Ana (Alejandro) Muñoz Gutiérrez

CDMX, Mexico | ana.clavius@gmail.com |+52 5511520035 | linkedin.com/in/anadatascientist | monfiz.github.io/Anas_Portfolio/

EDUCATION

Google Gen AI Training | GenAi in GCP   Mexico City, Mexico | **11.2023 – 12.2023**
- Implementation of language models for custom Chatbots, classic machine learning trainings with Model Garden and AutoML.

Autonomous Technological Institute of Mexico | *Master's in Data Science*   Mexico City, Mexico | **08.2021 – 10.2023**
- Master thesis **internship** in **Germany**, BAYLAT **scholarship**
- Host of the podcast Algorithms + Data + Society

Ibeoamerican University |*Bachelor's degree in **Physics Engineering***   Mexico City, Mexico | **08.2016 – 05.2021**
- Academic exchange at the National Institute of Applied Sciences, **Bioinformatics** department, Lyon **France** |Autumn 2019
- Summer school and **research seminar** at the National Institute of Astrophysics Optics and Electronics, Puebla Mexico | Summer 2018
- **Vice-president** and Academic Counselor, representing students | 2017-2018
- Popularization of science in the radio station **Ibero 90.9FM** | 2020-2022
- **Speaker** at **Vasconcelos Library**, Popularization of science | June 2019 and August 2021

UCSanDiegoX | *Data Science MicroMaster*   Remote | **11.2020 – 03.2021**
- Modules of: Python for data science, Statistics and Machine Learning.

SKILLS

Programming Languages and Libraries	**Software Tools**	**Communication & Interpersonal**
Python (advanced), R (intermediate), SQL (basic), TensorFlow, Pytorch, Scikit-Learn, Keras, Numpy, Pandas, Matplotlib, LangChain, Prophet, vertexai language models. Open AI API.	Microsoft Office, Docker, Github, Google Cloud, BigQuery, VertexAI, Lookerstudio, AutoML, Chat GPT, Google Sheets, GenAI.	Effective communicator, Ability to teach & mentor, Decision-maker & problem-solver. English (C1), French (B2), German (A1), Spanish (native language).

EXPERIENCE

Grupo Herdez, Multinational food company | **Data Scientist**   Mexico City, Mexico | **08.2023-Present**
- Led the development of custom **language models** using **GCP** services. These models were able to connect to data in BigQuery and return information required by the user such as tables, charts, and descriptive information.
- In charge of the weekly **massive forecast** for club chains (SAMS, Costco, City Club). Worked with historical sales data in **time series**. Implemented and adjusted models that were already used in other chains from scratch. Substantially improved the process of verification and correction of predictions, in addition to the adjustment of the models.
- Implemented from scratch the forecast for the company's retail chains, Moyo and Chilim Balam. Using classic machine learning algorithms and the **AutoML** tool in GCP. This involved deep EDA, data cleaning, model implementation, and comparison of metrics from various models.

Greenspin, Agricultural Tech Startup | *Data Scientist Intern*   Würzburg, Germany | **10.2022- 03.2023**
- Collaborated with a 3-person development team to implement a neural network capable of detecting wheat crops from satellite images and agricultural statistics. This was a collaboration between industry (Greenspin) and the **research center of artificial intelligence** CAIRO.
- Implemented **convolutional neural network models** from scratch in **Pytorch** and preprocessing of satellite data. Applied **transfer learning** techniques to areas with low-quality data.
- Outperformed the state-of-the-art metrics and opened new research opportunities. My code is used as base for new projects. Reduced data preprocessing time by 60%.

Private tutoring in mathematics, programming and physics | Independent   CDMX, Mexico | **06.2020 – Present**
- For over three years, I taught and regularized high school and university students in various subjects: calculus, statistics, linear algebra, classical physics, programming with Python. I enjoy teaching very much and am good at making technical or abstract concepts accessible.

PROJECTS

Hotel Cancellations Predictor	**Brain Tumor Detection**	**Wheat Remote detection**
Language: R	Language: Python	Language: Python
Dataset/Packages: Kaggle, Hotel booking demand data sets. Tidyverse, Dplyr.	Dataset/Packages: Cheng, Jun (2017): brain tumor dataset. Tensorflow, Numpy.	Dataset/Packages: MODIS satellite data, NASS agricultural statistics, CDL. Pytorch, Gdal, Numpy.
Source Code: https://github.com/Monfiz/Anas_Portfolio/tree/main/Project_2	Source Code: https://github.com/mig-calval/brain-tumor-detection	Sourse Code: https://github.com/Monfiz/Anas_crop_remote_sensing