

# ADRIÁN AUGUSTO FERRER ORGAZ

DATA SCIENTIST AND MACHINE LEARNING ENGINEER



(+52) 553-072-0140



[AdrianFO](#)

[ONLINE PORTFOLIO](#)



[aaf01416@gmail.com](mailto:aaf01416@gmail.com)



[AdrianFO-16](#)



## PROFILE

Data Scientist and Machine Learning engineer with international academic experience and keen interest in [Artificial Intelligence](#) and [Data Mining](#). Looking forward to growing in [multidisciplinary teams](#) through high attention to detail and commitment.

## LANGUAGES

- Spanish (native)
- English (advanced C1)
- French (beginner A1)

## EDUCATION

### Instituto Tecnológico y de Estudios Superiores de Monterrey, México

BS Data Science and Mathematics Engineering.

- August 2019 – 2023 (present).
- Beca al Talento Académico (Academic Talent Scholarship) (2019).

### École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Exchange Semester in Data Science and Computer Science Master program courses.

- Fall Semester 2022– 2023
- Highlighted courses taken:
  - [Applied Data Analysis \(CS-401\)](#)
  - [Machine Learning \(CS-433\)](#)
  - [Applied Machine Learning \(MICRO-455\)](#)

## PROFESSIONAL EXPERIENCE

### IT & Data Analyst @ Singular Beacon (May 2023 – present)

- Beginner [Odoo \(UI\) development](#): [Data engineering](#) and manipulation via xml-rpc and Qweb view module development.
- [Data engineering and analysis](#) solutions designer using [Python](#). Data manipulation and handling from REST sources.
- [Data engineering](#) with Azure Data Factory. [ETL designer](#) and data management supervision.

## RESEARCH EXPERIENCE

### “Hacking the Oscars data story” (2022)

“Hacking the Oscars” is a project developed as part of the [Applied Data Analysis EPFL course](#). Elaborate [data wrangling](#) performed on [CMU movie summary corpus dataset](#) along with IMDb datasets to understand [movie success](#) through [data-driven approach](#). Statistical differences between highest rated and lower rated movies are explored. Use of [unsupervised learning techniques](#) such as Latent Dirichlet Allocation for movie summary plots for feature engineering and interpretation. Participated actively in [code and analysis design](#), developed the [presentation web page \(data story\)](#).

### “Credit Scoring using Advanced Multivariate Linear Regression Analysis” (2021)

Project developed under the supervision of Mexican corporation Pretmex. Tasked with [expanding](#) a previous [empiric-developed credit score](#) scale based on previous client data. The data was analyzed with the aid of the newly created scale and with [advanced correlation](#) and [multivariate linear regression analysis](#). Separation and classification with these statistics and other machine learning techniques to predict client credit standing. Lead [code developer](#) and [project designer](#).

### “E-commerce vehicle capacity vehicle routing problem (CVRP) optimization” (2021)

Project developed under the supervision of Mexican nationwide department store Coppel. An [alternate solution methodology](#) is proposed for NP-hard CVRP using clustering [local optimization algorithms](#) and [regular travelling salesman](#) problem.

### “Semi-supervised anomaly detection based on autoencoders” (2020–2021)

Supervision and evaluation of various trained models, [development of code](#), [Neural Network architectures](#), and code execution on several environments. Development under supervision of Ph.D. Miguel Ángel Medina Pérez.

### “Multilayer Neural Network for death probability prediction in COVID-19 patients” (2020–2021)

Presented in 51 Congreso de Investigación y Desarrollo, Tecnológico de Monterrey. Prediction of [death probability](#) for a patient based on certain comorbidities and other factors. Concept as an [auxiliary tool](#) to aid in [prioritizing medical attention](#). Led [code architecture](#) and [web app development](#). A [streamlit web app](#) is presented with the results of the model iterations.

## SKILLS

- [Machine Learning](#)
- Data Analysis and Statistics
- [Programming](#)
- [Autodidact / Self – Taught](#)
- Problem Solving
- Flexibility and Adaptability

### Programming Languages:

- Experienced:
  - [Python](#), R, Matlab, Wolfram Mathematica
- Beginner/ Intermediate:
  - HTML5, CSS3, JavaScript, C#, Dart, C++

### Frameworks/Systems:

Scikit-Learn, Keras, PyTorch, Streamlit, Dash, seaborn, matplotlib, Odoo, Office.

### Complementary Skills:

- Advanced linear regression analysis
- Statistical Time Series Analysis
- Beginner web development

### Other Software:

- Adobe Photoshop (intermediate)
- DaVinci Resolve (beginner)
- Unity (beginner)

## CERTIFICATIONS

- IELTS Band 8 CEFR C1 Level Certification (2022)
- Machine Learning Scientist with Python, Datacamp (2020)
- Machine Learning Fundamentals with Python, Datacamp (2020)

## AWARDS

- 3rd place in an NDS Cognitive Labs sponsored machine learning team competition, HackMx (2021)
  - [Fraud detection](#) simulation in an e-commerce web application. Led modeling and machine learning tasks ([Anomaly detection](#) and classification algorithms).

# PROJECTS

## Online Portfolio (2022–present)

Web Portfolio developed with the purpose of learning HTML, CSS, JS, etc. and also complement my academic/professional profile while doing so. Linked in header, also available at [Portfolio](#).

## “Digital Signature (DS) scheme implementation application” (2022)

Project developed under the supervision of non-lucrative Mexican organization Teletón. A multidisciplinary project with the objective of making [cryptographic technology](#) accessible for [non-technical users](#). Main [designer of code, app functionality and user cycle](#). Developed the main logic and Python code blocks to be used in the final application, assisted in UI and app design. The final product consists of the app implementation of the DS algorithm and a promotion [prototype web page](#).

## “Topological Data Analysis (TDA) in gravitational wave data using time delay embedding” (2022)

Analysis of clean and noisy gravitational wave data using time delay embeddings. [Vietoris–Rips complex](#) data persistence homology analysis. Classification of noisy and clean gravitational waves based on engineered features from [persistence homology](#).

## “Control Theory for building vibration control in earthquakes” (2022)

Study of [modern building LTI models](#), active and passive actuators concepts for vibration control during earthquakes. System modeled in [Simulink](#) and code in state space representation. Analysis of [analytic and numeric solutions](#) facing seismic excitation before and after the implementation of control techniques. Lead code developer and project designer.

## Ruin Probability with Crammer–Lundberg Risk Model (2021)

Project developed under the supervision of nationwide corporation AXA. Provided a dataset, based on monthly income, client claim amount distribution, and corporation reserve, [ruin probability](#) is computed. Results obtained via [analytic solution](#) of the [Crammer–Lundberg](#) model and [simulation](#). Lead code developer and project designer.