Francisco Albornoz

Toronto, ON | (647) 619-5402 | francisco.albornozcambiaso@georgebrown.ca | Portfolio | LinkedIn

Education and Certifications

Postgraduate Certificate Applied AI Solutions Development   Jan 2024 – Dec 2024
George Brown College – Toronto, ON (GPA: 3.93)
- Developed Tableau dashboards and scorecards leveraging LAPD crime data, revealing actionable insights for a private security company operating in Los Angeles. Identified key business trends and areas for operational improvement.
- Carried out a Neural Network NLP project to summarize scientific articles into abstracts, fine-tuning open-source language models from Hugging Face to improve their performance and accuracy.
- Developed a real-time thief detection system using OpenCV that leverages a baseline image for comparison, accurately identifying and flagging significant deviations in a live video stream of the same scene.
- Implemented a prescription parser using an NLP technique named Conditional Random Fields (CRF) to accurately extract medication information, such as dosage and frequency, from unstructured text input.

Diploma in Applied Econometrics   Aug 2020 – May 2021
University of Chile – Santiago, Chile
- 10-month program covering experimental design, causal inference with observational data and time series analysis.

Diploma in Big Data   Aug 2017 – May 2018
Pontifical Catholic University of Chile – Santiago, Chile
- 144-hour program about the fundamentals of ML, recommenders, Hadoop, parallel computing, and graph theory.

Bachelor's in Economics   Mar 2010 – Dec 2015
University of Chile – Santiago, Chile
- Coursework included: Statistics, Econometrics (experimentation, causal inference and time-series), Finance (CAPM model), Microeconomics, Marketing and Business Intelligence.

Tableau Data Analyst Certificate   May 2024

AWS Cloud Practitioner Certificate   Aug 2022

Professional Experience

MACH (Fintech)   Aug 2022 – Jan 2024
Data Scientist – Santiago, Chile
- Enabled targeting of high-potential-value app users by integrating PySpark machine learning models with transactional data through SQL pipelines on Apache Airflow.
- Trained predictive machine learning models using AWS cloud infrastructure (SageMaker, Athena and Glue) and PySpark.
- Increased marketing spend efficiency by conceiving and evaluating high-ROI initiatives through ideation and A/B testing process, leveraging tools like Amplitude and Agile methodologies.
- Assessed ROI of referral campaigns via advanced causal inference techniques when A/B Testing was not feasible.

University for Development   Apr 2023 – Aug 2023
Data Science and AI Bootcamp Instructor – Santiago, Chile
- Delivered 5-hour weekly lectures on machine learning with Python' sklearn and graded students' work.

Marketing Advance   May 2022 – Aug 2022
Data Scientist (Contractor) – San Francisco, USA (Remote)
- Optimized advertising budget allocation through Marketing Mix Modelling, leveraging Meta's Robyn package.
- Performed validation and fine-tuning of Marketing Mix models with out-of-sample data.
- Estimated ROI for different advertising channels and provided compelling visualizations and reports to stakeholders.

WOM (Telecommunications)   Jan 2020 – Oct 2021
Data Scientist – Santiago, Chile
- Guided deployment of new network infrastructure, identifying areas where upgrades would reduce churn the most.
- Tracked network deployment results via geospatial dashboards blending network KQIs with churn and NPS metrics.
- Boosted conversion of tele sales by developing purchase propensity models with LightGBM.

Ministry of Education, Government of Chile   May 2020 – Aug 2020
Data Specialist (Contractor) – Santiago, Chile
- Developed R scripts to automatically create custom forms for a survey of higher education institutions, and then consolidated the filled forms into a single, ready-to-use Stata dataset by using Python scripts.

Walmart Chile

Data Scientist – Santiago, Chile **Nov 2018 – Jan 2020**
- Trained and deployed clustering and predictive models to personalize marketing communications for loyalty program members.
- Developed an internal R package to facilitate the delivery of advanced analyses (e.g. market basket using the *apriori* algorithm) and the execution of complex recurrent SQL queries.
- Quantified the impact of personalized marketing campaigns through A/B testing, leveraging the CUPED technique to increase statistical power in the presence of small treatment effects and limited sample sizes.

Data Analyst – Santiago, Chile **Nov 2017 – Nov 2018**
- Provided data-driven insights to marketing and merchants, leveraging econometrics and clear storytelling.
- Automated complex, ad-hoc reports through R Markdown and Tableau, increasing throughput of our team.
- Assessed the impact of key, C-level decisions via causal inference techniques such as Difference in Differences.

KMA Asset (Real Estate Investment Trust) **Aug 2017 – Jan 2018**
Data Consultant and Power BI Dashboard Developer – Santiago, Chile
- Leveraged R to extract, transform, and load (ETL) data from multiple CRM systems.
- Developed interactive Power BI dashboards showcasing critical sales funnel KPIs and trends.

Personal Projects and Coursework

Data Pipeline on Apache Airflow with AWS Integration   Jun 2024
End-to-end ETL data process with Apache Airflow that communicates with Amazon's cloud services such as S3 (bucket storage) and RDS (databases) to transform data into a format that is more suitable for dashboards and make it available on PostgreSQL.

Blog Posts about Causal Inference with R   Jun 2021 - Ongoing
Authoring a series of accessible blog posts explaining key causal inference concepts and methods (e.g., Potential Outcomes, A/B Testing and Matching) for a non-academic audience.

Advent of Code Solutions with Python and R   Dec 2023
Wrote solutions and detailed explanations to puzzles from this global competition, demonstrating coding and communication skills.

Big Data with PySpark (DataCamp Track)   Dec 2023
Completed 24-hour course track on PySpark, covering from Spark basics to predictive modelling and recommendation engines with large datasets.

Dashboard on AWS Monitoring Engagement with Politicians' Tweets   May 2022
Developed an R Shiny app displaying interactive visualizations about data retrieved using Twitter's API. Data was auto refreshed and stored on S3, and the app was deployed on Amazon's EC2.

DevOps on AWS (Coursera Specialization)   Mar 2022
Completed this 30-hour course track, gaining familiarity with CI/CD pipelines, infrastructure as code (IaC), and test automation using AWS services (e.g., CodePipeline, CloudFormation, CodeDeploy).

Analyzing My Music Collection with Python and R   Dec 2021
Imported and wrangled iTunes data from XML file with Python's *ElementTree* and *pandas*, then used the *reticulate* package to integrate R in a single analysis script and output a nicely formatted table.

Solutions to Exercises of "*Introduction to Statistical Learning*"   Sept 2020
Solved exercises from this foundational textbook, demonstrating proficiency in key machine learning algorithms and statistical concepts (e.g., XGBoost, Bootstrap, SVM).

Volunteering Experience

Collision Conference – Toronto, Canada **Jun 2024**
Volunteered at Collision, North America's fastest-growing tech conference, assisting with data collection and attendee tracking on the Measurement team, and welcoming and registering attendees on the Registration team.