in/diego-f-garciaa](https://linkedin.com/in/diego-f-garciaa) | [diegofgarcia.com](https://diegofgarcia.com)

## EDUCATION

### University of Missouri - Columbia

Columbia, MO

*Bachelor of Science in Computer Science, Minor in Mathematics - GPA 3.5*

*August 2022 – May 2026*

- **Relevant Coursework:** Cloud Computing, Big Data Analytics, Database Applications & Information Systems, Artificial Intelligence

## EXPERIENCE

### Undergraduate Research Assistant

May 2025 – August 2025

*University of Missouri - Columbia*

*Columbia, MO*

- Built and containerized 10+ DERMS microservices using Docker and Kubernetes, supporting scalable ingestion of real-time DER telemetry.
- Developed real-time streaming pipelines with Apache Kafka and ksqlDB, processing high-frequency DER data with low-latency transformations.
- Automated CI/CD workflows for ML and federated learning models, reducing deployment time and improving reproducibility across Kubernetes clusters.

## PROJECTS

### Stock Market Dashboard with ML Predictions | *FastAPI, PostgreSQL, SQLAlchemy, scikit-learn, Chart.js*

- Built FastAPI REST API with async endpoints, PostgreSQL database integration, and real-time stock data caching, improving request efficiency via a 10-minute cache window.
- Engineered ML prediction model using Random Forest with rolling average features (2, 5, 60, 250, 1000-day windows) to estimate directional stock movement with confidence scoring.
- Designed SQLAlchemy ORM schema with 3 normalized tables for stock history, predictions, and model versioning, supporting 1000+ days of historical data backfill via yfinance API.
- Developed responsive JavaScript frontend with Chart.js visualizations, auto-refresh functionality, and dynamic ML prediction display with confidence metrics.

### MLB Game Outcome Predictor | *Python, Scikit-Learn, Pandas, Matplotlib*

- Engineered end-to-end ML pipeline processing 1,560 games with feature engineering, demonstrating predictive analytics applicable to sports betting and business forecasting.
- Optimized Random Forest classifier via GridSearchCV across 180 configurations, achieving 58.01% accuracy and 0.56 ROC-AUC, outperforming Logistic Regression by 2.56%..
- Identified home field advantage and team win percentage as top predictors through feature importance analysis and correlation heatmaps during comprehensive EDA.
- Delivered production-ready code with version control and reproducible Jupyter notebooks, showcasing statistical modeling and data visualization best practices.

### QuizMe Web Application | *MongoDB, Express.js, React, Node.js*

- Developed full-stack MERN application with RESTful API architecture, implementing JWT authentication, quiz CRUD operations, and real-time scoring logic.
- Designed MongoDB schema with Mongoose ODM for flexible quiz storage, supporting nested question arrays, answer validation, and user performance tracking.
- Built responsive React frontend with hooks (useState, useEffect) and React Router, integrating Recharts library for data visualization of quiz analytics.
- Deployed application with Node.js backend handling 100+ concurrent users, implementing error handling middleware and input validation with Express.js.

## TECHNICAL SKILLS

**Languages:** Python, SQL, Java, C/C++, JavaScript, C#

**Data & ML:** Pandas, NumPy, Scikit-Learn, Matplotlib

**Databases:** PostgreSQL, MySQL, MongoDB

**Big Data & Analytics:** Apache Spark, Apache Hadoop, Apache Kafka, ksqlDB, Tableau, Excel

**Backend & Cloud:** .NET, Docker, Kubernetes, AWS, GCP, Node.js, Express

**Frontend:** React, HTML/CSS, Recharts