Jason Zafir, PhD

zafirov@gmail.com Personal Website

LinkedIn Github +971547152680

United Arab Emirates, Canada, EU, Panama

Education

PhD, Anderson School of Management,
University of California, Los Angeles (UCLA)
2018 – 2023

Master's of Science, Applied Economics,
Department of Applied Economics, HEC Montréal
2012 – 2015

Bachelor's of Commerce, Finance,
John Molson School of Business, Concordia University
2008 – 2011

Diploma of College Studies, Computer Science,
Dawson College
2005 – 2008

Publications — Authorships

A Comprehensive Look at the Empirical Performance of Equity Premium Prediction II, *The Review of Financial Studies (Top 5 Journal in field)*

FREEtree: A Tree-based Approach for High Dimensional Longitudinal Data With Correlated Features, *arXiv — Statistics: Machine Learning*

Improving Inference with Machine Learning: Application to CEO Turnover, *Dissertation*

Private Market Portfolio Forecasting and Optimization Using Machine Learning, *United States Patent and Trademark Office (Provisional Patent)*

Skills

Python • R • SQL • Web Scraping • NNs • NLP / LLMs • PySpark • Foundry • R Shiny • LangChain • XGBoost • LightGBM • PySpark

Professional Experience

ML Engineer — AI Lead,
Integrated Data Intelligence Ltd (Startup)
2023 – present | Abu Dhabi, United Arab Emirates
- Startup success forecasting using different AI algorithms, including gradient boost based and NLP based methods.

Data Scientist, *EPAM (NYSE: EPAM, MktCap: $20B)*
2022 – 2023 | Remote
- Consultant: Time series forecasting investor behavior using a variety of machine learning algorithms. Event categorization, article clustering and sentiment analysis through NLP.

Analyst, *Letko Brosseau (private fund, AUM: $27B)*
2014 – 2016 | Montreal, Canada
- Forecasting and modeling: emerging market economics, energy markets, corporate profits.

Intern — Quantitative Analysis,
CDPQ - Québec Public Pension Plans (AUM: $286B)
2013 | Montreal, Canada
- Built PCA autoregression tool to forecast inflation.

Intern — Macroeconomic Analysis,
CDPQ - Québec Public Pension Plans (AUM: $286B)
2013 | Montreal, Canada
- Forecasting and modeling: unemployment, gross domestic product, demographics, productivity.

Intern — Capital Markets, *Scotiabank*
2010 – 2011 | Montreal, Canada
- Data organization, interpretation & visualization.

Intern — Trading Floor,
Healthcare of Ontario Pension Plan (AUM: $114B)
2009 | Toronto, Canada
- Software development for trading floor operators.

Intern — Software Development, *CEDIB*
2008 | Lyon, France

Professional Training

Forecasting with Machine Learning, *ML Course*
- Time-series forecasting with XGBoost / LightGBM: data wrangling, model tuning and evaluation.

Machine Learning Projects

pandas-ai, *Contributor*
Library that adds Generative AI capabilities to pandas.

FREEtree: Tree Method for High Dimensional Longitudinal Data, *CRAN R package creator*

Earnings Sentiment Analysis: A Comparison of Transfer Learning Methods (BERT, ALBERT, XLNet), *Natural Language Processing*

PropCast— Deep Inception-Convolutional LSTM Networks for Forecasting Spatio-Temporal Transitions in Property Index Heatmaps

Publications — Research Assistance

Minimum Wage and Individual Worker Productivity: Evidence from a Large US Retailer,
Published, Journal of Political Economy
- Analysis of minimum wage changes on store and employee outcomes of large US retailer (~5,000 stores). Linking together HR data, performance, turnover, scheduling, payments, geography, revenues and web scraping variables.

Government Debt Management: The Long and the Short of It,
Published, The Review of Economic Studies
- Statistical analysis of US debt dynamics.

Unemployment Insurance with Limited Commitment Wage Contracts and Savings,
Published, BE Journal of Macroeconomics
- Researched conditions of unemployment benefits.

Household Search and the Aggregate Labor Market,
Published, The Review of Economic Studies
- Statistical analysis using US CPS household data.

Counterproductive Worker Behavior After a Pay Cut, *Published, Journal of the European Economic Association*
Difference-in-difference regression modeling of how workers reacted to a pay cut in a sales call-center.

Awards

UCLA Anderson doctoral fellowship, *$140,000*

FRQSC - Provincial doctoral scholarship, *$84,000*

Nominated - Best Master's Thesis (2015)

Academic Excellence Award (2012 & 2013)
Highest average of the Master's program.

Academic Excellence Award (2007)
Highest average of the Computer Science program.

S2DS, *Europe's largest data science training bootcamp*
2023 | London, United Kingdom
- Intensive, industry-linked project-based training for top PhD students.

Kenneth Woods Portfolio Management Program,
John Molson School of Business, Concordia University
2009 – 2012 | Montreal, Canada
- Hands-on training in investment management working with a real-life portfolio of over $5 million.

Teaching Assistance

MGMT 439: Tools and Analysis for Business Strategy (introduction to data science), *Graduate level (UCLA)*

MGMT 264A: Market Research (regression, dimensionality reduction and clustering),
Graduate level (UCLA)

MGMT 403: Financial Accounting,
Undergraduate level (UCLA)

ECO 680109: Applied Econometrics,
Graduate level (HEC Montréal)

ECO 600112A: Economic & Corporate Organization,
Graduate level (HEC Montréal)

ECO 380607: Introduction to Econometrics,
Undergraduate level (HEC Montréal)

Personal Tutor,
Economics, Computer Science, Data Science.

Machine Learning PhD Classes

ENGR C247: Neural Networks & Deep Learning,
UCLA Engineering Department, Grade: A

BIOSTAT 273: Classification and Regression Trees,
UCLA Biostatistics Department, Grade: A+

COM SCI 263: Natural Language Processing,
UCLA Computer Science Department, Grade: A

STATS 232C: Cognitive Artificial Intelligence (Reinforcement Learning),
UCLA Statistics Department, Grade: A+

Languages

English
(Native)

French
(Fluent)

Bulgarian
(Fluent)