Yabra Muvdi

yabramuvdi.github.io

Email : yabran.muvdi@gmail.com

EDUCATION

Barcelona School of Economics — Barcelona, Spain
Master Program in Data Science; GPA: 8.57/10.0 — *Sept. 2019 – Sept. 2020*
- **Relevant courses**: Machine Learning (A+), Data Warehousing and Business Intelligence (A), Data Visualization (A-), Text Mining (A+).

Universidad de los Andes — Bogotá - Colombia
Bachelor's Degree in Economics; GPA: 4.44/5.0 — *July 2013 – Dec. 2017*

EXPERIENCE

ETH Zürich - Data Science and Economics Group — Zürich, Switzerland
Doctoral Researcher, Member of ETH AI Center — *Sept. 2023 - Present*
- **LLMs in Economics**: Currently writing a chapter for the Handbook of Economic and Language about large language models and their applications in economics.
- **Recommender system**: Partnered with the Swiss Government (SECO) to develop a recommender system for the public employment service (RAV) to increase the quality of job matches. Integrated strategies to reduce potential biases and provided explanations for the machine learning model predictions.
- **Explainable AI**: Developing a taxonomy of explainable machine learning techniques and a field experiment to test how they affect people's perception of automated decisions.
- **Teaching**: Supported the development of three different courses (GitHub).

Machine Learning Consultant — Remote
Selected projects — *June. 2020 - Present*
- **Data extraction with LLMs**: Currently developing NLP tools for the Inter-American Development Bank to identify structural reforms in Latin American countries from the text of newspapers.
- **Custom chatbots for market research**: Designed and developed multiple custom chatbots to interact with automatically generated transcripts of market research interviews.
- **Income inequality visualization**: Developed a tool to interact with household income data en explore the distributional consequences of different shocks (Demo).

University College London - Barcelona School of Economics — Remote
Researcher in Economics and Machine Learning — *March 2021 - Aug. 2023*
- **NLP projects**:
  - Partnered with Lightcast to develop a language model (finetunned distilBert) to predict the presence of remote work on over 1 billion job postings. Webpage - Media - Model
  - Demonstrated methods for algorithmic text analysis in economics. Article - GitHub
- **Transactions data**: Developed a method to characterize firms within a trade network with low dimensional numeric representations (i.e. embeddings). GitHub - Poster - Blog post
- **Summer Schools**: Taught practical sessions in multiple summers schools (CEMFI 2022, IESE 2023, CEMFI 2023, IESE 2024) on the use of machine learning for textual and unstructured data.

Instiglio — Bogotá - Colombia
Data Analyst — *Jan. 2018 - Aug. 2019*
- **Model development**: Led the design of predictive models to forecast the performance of educational programs in Liberia. Supported the design of an innovative financing mechanism for a five-year educational project in the country.
- **Data analysis**: Collected, cleaned, analyzed, and visualized historical data of all the company's projects to produce relevant insights for the Board's annual meeting.
- **Business development**: Supported the business development efforts of the company by writing concept notes and formal proposals to prospective clients. Two of these efforts led to important projects for the company.
- **Achievements**: Got promoted from intern to analyst level 2 in a year.

Skills

- **Technical skills**: Causal inference, machine learning, explainable ML, natural language processing, large language models (chatbots, finetuning, RAG, prompting), algorithmic fairness, data visualization

- **Programming Languages**: Python, R, SQL, Stata   **Technologies**: GCP, AWS, Docker, Linux, Heroku

- **Languages**: Spanish (native), English (working proficiency), French (intermediate)