

DIEGO BERMUDEZ

412-519-6530 ◊ dabermud@andrew.cmu.edu ◊ linkedin.com/in/dabermudezs

EDUCATION

Carnegie Mellon University, Heinz College, Pittsburgh, PA Master of Information Systems Management – Business Intelligence and Data Analytics	December 2025
• GPA: 3.85 • Relevant courses: Applications of LLMs, AI Engineering, Deep Learning, Entrepreneurship.	
University of Los Andes, Bogotá, Colombia Master of Arts in Economics	April 2023

• **Thesis:** Access to financial services as a mechanism for consumption smoothing in the context of violence.
• **Relevant courses:** Advanced Econometrics, Topics in Machine Learning, Financial Risk Analysis.

University of Los Andes, Bogotá, Colombia
Bachelor of Arts in Economics

• **Relevant courses:** Financial Microeconomics, Advanced Macroeconomics, Data Mining and Applications.

SKILLS

- **Programming Languages:** Python, SQL, R, Java, VBA, JavaScript, STATA, MATLAB, HTML5, Bash.
- **Machine Learning & AI:** PyTorch, TensorFlow, scikit-learn, spaCy, Transformers, XGBoost, LightGBM, HuggingFace, NLP, A/B, LangChain, LlamaIndex, FAISS, RAGAs, vector databases, prompt engineering, semantic search.
- **MLOps, Data Engineering & Cloud Infrastructure:** MLflow, Weights & Biases, DVC, Docker, Airflow, Jenkins, GitHub Actions, CI/CD Pipelines, Apache Kafka, Spark, Hadoop, Dataiku, Airbyte, PostgreSQL, AWS, GCP, Azure.
- **Visualization & BI Tools:** Power BI, Tableau, Streamlit, Plotly Dash, GIS (GeoPandas, Folium).
- **Statistical & Econometric Analysis:** Impact Evaluation, Time Series Modeling, Regression, Clustering, Structural Breaks.
- **Languages:** English (Fluent), Spanish (Native).

PROFESSIONAL EXPERIENCE

Davivienda Bank (Daviplata) Data Scientist	Bogotá, Colombia April 2023 – August 2024
---	--

- Designed and deployed ML pipelines for georeferencing and customer segmentation, predicting business locations from transaction data to enhance targeted marketing across 1M+ businesses (**Python, scikit-learn, Docker, MLflow, PostgreSQL**).
- Integrated MLflow and Weights & Biases for experiment tracking and model versioning, ensuring reproducibility and governance across data teams (**MLflow, W&B, Git, Python**).
- Led A/B testing and causal analysis to identify behavioral drivers of digital wallet usage, increasing engagement and savings transactions by 20% (**Pandas, scikit-learn, CausalML, Airflow**).
- Built event-driven fraud detection and payment automation workflows via Kafka-based streaming pipelines, reducing operational latency by 40% (**Kafka, Airflow, FastAPI, Docker**).
- Migrated analytics infrastructure to GCP and implemented containerized deployments using CI/CD pipelines and version-controlled environments (**GCP, Docker, GitHub Actions, BigQuery**).

Business Intelligence Analyst	November 2021 – April 2023
-------------------------------	----------------------------

- Automated reporting workflows to the Colombian Financial Superintendence for 50K+ client petitions per month, ensuring compliance and traceability (**Python, SQL, ETL Pipelines**).
- Developed NLP models to classify customer feedback and measure satisfaction, reducing manual review by 6 hours per day (**spaCy, scikit-learn, REST APIs, Power BI**).
- Deployed predictive dashboards and BI systems for fraud-related payments and satisfaction metrics, enhancing decision-making efficiency (**Power BI, Docker, Dataiku**).

The Barometer Data Analyst	Bogotá, Colombia December 2020 – January 2022
-------------------------------	--

- Developed and implemented unsupervised NLP models to analyze 500K+ social media and web posts, identifying nine thematic clusters related to migration and inclusion (**Python, LDA, k-means, pandas**).
- Designed and optimized a text classification pipeline to detect sentiment and discrimination patterns, monitoring spikes in xenophobic discourse following major migration events (**scikit-learn, spaCy, TF-IDF, Python**).
- Created network analysis visualizations to map the spread and influence of online conversations, uncovering regional narratives and actor clusters (**NetworkX, Gephi, Python**).
- Produced data-driven reports in collaboration with the International Organization for Migration (IOM), guiding policy responses and awareness campaigns across five Latin American countries.

Ecopetrol S.A. – Business & Commercial VP Crude Trading Intern	Bogotá, Colombia January 2020 – July 2020
---	--

- Developed a quantitative model to measure value generation in crude trading operations, increasing transparency and negotiation efficiency by 25% (**SQL, Excel, Python**).

ACADEMIC EXPERIENCE

Carnegie Mellon University – FLM Haiti Capstone Project	Pittsburgh, PA August 2025 – Present
--	---

- Collaborating with FLM Haiti to design and implement an offline-first Electronic Medical Record (EMR) system adapted to Haiti's infrastructure limitations, such as intermittent internet and electricity (**Flutter, Supabase, PostgreSQL**).
- Leading the development of a cross-platform mobile and desktop application with synchronization features for secure data storage and access in low-connectivity environments (**Flutter, Supabase Auth, REST APIs**).
- Integrated **Azure Forms OCR** for digitization of historical paper medical records, automating patient data extraction and reducing manual data entry effort.

Carnegie Mellon University	Pittsburgh, PA
Research Assistant – Work Zone Crash Risk Project	November 2024 – Present
• Built causal inference and deep learning models integrating GIS and time-series data to estimate the impact of work zones on crash risk (TensorFlow , GeoPandas , scikit-learn , CausalML , GCP).	
University of Los Andes	Bogotá, Colombia
Teaching Assistant – Financial Regulation	January 2022 – July 2022
• Mentored students in financial regulation frameworks and policy design, incorporating simulations and real-world regulatory case studies.	
Center for Economic Development Studies (CEDE), University of Los Andes	Bogotá, Colombia
Professional Research Assistant	November 2020 – July 2021
• Applied econometric (Bai–Perron, Chow, Structural Breaks) and ML (DTW, Clustering) methods to analyze public transport demand differentials by income levels (R , Python , statsmodels , scikit-learn).	
• Processed mobility and smartphone usage data during COVID-19 using automated web scraping and transformation pipelines (Python , BeautifulSoup , Pandas , NumPy).	

AWARDS

- Merit-Based Scholarship – Heinz College, Carnegie Mellon University (Mar 2024)
- Asobancaria 7th Call for Papers – Best Paper in Economics (Sep 2023)
- COLFUTURO Loan-Scholarship – Academic Merit and Professional Excellence (May 2023)