

Jefferson Conza

Quito, Ecuador (GMT -5)

Número telefónico: +593 992701821, +51 953219234

E-mail: jeffconza3@gmail.com

LinkedIn: [jefferson-conza-46779b59/](https://www.linkedin.com/in/jefferson-conza-46779b59/)

GitHub: [JeffersonConza](https://github.com/JeffersonConza)

Estudiante de Matemática | Ingeniero de Machine Learning

Estudiante senior de Matemática e ingeniero de Machine Learning con más de 2 años de experiencia en el desarrollo de modelos, despliegue de soluciones de visión artificial y creación de herramientas impulsadas por LLM. Sólida experiencia en Python, R, SQL, Stata y aplicaciones contenedorizadas como Docker. También cuento con experiencia en TensorFlow, PyTorch y NumPy. Busco un rol desafiante e innovador como Ingeniero de Machine Learning donde pueda aplicar y desarrollar mis habilidades.

Deseoso de contribuir en proyectos orientados a datos enfocados en la resolución de problemas complejos.

Experiencia Laboral

Analista de Datos en Ecuadorian Development Research Lab (LIDE)

Ene 2025 a Oct 2025

- Refactoricé y generalicé algoritmos de R para Métodos de Control Sintético, mejorando la reproducibilidad del código y permitiendo la aplicación del modelo en distintos conjuntos de datos con parámetros temporales variables (años).
- Realicé ingeniería de características integrando indicadores económicos externos (PIB, Población) para resolver valores faltantes y llevé a cabo Análisis Exploratorio de Datos (EDA) para validar la calidad de la información.
- Implementé modelos de inferencia estadística utilizando R y Stata para cuantificar el impacto de las políticas públicas en Ecuador, apoyando directamente los procesos de toma de decisiones basados en datos.
- Documenté hallazgos de investigación en informes técnicos, mejorando la difusión de conocimientos en control sintético (modelado para comparación de resultados).

Tecnologías Principales: R, STATA, Excel, python

Analista de Datos en Belozfi (YC W22) (Estacional)

Ago 2025 a Sep 2025

- Desarrollé y perfeccioné modelos de aprendizaje automático para la detección de fraude en tiempo real utilizando Python.
- Diseñé y propuse modelos de campañas orientados a datos basados en análisis SQL para maximizar las tasas de conversión.
- Gestioné y optimicé los esfuerzos de publicación de Optimización de Motores de Búsqueda (SEO) mediante el uso práctico de Google Search Console. Esto incluyó el monitoreo del rendimiento del sitio web, la identificación de problemas técnicos de SEO, el análisis de consultas de búsqueda para informar la estrategia de contenido y garantizar la visibilidad del contenido y el crecimiento del tráfico orgánico.

- Presenté informes semanales (1-2 páginas) para mostrar hitos del proyecto, métricas de SEO e indicadores clave de rendimiento, asegurando la alineación de las partes interesadas y la transparencia sobre el progreso.

Tecnologías Principales: Python, SQL, Slack, MetaBase, Google Search Console

Machine Learning Fellow en Anyone AI (Tiempo parcial)

Ene 2025 a Jun 2025

- Desarrollé modelos de análisis de sentimientos y de incrustación de palabras para clasificar reseñas de películas, aprovechando BoW, TF-IDF y técnicas de aprendizaje profundo para detectar la polaridad en texto no estructurado.
- Predije la capacidad de pago de créditos hipotecarios utilizando un conjunto de datos de más de 350 mil transacciones, logrando un ROC AUC >0.72 con modelos DecisionTree, XGBoost y LightGBM.
- Construí una tubería de datos y visualizaciones para el análisis del comercio electrónico en Latam (2016–2018), revelando tendencias en ingresos, categorías de productos y retrasos en las entregas por regiones.
- Desplegué un modelo de clasificación de imágenes con 82% de precisión para identificar la marca y el modelo de automóviles a partir de imágenes; contenedorizado como un servicio web basado en API utilizando Docker y AWS.

Tecnologías principales: Python, Numpy, Pandas, Scikit-learn, Keras, PyTorch, Docker, AWS, SQL, Seaborn, Matplotlib

Anotador de Datos en TALOV

Abril 2022 - Agosto 2022

- Recolecté y curé un conjunto diverso de imágenes de moneda, asegurando la representación de diversas condiciones (p. ej., bien conservadas, dañadas, dobladas, entintadas) para mejorar la precisión de la detección de objetos para usuarios con discapacidad visual.
- Realicé aumentación de imágenes en varios conjuntos de datos de más de 100 imágenes (para moneda), generando variaciones rotadas (90°, 180° y 270°) para mejorar la generalización y robustez del modelo.
- Actualicé el conjunto de datos con moneda de reciente lanzamiento, manteniendo la relevancia y precisión del modelo en escenarios del mundo real.
- Contribuí a una red neuronal que logró más del 90% de precisión de reconocimiento en la detección de moneda, mejorando la accesibilidad para personas con discapacidad visual.

Tecnologías Principales: Roboflow, Python, Excel

Proyectos

Reconstrucción Isotrópica de Datos de Bio-Imagen (Investigación Confidencial)

Ago 2025 - Presente

- Desarrollo de una arquitectura *CycleGAN* en PyTorch para realizar traslación de imagen a imagen no pareada, mapeando dominios fuente axiales (xz) a objetivos laterales de alta fidelidad (xy) para corregir la distorsión anisotrópica.
- Diseñé una tubería de ingestión de datos de alto rendimiento para procesar más de 120,000 parches de imágenes (resolución 128x128) a través de 12 diversos conjuntos de datos (Núcleos, Membranas, Canaliculos Biliares), manejando la extracción de tensores complejos **.npz**.
- Documentación de la arquitectura de la red neuronal y los resultados de la experimentación en LaTeX, estableciendo una base de referencia para la investigación del equipo de ingeniería.

Tecnologías Principales: Python, PyTorch, GitHub, NumPy, LaTeX

Centro de Pronóstico de Energía | [Ver en GitHub](#)

- Implementé y realicé pruebas comparativas de modelos de regresión (*Regresión Lineal*, *Random Forest* y *XGBoost*) tanto en Python como en R, con el fin de comparar, seleccionando Random Forest como el modelo de producción basado en un rendimiento óptimo de RMSE.
- Refactoricé una base de código heredada en una estructura de paquete modular, integrando tuberías de datos automatizadas para entrenamiento histórico y pronóstico de varios días.

Tecnologías Principales: Python, R, Docker, FastAPI, Streamlit, XGBoost, Scikit-learn

Detección de Neumonía con Transformers | [Ver en GitHub](#)

- Implementé la arquitectura *CheX-DS* (*DenseNet121* + *Swin Transformer*) en PyTorch, utilizando una Pérdida Asimétrica Ponderada para manejar eficazmente el desequilibrio de clases en conjuntos de datos médicos.
- Logré un 99% de Recuperación (Sensibilidad), superando un modelo de referencia ResNet50 al reducir la Pérdida de Prueba (0.29 frente a 1.03) y minimizando los falsos negativos críticos para la seguridad del paciente.
- Desplugué el modelo de diagnóstico a través de un tablero interactivo de Streamlit, contenedorizado con Docker para garantizar una inferencia reproducible en diversos entornos informáticos.

Tecnologías Principales: Python, PyTorch, Swin Transformer, Docker, Streamlit.

Habilidades

Habilidades Técnicas: Python, R, Julia, Kotlin, SQL, Stata, Bash

Frameworks y Librerías: PyTorch, TensorFlow, Keras, Scikit-learn, Pandas, NumPy

Metodologías Ágiles: Scrum

Herramientas y DevOps: Docker, Git/GitHub, VS Code, PyCharm, Antigravity CLI, Wolfram Mathematica, LaTeX

Idiomas: Español, nativo; Inglés, C1

Educación

Universidad Yachay Tech

Feb 2019 - Jun 2027

Licenciatura en Matemáticas

Anyone AI

Ene 2025 - Jun 2025

Programa de Ingeniería de Machine Learning

AWS Cloud Institute

Oct 2025 - Jun 2026

Certificado de Desarrollador de AWS Cloud

Cursos

DataCamp | Fundamentos de Llama

Febrero 2025

DataCamp | Aprendizaje Profundo en Python (PyTorch)

Enero 2025

NVIDIA | Fundamentos de Computación Acelerada con CUDA Python

Septiembre 2024

Platzi | Visualización y Storytelling con Latinometrics

Septiembre 2024