Jaime Andrés Castañeda

Colombia
Phone: +57 300 222-4658
Email: jacasta2@gmail.com
URL: www.linkedin.com/in/jaimeandrescastaneda-datascience
Portfolio: https://jacasta2.github.io

PROFILE

Computer Science Engineer, Master in Systems Engineering and PhD in Economics. Has extensive experience leading and working on experimentation and data science projects to strategically support decision-making and drive business results. Has developed experimentation and data science apps, supervised and unsupervised machine learning models using Python, Power BI dashboards and statistical analyses to support decision-making and the development of improved business strategies. Skilled in Python, SQL, Machine Learning, NLP, Pandas, Numpy, Matplotlib, Seaborn, Scikit-learn, Keras, Streamlit, FastAPI, VS Code, Git, AWS SageMaker, Postgres and Power BI. Interested in data science and machine learning roles with career development as a technical leader.

PROFESSIONAL EXPERIENCE

Ministry of Information and Communication Technologies of Colombia

Data Scientist   *03.2023 – 12.2023*
- Implement extract, transform and load (ETL) processes and components that support data science initiatives
- Develop data science and machine learning products to support decision-making in different areas and drive business results
- Manage data science projects addressing data collection, preparation and storage and its exploitation for evaluating strategies and value generation

Achievements
- Led the construction of a composite index in Python with more than 40,000 school records to support the Internet connectivity prioritization policy
- Developed Python code to clean and extract text from legal documents through regular expressions and prompt engineering using the OpenAI API's ChatGPT
- Built 5 dashboards using Power BI to track the effectiveness of Internet connectivity and Information Technology training projects
- Created 2 updated map templates of Colombian towns and departments using Python and Power Query for the standardization of Power BI dashboards

Correlation One

Data Science Fellow   *03.2022 – 07.2022*
- Identify relevant data sources and collect the data necessary to address business problems using tools such as Python and its interaction with APIs
- Perform data preparation processes (cleaning, transformation and feature engineering) to ensure data quality and integrity
- Build and train predictive models using data science and machine learning techniques to support decision-making

Achievements
- Performed an extract, transform and load (ETL) process with more than 1.5 million traffic accident records using Python, SQL and Postgres
- Built heatmap visualizations using Python, SQL and Seaborn to analyze motorcycle accident risks, identifying that more than 50% of accidents occur with passenger vehicles and other motorcycles
- Developed an unsupervised machine learning model using Python and SQL that clusters highway corridors for road safety operations, prioritizing corridors that total close to 50% of accidents and more than 40% of fatalities

Universidad del Rosario

Associate Data Researcher *09.2021 – 01.2022*

- Design, direct and coordinate research programs and projects that generate insights for decision-making using data science solutions
- Collect relevant data using experimentation, A/B tests and surveys, among others, to address and provide solutions to strategic business questions
- Implement data science and machine learning solutions that drive business results using Python, R, Stata and other analytics tools

Achievements

- Designed and directed the implementation of a decision-making experiment with more than 500 participants to generate strategies to improve organizational service levels.
- Trained supervised machine learning models using Python that show that expedited shipping increases inventory availability by around 20%

Universidad del Rosario

Assistant Data Researcher *04.2016 – 09.2021*

- Direct, coordinate and participate in research and innovation spaces that support decision-making and generate value using data science and machine learning solutions
- Design and implement data collection instruments such as experiments, A/B tests, surveys and interviews that support analytics initiatives
- Develop machine learning models using analytics tools such as Stata, SPSS and similar to evaluate the impact of inventory policy and risk management on business performance

Achievements

- Designed and led the implementation of a decision-making experiment to test risk mitigation strategies that scale manufacturing levels
- Built supervised machine learning models using Stata that show how risk-sharing contracts increase inventory availability by around 5%
- Performed data science analyses that show how inventory transshipment increases supply chain efficiency by more than 10%

ADDITIONAL EXPERIENCE

Massachusetts Institute of Technology (MIT), *Data Researcher - Postdoctoral Fellow 2014 - 2015*
Università della Svizzera italiana (USI), *Data Researcher - Doctoral Assistant 2009 - 2013*

EDUCATION

Università della Svizzera italiana, *PhD in Economics 2009- 2013*
Universidad Nacional de Colombia, *Master's in Systems Engineering 2008 - 2009*
Universidad Nacional de Colombia, *Bachelor's in Systems and Informatics Engineering 2003 - 2007*

SKILLS

Technical skills: Data science, Machine Learning, Analytics, Natural Language Processing (NLP), Extract Load Transform (ETL), Web scraping, Data wrangling, Data visualization, Pandas, Numpy, Matplotlib, Seaborn, Scikit-learn, Keras, Streamlit, FastAPI, VS Code, Git, AWS SageMaker, Postgres, Power BI, Stata, SPSS, Excel, Qualtrics
Soft skills: Rigor, Organization, Teamwork, Proactivity, Planning, Communication