Cristóbal **Alcázar C.**

DATA SCIENTIST · MACHINE LEARNING ENGINEER

Santiago, Chile

☐ (+569) 951-63-585 | ✉ alcazar.cristobal90@gmail.com | 🏠 alkzar.cl | alcazar90 | alcazar90

Work Experience

European Southern Observatory (ESO)
Paranal, Chile

MACHINE LEARNING ENGINEER (GRADUATE INTERNSHIP) - SOFTWARE GROUP — *Aug. 2023 - Jan. 2024*

- Engineered a RESTful API to streamline access to diverse data sources in the observatory, including Jira tickets and software version control logs, fostering enhanced accessibility and integration of user information across various platforms.
- Devised and implemented a reporting tool application using Streamlit Python's library, featuring customizable summarization levels tailored to user perspectives through utilizing Large Language Models (LLMs). Integrated Jira tickets and associated commits into comprehensive reports spanning specified timeframes.
- Spearheaded prototyping initiatives utilizing vector databases focusing on OpenSearch and the Retrieval Augmented Generation (RAG) pattern. Contributed to developing projects showcasing applications such as similarity search, chatbots, and Q&A apps.

Compraquí - BancoEstado subsidiary
Santiago, Chile

MANAGER PAYOPS AND ANALYTICS — *Oct. 2020 - May. 2021*

- Oversight of the complete payment process for merchants, leading a team of four analysts to ensure a seamless end-to-end payment workflow.
- Create and lead the analytics teams to provide data and insights to different company areas and the management team, facilitating informed decision-making for the business. Push forward initiatives to train people and evangelize the company using scripting languages like Python and R for analytics and developing a data products mindset.
- Designed, developed and documented the critical processes within the Payops and Analytics area in the company's early stage.

PAYOPS AND RISKFRAUD LEAD — *Oct. 2019 - Oct. 2020*

- Led the Payops and RiskFraud team in processing payments for over 50k merchants nationwide, focusing on minimizing fraud and preventing service misuse in the card payment network.
- Developed various data products, including reports and dashboards, utilizing code to automate data extraction, preprocessing, and presentation.
- Implemented the risk and fraud process, creating features based on transaction details like transaction amount, deviation from average tickets by merchant category, and transaction frequency. Developed an outlier detection model to calculate a transaction score.

Central Bank of Chile
Santiago, Chile

ECONOMIC FINANCIAL ANALYST - MACROECONOMIC STATISTIC DIVISION — *Oct. 2016 - Oct. 2019*

- Responsible for deploying and maintaining the statistical data process to elaborate the local fixed-income quarterly publications. Includes processing multiple data sources using technologies such as SQL, Python, and R. Translating from granular data to macro-financial concepts consistently with other macroeconomic statistics.
- Write technical reports, publish statistics, and make periodic presentations to stakeholders about local fixed-income markets' economic and financial scenarios.
- Implement and publish the Handbook of Securities Statistics for Chile, one of the nine Special Data Dissemination Standards (SDDS plus) items, the most advanced standard for economic-financial statistics dissemination worldwide. At that time, Chile was one of the two countries in LATAM to meet these standards and have comparable statistics with other countries. More information here.

Teaching Experience - Extracurricular

Universidad de Chile
Santiago, Chile

MDS7203: DEEP GENERATIVE MODELS (TEACHING ASSISTANT) — *Aug. 2023 - Dec. 2023*

- Contributed to designing and leading a master's level data science course, emphasizing theoretical fundamentals and algorithms for building and training generative models. Covered key topics like Variational Autoencoders, Score-based models, Diffusion models, and Autoregressive models. Actively participated in crafting curriculum materials and hands-on activities for a dynamic learning experience.

Data Visualization Talk, Women in Data - Guatemala Chapter (NPO)
Remote

SPEAKER — *Oct. 2021*

- A talk focused on the skills for moving from a beginner to an intermediate user in data visualization. The topics included composing data visualisations by layers, knowing some of the best practices, and customising many visual features to maximise communication using python.
- Slide, video recording, and code for my talk: `https://github.com/alcazar90/WomenInData-Guatemala2021`

Data Visualization Workshop, Ministry of Social Development and Family
Santiago, Chile

SPEAKER — *Aug. 2019 - Sep. 2019*

- A five-session Data Visualization workshop focused on both hands-on and theoretical material addressed for professionals in the public sector.

Universidad de Chile   *Santiago, Chile*

EI3001-2: DATA JOURNALISM WITH PYTHON (PART-TIME PROFESSOR)   *Mar. 2019 - Aug. 2019*

- An interdisciplinary course dictated for engineering and journalism students; focused on providing data science tools to apply to data journalism projects.

ECP7-1: DATA ANALYSIS WITH R PROGRAMMING (PART-TIME PROFESSOR)   *Sep. 2018 - Dec. 2018*

- This was the first course imparted at the Institute of Communication and Image (ICEI) that gives data science tools for data wrangling and visualization for applying to data journalism projects.

Education

Universidad de Chile   *Santiago, Chile*

MSc IN DATA SCIENCE   *Mar. 2022 - Expected Jul 2024*

- HuggingFace Student Ambassador: I volunteer to organize talks and events in my local community.

Universidad Adolfo Ibañez   *Viña del Mar, Chile*

B.S. IN BUSINESS ADMINISTRATION, MSc IN FINANCE   *Mar. 2009 - Dec. 2015*

- Minor in Political Science.

Skills

Programming	Python scientific stack (numpy, pandas, scipy, matplotlib), PyTorch, JAX, open-source ML ecosystem (HuggingFace, W&B), R (tidyverse, ggplot2), SQL
Languages	Spanish, English
EDX	Introduction to Computer Science and Programming Using Python, Introduction to Computational Thinking and Data Science, Artificial Intelligence (CSMM.101.x), Machine Learning (CSMM.102.x), Robotics (CSMM.103.x)
Coursera	Introduction to Machine Learning in Production, Machine Learning, "Graph Search, Shortest Paths, and Data Structures", "Divide and Conquer, Sorting and Searching, and Randomized Algorithms", The Data Scientist's Toolbox, R Programming, Getting and Cleaning Data, Exploratory Data Analysis, Reproducible Research, Statistical Inference, Regression Models, Managing Big Data with MySQL

Writing

Digital literacy, computational thinking, and literate programming.   *Lom Ediciones*

AUTHOR   *Oct. 2019*

- A chapter published in the book "Inteligencia Artificial y Bienestar de las juventudes en America Latina" (Brossi) that reviews and reflects on ideas such as computational thinking and literate programming, particularly highlighting the importance of establishing programming as an elementary skill in Latin America.