

JORGE CAMPOS

Data Scientist

Email:
hirejorgecampos@gmail.com

Phone: +1 619-274-6045

LinkedIn:
<https://www.linkedin.com/in/jorgerochacampos/>

Location: San Diego, CA

SUMMARY

Former mechanical engineer turned data scientist with 3+ years leading eng teams from conceptualization to high-impact products. Experienced in building robust data pipelines from pre-processing to applied machine learning models.

EVERYDAY TOOLS

Python, Docker, Git, Pandas, Sklearn, PyTorch, Keras, SQL, TensorFlow, NumPy, MatLab, Microsoft Suite, Generative AI, Data Analysis, PowerBI, Solidworks, Autocad, Excel, Relational Databases, Llama index, Langchain, Vertex AI, ChromaDB, RAG, Google Cloud Platform, Vertex AI, Big Query,

RELEVANT EXPERIENCE

KaggleX Fellow
Google (Kaggle)
San Diego, CA
Aug 2024 - Present

- Contributing to the open-source Learning Path Index project, which creates customized learning paths for users by leveraging **Gemma2**.
- Enhancing an existing retrieval-augmented generation (**RAG**) pipeline by implementing **hybrid search techniques** that combine keyword and vector search, along with **few-shot prompting** to optimize Gemma2's outputs.
- Developing a synthetic Q&A dataset of 1,000 entries using Gemini 1.5 Pro on **Vertex AI** and evaluating it with local large language models (LLMs).
- Building a RAG pipeline with the **LlamaIndex** framework and **Weaviate** vector database, then benchmarking its performance against a fine-tuned Gemma2 model on synthetic Q&A data.
- Integrating **Langfuse** for real-time performance monitoring and evaluation, followed by **Dockerizing** the entire pipeline for deployment.
- Participating in technical and professional workshops led by AI experts to deepen knowledge and refine skills.

Data Scientist Intern
SteadyScript
San Diego, CA
Jan 2024 - Sept 2024

- Collaborate with the AI team to research and develop innovative AI features for a healthcare-focused CRM software.
- Implement a **CNN model** to extract patient information from forms, improving submission efficiency by **50%**, streamlining the data entry processes for healthcare providers, reducing manual effort and errors.
- Presented the final project to the CTO and stakeholders, effectively demonstrating the model's operational impact.

Data Scientist
Loan Project
San Diego, CA
Sept 2023 - Jan 2024

- Designed and deployed a machine learning model to automate loan approval processes, effectively addressing class imbalance challenges
- Apply **SMOTE** oversampling techniques to remove class bias and improve model accuracy by **~18%** in approving loans
- Experiment and test various supervised ML techniques using Sklearn, random forests, and **XGBoost**
- Improve model accuracy by using **GridSearchCV** to find the best hyperparameters for each model
- Benchmarked the performance of each machine learning model using cross-validation techniques.
- Optimized the final model with XGBoost, achieving **90%** accuracy in loan approvals.

Lead Engineer
Global Engineering Brigades (GEB)
Tegucigalpa, Honduras
July 2021 - Aug 2021

- Led a team of multi-disciplinary engineers mapping, surveying and designing water system for local community
- Create a budget and materials list while taking into account community feedback
- Develop an **data visualization** of community water pressure to discover lack of water pressure in households and business using **Pandas** and **Seaborn**
- Conducted data cleaning checking for missing values, data distributions, removing duplicates and outliers using **Excel**
- Present final design, visualizations and budget to community members, while taking feedback at the end

Teaching Assistant
UC Irvine
Irvine, CA
Sept 2020– Dec 2020

- Facilitated student mastery of coding best practices and data structure fundamentals in **MATLAB**
- Facilitate weekly group discussions of over 20 students, teaching MatLab fundamentals, and providing feedback on coding projects
- Support in-class lectures by answering student questions and presenting previous Matlab projects involving Machine Learning

EDUCATION & CERTIFICATIONS

UC Irvine
B.S. in Mechanical Engineering
Irvine, CA
Sept 2017 – June 2021

DataCamp
Data Science Professional Certificate
San Diego, CA
Nov 2023 – Jan 2024