Brayan David Aroca Gonzalez

✉ bdarocag@gmail.com   📞 +57 3187777645   in linkedin.com/in/brayan-david-aroca-gonzalez/
📍 Bogotá D.C., Colombia   ⌨ github.com/bdarocag   🌐 bdarocag.github.io

Profile

Skilled Data Scientist with a strong statistical background and over 5 years of experience in predictive modeling, data processing, and data mining to tackle complex problems. Proficient in Python (my principal programming language), R, and SQL, with extensive experience using Python libraries and frameworks such as TensorFlow, PyTorch, and Pandas to generate predictions, process data efficiently, and derive actionable insights.

Throughout my career, I have developed scoring models for financial institutions aimed at stabilizing and improving loan recovery rates. To achieve these objectives, I have designed neural networks using TensorFlow and PyTorch to classify profiles and calculate payment probabilities. Additionally, I have conducted descriptive analyses and applied multivariate techniques to build robust models. My expertise extends to leveraging methodologies from biostatistics and epidemiology, including survival analysis, risk assessment, and medical and demographic profiling, particularly within the healthcare sector.

Education

Manuela Beltrán University
MsC. Biostatistics   2022
GPA 4.0.

Francisco José de Caldas District University
Bachelor of Science in Biology   2018
GPA 4.0.

SENA
Software Programming Technician   2011
GPA 4.0.

Work Experience

Colsubsidio
Data Scientist   Feb 2024 - Act.
Applying statistical, artificial intelligence, and business techniques to develop projects with a significant impact on financial management within the healthcare sector. Designed innovative dashboards, applications, and reports to support various healthcare programs, including Breast Cancer, Cardiovascular Risk, and Hemophilia, among others. Proficient in R, Python, Dash, and Power BI to deliver actionable insights and improve decision-making processes.

Group COS BPO
Data Scientist   Nov 2019 - Jan 2024
Building and implementing data mining and machine learning models to solve complex problems and automate algorithms, enhancing efficiency and reducing manual effort. Delivering results through dashboards and leveraging business intelligence and programming tools such as Power BI, Python, and R.

Infometrika SAS
Asistente estadístico   Feb 2018 - Oct 2019
Supporting the production, analysis, and consolidation of databases to provide actionable insights for clients. Experienced in ETL processes and data structuring for advanced analytics.

Francisco José de Caldas District University
Field Assistant   Nov 2016 - Mar 2017
Creating reports, analyses, and summary tables to present key insights. Proficient in analyzing results
using the R programming language.

Skills

- SQL
- R
- Python

- Git
- Machine Learning
- Power Bi

- Azure
- STATA
- Epidemiology

Languages

- English - B2

- Portuguese - A2

Certifications

Course in Data Science in Public Health and Infectious Disease Modeling
Javeriana University, edX   2024
Applied Data Science

Data Science Training
Correlation one   2022
Data Science training Colombia cohort 4.0 - MinTIC. Certificate

R Programming
Coursera - Johns Hopkins University   2020
Advance skills in R Programming. Certificate

Publications

- Benavides I, Moritz S, Aroca-Gonzalez B, Romero J, Santacruz M, Selvaraj J. (2022). "kssa: Known Sub-Sequence Algorithm.". R package version 0.0.1.
- Aroca-Gonzalez BD, Gradstein R, González-Nieves LM. (2021). "¿En peligro o no? Distribución potencial de la hepática Pleurozia paradoxa en Colombia.". Rev. Acad. Colomb. Cienc. Ex. Fis. Nat. 45(174):260-271.
- Aroca, BD, Fernandez, D., Castro, V., Mejía, A., Tijaro, M.H., Cataño, E. Garzón, O., Neira, A., Olave, C., Niño, G., Martín, G., Coca, L. Betancourt y Pineda, W. (2017). "Líquenes y Briófitos del Piedemonte Llanero, Cubarral- Meta". Ed. Corporación para el Desarrollo Sostenible del Área de Manejo Especial La Macarena. ISBN 978-958-56208-3-4, 2017.

References

Works - Bosses

Esteban Gutierrez
Senior Data Scientist
+57 324 683 15 66

Christian Castro
Chief Data Officer
+57 320 816 36 91

Personal

Catalina Barrera
Fisioterapeuta
+57 350 794 02 06