

# Ana (Alejandro) Muñoz Gutiérrez

CDMX, México | [ana.clavius@gmail.com](mailto:ana.clavius@gmail.com) | +52 5511520035 | [linkedin.com/in/anadatascientist](https://linkedin.com/in/anadatascientist) | [monfiz.github.io/Anas\\_Portfolio/](https://monfiz.github.io/Anas_Portfolio/)

## EDUCACIÓN

- Capacitación Google Gen AI | GenAI in GCP** CDMX, México | **11.2023 – 12.2023**
- Implementación de modelos de lenguaje para Chatbots personalizados, entrenamientos clásicos de machine learning con Model Garden y AutoML.
- Instituto Tecnológico Autónomo de México | Maestría en Ciencia de Datos** CDMX, México | **08.2021 – 10.2023**
- Tesis de Maestría **Internship** en **Alemania**, BAYLAT beca
  - Host del podcast Algoritmos + Datos + Sociedad
- Universidad Iberoamericana | Licenciatura en Ingeniería Física** CDMX, México | **08.2016 – 05.2021**
- Intercambio académico en el Instituto Nacional de Ciencias Aplicadas, departamento de **Bioinformática**, Lyon **Francia** | Otoño 2019
  - Escuela de verano y **seminario de investigación** en el Instituto Nacional de Astrofísica Óptica y Electrónica, Puebla México | Verano 2018
  - Vicepresidenta** y Consejera Académica, representante de estudiantes | 2017-2018
  - Divulgación científica en la estación de radio **Ibero 90.9FM**, programa semanal | 2020-2022
  - Conferencista** en la **Biblioteca Vasconcelos**, Divulgación científica | Junio 2019 y Agosto 2021
- UCSanDiegoX | Data Science MicroMaster** Remoto | **11.2020 – 03.2021**
- Módulos de: Python for data science, Statistics & Machine Learning.

## HABILIDADES

### Lenguajes de programación y librerías

Python (avanzado), R (intermedio), SQL (intermedio), TensorFlow, Pytorch, Scikit-Learn, Keras, Numpy, Pandas, Matplotlib, LangChain, Prophet, vertexai language models. API Open AI.

### Herramientas de software

Microsoft Office, Docker, Github, Google Cloud, BigQuery, VertexAI, Lookerstudio, AutoML, Chat GPT, Google Sheets, GenAI.

### Comunicación & Interpersonal

Comunicadora efectiva, Habilidad de enseñar y divulgar, Toma de decisiones & problem-solver. Inglés (C1), Francés (B2), Alemán (A1).

## EXPERIENCIA

- Grupo Herdez**, Empresa multinacional de alimentos | **Científica de Datos** CDMX, México | **08.2023- Actualidad**
- Lideré el desarrollo de **modelos de lenguaje** personalizados usando los servicios de **GCP**. Estos modelos eran capaces de conectarse a datos en BigQuery y regresar información requerida por el usuario como tablas, gráficas e información descriptiva.
  - Encargada del **forecast masivo** semanal para las cadenas de clubes (SAMS, Costco, City Club). Trabaje con datos históricos de ventas en **series de tiempo**. Implementé y ajusté los modelos que ya usaban en otras cadenas desde cero. Mejoré sustancialmente el proceso de verificación y corrección de predicciones, además del ajuste de los modelos.
  - Implementé desde cero el forecast para las cadenas de retail de la empresa, Moyo y Chilim Balam. Usando algoritmos clásicos de machine learning y la herramienta de **AutoML** en GCP. Esto implicó un EDA profundo, limpieza de datos, implementación de modelos y comparación de métricas de varios modelos.

**Greenspin, Tech Startup de Agricultura | Data Scientist Intern**

**Würzburg, Alemania | 10.2022- 03.2023**

- Colaboré con un equipo de 3 personas para implementar una red neuronal capaz de detectar cultivos de trigo a partir de **imágenes satelitales** y estadísticas de agricultura. Fue una colaboración entre la industria (Greenspin) y el **centro de investigación en inteligencia artificial CAIRO**.
- Implementé desde cero modelos de **redes neuronales convolucionales** en **Pytorch** y preprocesamiento de datos satelitales. Apliqué técnicas de **transfer learning** a áreas con baja calidad de datos.
- Superé las métricas del estado del arte y abrí nuevas oportunidades de investigación. Mi código se usa como base para nuevos proyectos. Reduje el tiempo de preprocesamiento un 60%.

**Tutorías privadas en física, matemáticas y programación | Independiente**

**CDMX, México | 06.2020 – 08.2023**

- Durante más de tres años enseñé y regularicé a alumnos de preparatoria y universidad en diversas materias: cálculo, estadística, álgebra lineal, física clásica, programación con Python. Disfruto mucho enseñar y soy buena haciendo accesibles conceptos técnicos y o abstractos.

---

## PROYECTOS PERSONALES

---

### Predictor Cancelaciones Hotel

Lenguaje: R

Dataset/Packages: Kaggle, Hotel booking demand data sets. Tidyverse, Dplyr.

Código fuente:

[https://github.com/Monfiz/Anas\\_Portfolio/tree/main/Project\\_2](https://github.com/Monfiz/Anas_Portfolio/tree/main/Project_2)

### Detección Tumores Cerebrales

Lenguaje: Python

Dataset/Packages: Cheng, Jun (2017): brain tumor dataset. Tensorflow, Numpy.

Código fuente: <https://github.com/mig-calval/brain-tumor-detection>

### Detección Remota de Trigo

Lenguaje: Python

Dataset/Packages: MODIS datos satelitales, estadísticas agricultura NASS, CDL. Pytorch, Gdal, Numpy.

Código fuente:

[https://github.com/Monfiz/Anas\\_crop\\_remote\\_sensing](https://github.com/Monfiz/Anas_crop_remote_sensing)