



Erick Díaz

Data Scientist

ErickDiaz

erickdiazs

ErickDiaz

Technical Expertise

Data Science
ML Ops
Generative AI
Bioinformatics

Education

Master's Degree in Mathematical Engineering and Computing | Nov. 2020 - 2022
UNIR - Universidad Internacional de La Rioja

Big Data and Artificial Intelligence Specialization | 2020-2021
UNIR - Universidad Internacional de La Rioja

Bioinformatics Specialization | 2019
Universidad San Carlos de Guatemala

Computer Science Engineer | 2013
Universidad San Carlos de Guatemala

Skills

Core Expertise: Machine Learning, Natural Language Processing (NLP), Large Language Models (LLMs), Deep Learning, Computer Vision

Scripting Languages: Python (7 years), R, PySpark

ML Frameworks: TensorFlow, PyTorch, Sklearn, EvalML

Big Data Ecosystems: Apache Spark, Hadoop, MapReduce, HDFS, HBase, Zookeeper, Hive

Database Management: SQL (6 years), Data Warehousing (2 years)

Languages: English, Spanish

Work Experience

Jul 2022 Today	Sr. Data Scientist - Vana Remote -Technical Lead of the Data Science Team, responsible for designing and overseeing data science projects. - Propose new Data Science experiments to explore areas of opportunity or improvement within the company. - Create and maintain risk data science services - Create models for micro-loan acceptance	3yrs 4 mos
Jan 2025 Today	Lecturer and Researcher - Universidad San Carlos, Escuela de Ciencias Físicas y Matemáticas	7 mos
Mar 2021 Jul 2022	Sr. Data Eng./Data Scientist - Zubale Remote - Develop a forecasting model for in-store demand (100+ Stores) - Asses and interview new members of the DS team - Collaborate in the development of a library in python for ETL's	1yrs 5 mos
Jan 2020 Mar 2021	Data Scientist - HealthCare.inc Guatemala - Working on ETL development using python and SQL - Creation of a machine learning model for cost and revenue forecast using Pythorch	1yr
Apr 2018 Dec 2019	Data Scientist 1 - Xoom Guatemala - Data wrangling and data pipelines for machine learning models using PySpark and Tensorflow	1 yr 9 mos
Feb 2016 Mar 2018	Big Data Engineer - Xoom Guatemala - Data pipelines in PySpark, Hadoop	2 yrs 2 mos

Research Projects and Publications

- Detect and classify *P. Vivax* malaria parasites in blood films, the goal is to expedite the malaria parasite detection. GitHub Repo, ICLR 2020 - Workshop Conference Poster

Certifications

2025	Retrieval Augmented Generation (RAG)	Coursera
2024	Generative AI with Large Language Models(LLMs)	Coursera
2020	Mathematics for Machine Learning Specialization - Linear Algebra - Multivariate Calculus - PCA	Coursera
Apr 2019	Introduction to TensorFlow for Artificial Intelligence, Machine Learning, and Deep Learning	Coursera
Mar 2018	Machine Learning	Coursera
Dic 2017	Neural Networks and Deep Learning License WEZQ3JAHWQAT	Coursera, deeplearning.ai
Aug 2017	Deep Learning Nanodegree Foundation	
Apr 2017	Big Data and Hadoop Developer Online Course License 20270414435	Grey Campus