Diego Cerda, M.Sc.

GitHub | diegocerdac | dcerdac@gmail.com | (+52) 331.114.1974

Summary

Seasoned Senior Data Scientist with over 6 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 - 2024

- Streamlined onboarding by architecting data pipelines and automating ETL processes, reducing time-to-value to 60 days. Led cross-functional teams as Product Owner, managing development and delivery of innovative products to meet business objectives and user needs.
- Analyzed OCPP data from ABB and Schneider Electric EV charging stations and implemented NLP sentiment analysis to identify electric vehicle charging system anomalies, reduce incidents, and build a knowledge database, significantly improving customer success rates 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.
- Developed and implemented advanced monitoring system for Battery Energy Storage Systems (BESS), leveraging BMS data to identify voltage anomalies across multiple levels (cell, rack, etc.). Calculated degradation levels based on market participation in ERCOT and CAISO, and created interactive dashboards for real-time monitoring and analysis.

Data Scientist / Machine Learning Engineer 2018 - 2020

- **Advanced Energy Analytics:** Deployed analytics for 15+ GW across diverse energy applications, including RWE-Innogy's offshore wind turbines. Implemented predictive maintenance for Enel, Longroad, RWE, and Novatus using anomaly detection methods (Isolation Forest, PCA) to optimize operations.
- **Predictive Solutions:** Implemented predictive analytics and anomaly detection for energy giants including ExxonMobil, Shell, and Enel Solar, leveraging advanced techniques (neural networks, LSTM, pattern recognition) to optimize operations and enhance equipment health.
- **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.

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 (Fluent)

Projects & Certifications

HCIA-AI - Huawei Certified. Shenzhen, China (2024)
Wildfire Drone Detection - OpenAI Hackathon Latinoamerica. Santiago, Chile (2024)

Skills

Machine Learning	Scikit-learn, TensorFlow, Keras, PyTorch, BERT, Transformers, Prophet, Catboost, LightGBM, XGboost
Data Analysis	NumPy, SciPy, Pandas, Jupyter Notebook, Statsmodels
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
DevOps	Docker, K8s, Circle CI/CD
Programming Languages	Python, Java, Matlab, C, Unix, Shell-Bash Scripting
Web Development	Javascript, HTML, PHP, Node.js, Flask, Django, Celery
Databases	SQL, NoSQL, Postgres
APIs	REST API, SCADA Systems
Development Tools	Eclipse, Jira, Confluence, GitHub