DANIEL GONZALEZ

• Toronto, ON • 4376620134 • daniel.andres.gonzalez.menza@gmail.com • linkedin.com/in/danielgonzalezmenza

Professional Summary

Dedicated and high-performing student actively pursuing a postgraduate degree in Big Data Analytics, supported by a bachelor's degree in Statistics and extensive hands-on experience as a Data Scientist and Machine Learning Engineer. My strength lies in developing sophisticated machine learning models, implementing MLOps strategies, and leveraging cutting-edge cloud technologies. With a strong dedication to ongoing growth, I keep pace with the latest advancements in technology. I have a proven history of deploying high-impact models that drive significant improvements in business operations. I excel in collaborative settings, effectively engaging with cross-functional teams to create actionable insights and innovative solutions. My contributions have received recognition for enhancing process efficiency, increasing scalability, and strengthening data-driven decision-making. I skillfully utilize tools such as Python, R, AWS, and GCP for model deployment, automation, and comprehensive data analysis, ensuring I consistently deliver outstanding results that elevate organizational success.

Skills

- ✔ Machine Learning Model Development
- ✔ Statistical Data Analysis
- ✔ Cloud Computing
- ✔ Data Pipeline Optimization
- ✔ Interpersonal Relationships
- ✔ Multitasking, fast-paced environment
- ✔ Team Work
- ✔ Model Deployment in Production
- ✔ Data Visualization and Reporting
- ✔ CI/CD for Machine learning models
- ✔ Listening Effectively
- ✔ Personal Initiative, self-driven
- ✔ Ability to adapt to new concepts
- ✔ Detail – Oriented
- ✔ Computer Literacy
- ✔ Problem-Solving
- ✔ Effective Communication with stakeholder
- ✔ Team Collaboration
- ✔ Critical Thinking
- ✔ Data Analytics
- ✔ Personal Development
- ✔ Oral – Written Communication
- ✔ Data Interpretations

Other Technical Skills & Tools

Data Science | Python Programming | R-Studio | GitHub | Docker| SQL | AWS S3 | AWS Step Functions | AWS ECR | GCP | Spark | Data Wrangling | Hadoop | Spark | Hbase | Machine Learning | Deep Learning | Artificial Intelligence | Natural Language Processing | Neural Networks | Linux | Cloud Computing | PowerBI | Tableau | ETL | Jupyter Notebook | Software Development | Continuous learning | Social Media Analytics | Unix Shell Scripting | Snowflake | Data architectures

Work History

MLOps Engineer | January 2024 – Present
Teamcore | Remote

- Built and maintained workflows using AWS services, including SQS for message queue management, Lambda for task orchestration, Step Functions for data processing, S3 for scalable data storage, and ECR for deploying containerized Python images with the correct libraries, ensuring optimal deployment, execution of models, and seamless cloud deployment.
- Standardized and optimized the continuous delivery pipeline for existing retail machine learning models, improving deployment efficiency and ensuring consistent high performance across production environments.
- Contributed to migrating machine learning models from AWS to Google Cloud Platform (GCP), enhancing scalability, reducing infrastructure costs, and streamlining cloud operations.
- Consistent application of agile methodology, with current experience using Jira to enhance team communication and ensure the continuous resolution of blockers, leading to more efficient project delivery.

Data Scientist | February 2019 – December 2023
Grupo Bit | Colombia

- Extensive experience using Ubuntu as my operating system and working with Docker images for Python to generate the correct library versions and deploy containers. My work involved regular use of the Unix command line for system management and executing various tasks. The use of Ubuntu eliminated 100% of compatibility issues with programs, and Docker facilitated the maintenance of models.
- Designed and implemented a DevOps-driven AWS pipeline for continuous integration of Lambda functions, triggered by code updates in GitHub. This solution eliminated 100% of deployment errors, reduced deployment time by 80%, and ensured seamless integration and rapid delivery of updates. By applying DevOps principles, this solution improved system reliability and efficiency, fostering a culture of continuous improvement and collaboration.
- Programming and deployed machine learning, deep learning, and statistical models in Python to generate actionable business insights, including a Recommendation System that increased client sales by 21%.
- Designed and deployed econometric models with exogenous data, applying data analytics to uncover additional insights beyond the client's data, creating new business opportunities. These insights increased information consumption by 15%.
- Took the initiative to research and calculate recommendation system metrics using a partition matrix, and developed an alternative data processing approach that reduced processing time and cut costs by up to 50%. Driven by curiosity and a focus on innovation, I researched and implemented new technologies, including the Jax library, to leverage GPU computational power, achieving a 95% reduction in processing time.
- Optimized algorithms for a forecasting model by leveraging Python, Numpy, and Polars libraries, resulting in a 98.47% improvement in processing time and an 85% reduction in peak memory usage for efficient implementation.
- Collaborated with cross-functional teams to develop client-focused communication channels, earning first place for innovation in the Grupo Bit Business Challenge.
- Worked with Jupyter Notebook hosted in Docker containers, leveraging custom Docker images to create isolated, reproducible, and scalable environments for data analysis and model development.
- Utilized the ETL program Spoon to inject data directly into Office Excel from various CSV sources for generating sales reports. Additionally, the implementation of slicers in the pivot table segment allows the sales team to interact with customers more effectively, resulting in a 5% increase in customer engagement.

Education

Postgraduate Degree in Big Data Analytics – Lambton College, Canada | 2025

- Big Data Tools – Final Project: Developed a project aimed at improving the accuracy of house price predictions by applying multiple linear regression and regularization techniques. Designed and implemented an interactive Tableau dashboard showcasing key visualizations of the dataset used for training and testing the predictive model, providing actionable insights and enhancing data interpretability.
- Big Data Fundament – Laboratory: Utilized Apache Spark to perform data analysis on a dataset containing over 10 million flight records, optimizing database queries to extract and transform data efficiently for insights. Designed and executed scalable ETL processes, leveraging Spark's distributed computing capabilities to handle large-scale data processing.
- Big Data Technology Solutions – Final Project: Managed a dataset containing 50,000,000 records by sending it to the Hadoop ecosystem, utilizing Pig's SQL-like language for Data Wrangling, and leveraging Spark to perform calculations, deriving actionable insights from the data.

Bachelor of Statistics – Universidad del Valle, Colombia | 2018

- Consulting - Credit Risk Management Model: Developed a logistic regression model to calculate the probability of default on credit payments. The odds analysis revealed that an exogenous credit score, which the company had initially excluded from its process, is crucial to consider due to its strong correlation with credit default probability.

Specialization in Business Intelligence and Big Data – Universidad del Valle, Colombia | 2018

Affiliations

Member, Sigma Statistical Studies Center | 2014 – 2018

- Provided tutoring and advisory services to students, supporting their understanding and application of statistical concepts and methodologies.
- Collaborated with peers to promote academic excellence and enhance statistical literacy among students.