

Diego Cerda, M.Sc.

 GitHub |  diegocerdac |  dcerdac@gmail.com |  +1 (346) 232 5758

SUMMARY

Senior Energy Analyst with over 7 years of experience in the energy sector. I specialize in driving innovation and developing advanced analytical tools for solar, battery storage, geothermal, wind, and oil & gas industries. Proficient in machine learning for regression, clustering, classification, and forecasting. Highly skilled in data analytics, SQL, and Python.

WORK EXPERIENCE

NarrativeWave Inc.

Senior Data Science Manager

2022 -

- Performed Root Cause Analysis on OCPP data from ABB and Schneider Electric EV charging stations, using NLP sentiment analysis to identify and resolve anomalies in charging systems. This effort reduced system incidents by 24%. Additionally, it developed a knowledge database that improved troubleshooting efficiency for major power infrastructure providers.

Lead Data Scientist

2020 - 2022

- Integrated LLMs with Canary and Cygnet SCADA systems to extract insights from historical operator notes in WellView software. Fine-tuned BERT model for job cost prediction, enabling proactive budget management and cost deviation prevention.
- Developed predictive models for energy generation, enabling advanced forecasting, power curve analysis, and proactive identification of equipment wear and long-term operational challenges, ensuring optimal O&M responsiveness.

Data Scientist / Machine Learning Engineer

2018 - 2020

- **KPI Development & Visualization:** Designed and implemented IEC-compliant dashboards tracking critical solar performance metrics including performance ratio, soiling, degradation, DC health, tracker alignment, and expected power. Created visualizations such as waterfall losses, heatmaps, and power curves. Calculated power loss due to underperformance and downtime allocation, integrating PVLib and PVSyst data for comprehensive analysis and optimization insights.
- Developed and deployed machine learning models for upstream oil production optimization, enabling automated detection of flow anomalies, equipment blockages, and performance deviations across wells and facilities. Improved operational efficiency by identifying issues in real time, including unexplained production drops, abnormal fluid ratios, and instrumentation drift.

Tafer Hotels & Resorts

Data Analyst

2017 - 2018

- Developed a forecasting model for hotel demand using mixed-integer linear programming and an additive time series method (capturing yearly, weekly, and daily seasonality), integrating data from multiple sources, leading to increased occupancy rates and optimized revenue streams.
- Collaborated closely with sales and marketing teams, employing the model's constraints and outputs to determine competitive pricing and innovative marketing strategies, aligning with demand forecasts and market dynamics.

Hewlett Packard Enterprise

Systems and Solutions Architect

2014 - 2016

- Developed a cross-platform support application for SAS, Qlikview, and Informatica, significantly enhancing incident response times.
- Led the adoption of Agile methodologies, ensuring consistent delivery of business value through efficient, iterative sprints.

EDUCATION

- 2016 - 2017
- M.Sc. in Advanced Computer Science at **University of Sheffield**
Dissertation: Developed and compared machine learning models, including SVM, Decision Tree, Multilayer Perceptron, Regression, Logistic Regression, and Mixed Gaussian Model, for predicting energy efficiency in buildings; leveraged advanced statistical techniques for performance analysis. [Code Example](#)
- 2009 - 2014
- B.Sc. in Electronics Engineering at **Instituto Tecnologico y de Estudios Superiores de Monterrey**

LANGUAGES

Spanish (Native), English c (Fluent)

PROJECTS & CERTIFICATIONS

- [Climatebase Fellow](#) - Cohort 7 - Climatebase Fellowship. (2025–Present).
- [EcoVentus](#) - 2nd Place Winner - Huawei Developer Competition. Mexico City, Mexico (2025)
- [HCIA-IA](#) - Huawei Certified. Shenzhen, China (2024)
- [Wildfire Drone Detection](#) - OpenAI Hackathon Latinoamerica. Santiago, Chile (2024)

SKILLS

Machine Learning	Scikit-learn, YOLO, TensorFlow, Keras, PyTorch, BERT, Transformers, Prophet, Catboost, LightGBM, XGboost
Data Analysis	NumPy, SciPy, Pandas, Jupyter Notebook, Statistics, Excel
Data Visualization	Matplotlib, Seaborn, Plotly
Natural Language Processing (NLP)	NLTK, Gensim, TfidfVectorizer, Word2Vec
Explainable AI	SHAP, Lime
Cloud Computing	EC2, S3, Lambda, RDS, CloudFormation, IAM, EKS, Route53, AWS-CLI, VPC, SageMaker
MLOps - DevOps	Docker, K8s, Circle CI/CD, MLFlow, Weight & Biases
Programming Languages	Python, Java, Matlab, C, Unix, Shell-Bash Scripting
Web Development	Javascript, HTML, PHP, Node.js, Flask, Django, Celery, FastAPI
Databases	SQL, NoSQL, Postgres, MongoDB
APIs	REST API, SCADA Systems
Development Tools	Eclipse, Jira, Confluence, GitHub