Emmanuel De Vera

Entry Data Scientist / ML Engineer

A machine learning enthusiast with versatile academic and practical experience across applied deep learning, database management, and data communication. I mission to leverage data-driven insights to propel innovation in any application.

+64 027 897 0720   edevera.ortiz@gmail.com

Relevant Experience

Big Data in Practice │ Spark MLlib │ Jun 2025

- Developed a complete regression pipeline using Apache Spark ML to predict student GPA from a dataset of over 900,000 records and 35 features. Exploratory data analysis through heatmaps and histograms informed appropriate scaling strategies and feature selection.

- Deployed code on a Hadoop cluster using HDFS, Yarn, and shell scripting, demonstrating hands-on experience with large scale analytics in a distributed computing environment.

- Investigated the impact of PCA dimensionality reduction and analyzed feature coefficients which were communicated verbally alongside a well-paced and structured presentation.

Natural Language Processing │ Keras API │ Mar 2025

- Engineered machine learning solutions to categorize real news articles as one of four genres: world, sports, business, and science. TF-IDF vectorization was explored in conjunction with scikit-learn classifiers Naive Bayes, KNN, and DTrees to benchmark classical approaches.

- Executed essential NLP preprocessing techniques such as tokenization, stopword removal, punctuation filtering, and lemmatization to convert raw text into clean and structured input.

- Architected a CNN using Sequential API, integrating pre-trained GloVe word embeddings to enhance semantic understanding, obtaining a reliable model with a test accuracy of 92.3%.

Computer Vision Project │ Pytorch │ Sep 2024

- Designed a convolutional neural network using Pytorch for multi-class image classification, achieving 30% greater accuracy over an MLP via attentive tuning of convolutional layers, filter sizes, pooling, and batches, alongside dropout and weight decay to minimize overfitting.

- Faster and efficient tuning was attained by learning to capitalize CUDA GPU acceleration.

- An inventive solution to address limited training size was utilizing stable diffusion to generate synthetic images. This enriched model robustness by training under diverse compositions.

Tertiary Education
Victoria University of Wellington

Master of Artificial Intelligence
BSc, Computer Science, Data Science
2022 - 2025

- BSc consists of a joint specialization in Artificial Intelligence and Cybersecurity
- Tangiwai scholarship received in 2022
- Dean's List earned in 2022, 2023, and 2024
- Graduated BSc with an 8.58 GPA / 3.88 GPA

Secondary Education
Tawa College
2017 - 2021

- Attained consistent top in class awards for physics, maths, economics, and design IT.
- Awarded Proxime Accessit to the Dux, 2nd highest academic student in my final year.

Supplementary Experience

Relational Database Modeling with PostgreSQL

Mar 2025

- Transformed raw data and a project synopsis into a functional 8-table relational database via postgreSQL.
- Enforced apt foreign key, check, delete, and update constraints to guarantee data integrity and fluidity.
- Crafted optimized SQL queries featuring aggregation, joins, and nesting to facilitate precise data extraction.

Data Acquisition and Communication through R

Mar 2024

- Automated web scraping through the rvest package, streamlining data retrieval from diverse web sources.
- Leveraged the tidyverse suite to perform effective data cleaning and transformation in preparation for modeling.
- Formulated polished visualizations using ggplot2 and RMarkdown to convey findings clearly and insightfully.

Team Management and Software Design using Java

Sep 2023

- Led collaborative software architectural efforts using CRC cards and UML diagrams to deshroud team goals.
- Oversaw git version control, resolving merge conflicts and issues to establish a smooth development pipeline.
- Applied Java principles polymorphism, generics, and stream operations to recreate 1989 Chip's Challenge.

FSM 2D Physics Passion Project via Javascript

Sep 2022

- Utilized matrices and trigonometry to produce robust and streamlined physics, enemy tracking, and collisions.
- Built modular finite state machines to administer state transitions with simplicity and maintainability in mind.
- Experimented with p5.js to port the project onto an accessible website with integrated responsive design.