Isaac Vergara

Duke University

📞 +1 619 3370

✉ isaacvergaram@hotmail.com

○ [isaacvm98](#)

in [Isaac Vergara](#)

Education

MSc in Data Science (Quantitative Finance Concentration)
2025-2027

Duke University
Durham, NC

Bachelor of Economics
2017-2021

Universidad Panamericana, Ciudad de México
CDMX

Microsoft Certified: Azure Data Scientist
2022

Microsoft Learn
Online

Natural Language Processing with Python
2022

Udemy
Online

Data Politics
2021

Universidad Panamericana, Ciudad de México
Online

Political and Economic Summer Program
2018

London School of Economics
London

Google Project Management Professional Certificate
2025

Coursera
Online

Experience

Fiscal Mind
2025-

Co-Founder

- Co-founded legal tech startup developing an AI-powered tax advisory platform for Mexican VAT (IVA) compliance, combining domain expertise with advanced ML techniques.
- Architected end-to-end ML pipeline using multi-agent workflows with Google Gemini 2.0 and OpenAI models, implementing RAG (Retrieval-Augmented Generation) with Neo4j knowledge graphs for legal document processing.
- Built production-ready FastAPI backend with Supabase integration, implementing intelligent caching using Voyage AI embeddings (cosine similarity >95% threshold) to optimize response times and reduce API costs.
- Engineered automated legal document processing system using PyMuPDF and regex parsing to extract and structure 500+ articles from Mexican tax law (LIVA, RLIVA, CFF), creating searchable knowledge bas.e

Enable Global
2023-2025

Semi Sr. Data Scientist

- Spearhead data science projects focused on Operational Technology (OT) data solutions to drive business outcomes. Communicate findings and results to stakeholders.
- Lead a comprehensive analysis to identify the root causes of defective steel pieces in the transformation process for a chassis fabricator, using several techniques like ANOVA, Anomaly Detection, Parametric and Non-Parametric tests. This analysis helped our client detect the causes and address them by installing a new heater system.
- Engineered and implemented an automated data quality web app using Python, incorporating cosine similarity algorithms and clustering techniques for robust deduplication. Leveraged Pydantic for data validation and schema

enforcement. This application has been used for cleaning master and spare parts data for several manufacturers

Entropia AI
2021-2023

Data Scientist

- Use state-of-the-art (SOTA) models, and cloud technologies like AWS and Azure to deliver high value. solutions to clients across a diverse range of industries.
- Create understandable dashboards with Dash as deliverables and end-to-end Data Science solutions.
- Act as Product Manager for several projects, leading a team of data analysts and scientists to deliver results on time and within budget.
- Deployment of models using Prefect for data pipeline workflow orchestration framework and an inference service using AWS ElasticBean for sales and demand forecasting.
- Architected and deployed a production ready Almost Ideal Demand System (AIDS) model, translating advanced economic theory (Deaton & Muellbauer, 1980) into a real-time production system. Implemented an automated pipeline that updates price elasticity estimations upon availability of new economic indicators, ensuring the model remains calibrated with current market conditions.
- Implemented a multi-level SARIMAX forecasting system for pharmaceutical distribution, improving prediction accuracy from 88% to 92%. Implemented hierarchical predictions across various business dimensions (SKU, store, region) to optimize budget allocation.

IQVIA
2020-2021

Consulting Intern

- Provided strategic support for pharmaceutical product launches, analyzing both quantitative and qualitative survey data from the healthcare sector.
- Designed and conducted structured interviews with healthcare professionals, managing end-to-end survey execution from questionnaire design to data collection.
- Synthesized complex market research findings into executive-ready PowerPoint presentations, translating data insights into clear recommendations for product launch strategies

Teaching Experience

Coderhouse
2022-2023

Tutor

- Tutor for the Data Science course, providing instruction and guidance to students in areas such as data manipulation, statistical analysis, and machine learning with Python.

Awards and honors

CONACYT Postgraduate Scholarship Abroad
2025-2027

Consejo Nacional de Humanidades, Ciencias y Tecnologías, Mexico

- Selected for prestigious " Postgraduate Scholarships in Science and Humanities Abroad 2025" scholarship to pursue master´s degree in Data Science.

Academic Scholarship
2025-2027

Duke University

- Awarded half-tuition merit scholarship for the master´s degree in Data Science.

Academic Scholarship
2017-2021

Universidad Panamericana, Ciudad de México

- Awarded half-tuition merit scholarship for the bachelor's degree in economics.

Basketball Scholarship
2014-2017

Universidad del Valle de Mexico, Lomas Verdes

– Awarded full-tuition athletic scholarship for High School education, granted selectively to only 5 players annually
– Maintained scholarship throughout the entire program by sustaining a minimum /10 GPA requirement

Projects and Competitions

– Developed a Spatial Lag Model to analyze poverty-crime relationships across Mexico's 32 counties, incorporating spatial dependence through Moran's I statistics under Dr. Jose Miguel Torres' guidance.
– Participant in 'Reto Banxico', Mexico's prestigious central bank competition, where our team presented comparative monetary policy analysis of emerging markets to Banco de México officials, benchmarking Mexico's monetary framework against other developing economies
– Logistic Regression model for Econometrics II to predict changes in European Soccer forward's market value where I had to obtain the market value data and statistics from transfermrkt.com.
– Published "Analyzing NBA Timeouts" article, implementing multilevel logistic models to analyze timeout effectiveness using play-by-play data; selected by Medium's editors for broader distribution through Boost platform.
– Secured 12th place (Silver Medal, top 1%) out of 1,200 teams in Kaggle's NCAA March Madness competition, implementing XGBoost and Generalized Linear Mixed Effect models for tournament prediction.
– Participated in NFL's Big Data Bowl 2024, developing a methodology to evaluate tacklers using tracking data, creating two new metrics: Time to Tackle Opportunity Above Replacement and Tackles Above Replacement.
– Created [Player Profile](#) a comprehensive NBA analytics platform built with Dash, featuring interactive dashboards for player evaluation, incorporating advanced metrics with an Added Value metric relative to contract, play-type analysis, year-over-year performance comparisons, expected points from shot selection and diagrams to evaluate player's assists.
– Developed Advanced Portfolio Optimization Suite implementing comprehensive pipeline integrating Markowitz mean-variance optimization, Black-Litterman model with investor views, and transaction cost modeling for Mexican equity portfolio; engineered Hidden Markov Model for regime detection enabling adaptive allocation strategies across bull/bear market cycles with automated regime transition analysis.
– Built VaR optimization framework using CVXPY with CVaR analysis, risk parity (Equal Risk Contribution), and Hierarchical Risk Parity methodologies; engineered production-ready backtesting system with realistic transaction costs, market impact modeling, and comprehensive performance metrics including Sharpe ratio, maximum drawdown, skewness, and kurtosis across 11 distinct portfolio strategies including robust Markowitz optimization under parameter uncertainty.

Leadership

Treasurer
2020-2021

Economics Student Society

– Served as Treasurer of the Economics Student Society, managing budget and organizing academic and professional development events

Basketball Team Captain
2016-2017

Universidad del Valle de México

– Led team to National High School Championship title as Team Captain, demonstrating leadership and team management skills.

Cloud Club Co-captain
2025-

Duke University

– Co-lead AWS Cloud Club initiatives focused on developing hands-on cloud computing projects to enhance student portfolio development

Technical Skills

Languages: Python, R

Modeling Experience: Time Series, Boosting Trees, Regression (Linear, Logistic), Clustering, Dimensionality Reduction, Neural Networks, NLP, Conditional Density Estimation, Portfolio Optimization, Risk Parity, Regime Detection, VaR Modeling, Multilevel Modeling, Spatial Lag Models, ANOVA, Parametric and Non-Parametric Tests

Libraries: Sklearn, Pandas, TensorFlow, PyTorch, XGBoost, SpaCy, Keras, sktime, CVXPY, HMMlearn, QuantLib, PyPortfolioOpt, PyMuPDF, OpenAI, Google Gemini API, yfinance

Data Visualization: Matplotlib, Seaborn, Plotly

Dashboards: Dash, Streamlit, PowerBI, Tableau

Cloud/Databases: MongoDB, MS SQL, PostgreSQL, MySQL, Neo4j, Supabase

Model Serving: FastAPI

Machine Learning Platforms: Azure Machine Learning Studio, AWS SageMaker, GCP Vertex

Other relevant skills: Docker, Prefect, Git, GitHub

OS: Windows, Linux

Mathematical Finance: Stochastic Calculus, Portfolio Optimization Markowitz, Black-Litterman, Risk Modeling, VaRCVaR Analysis, Regime Detection, Risk Management, Almost Ideal Demand Systems (AIDS), Price Elasticity Modeling, Demand Forecasting

Economics/Finance Methods: Spatial Econometrics, Behavioral Economics, Monetary Policy Analysis, Market Research, Survey Design

Advanced Optimization: Convex Optimization, Risk Parity, Hierarchical Risk Parity

AI/ML Technologies: RAG (Retrieval-Augmented Generation), Multi-agent Workflows, Voyage AI Embeddings, Knowledge Graph Construction

NLP: Legal Document Processing, Text Extraction, Regex Parsing

Business Applications: Legal Tech, Tax Compliance Systems, RegTech, Pharmaceutical Distribution, Manufacturing Quality Control, Healthcare Sector Analysis, Product Launch Strategy