

MARTÍN ARAMAYO

Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina

📧 [m-g-aramayo](#) 📧 [MartinAramayo](#) 📧 [MartinAramayo](#) 📧 [AllMylinks](#)
Mis blogs: en Inglés [blog/Martín](#), En español [blog/Martín/es](#), (cada link tiene posts diferentes).

Educación terminada

Instituto Balseiro, beca completa, Comisión Nacional de Energía Atómica Jan. 2018 – Dec. 2021

Maestría en Física estadística: Machine learning y simulaciones aplicadas & Lic. en Física Bariloche, Río Negro, Argentina

Experiencia

Mercado Libre Dic. 2022 – Oct. 2023 (11 meses)

Data Scientist Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina

- Implementación 3 iniciativas de negocio para la moderación del marketplace.
- Refractor/update de Dashboards Tableau del Mod Dev Team (800+ visitas) incluyendo: jobs SQL, métricas y UI/UX
- Implementación de un modelo de mejorado de imágenes para aplicarse a 3M+ imágenes del sitio.
- Participación en la Hackathon interna ChatGPT-4 de OpenAI con un resumidor de texto para moderación
- Tech: OpenAI, ChatGPT, Tableau, Bash, AWS, Big Query, Google cloud, GAN models, Pandas, Go, SQL, Computer Vision, Python, modelos & BI tools in-house.

Intellignos, Havas Mar. 2022 – Dic. 2022 (10 meses)

Data Scientist Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina

- Convertir una costosa métrica de $O(ne^n)$ a una de $O(n)$ con n el número de features consideradas en la métrica.
- ETL, análisis, soporte, documentación y modelado para clientes globales. Optimicé dos ETLs de dos clientes a una.
- Diseño de configuración, manejo, automatización y producción de la ETL para 3 clientes US.
- Tech: Bash, DataBricks, StreamSets, BigQuery, Dataproc, PySpark, Azure, Python, modelos & BI tools in-house.

UNICEF, ONG Jun. 2020 – Feb. 2022 (1 año 9 meses)

Data Scientist trainee Bariloche, Río Negro, Argentina

- Análisis de datos GIS y microcensos para un proyecto de estimación demográfica.
- Clasificación binaria del estado de ocupación de viviendas mediante machine learning.
- Tech: Scikit-learn, Numpy, Pandas, SPSS datasets, Scipy, SQL.

CNEA, Centro Atómico de Bariloche Jun. 2018 – Ene. 2022 (3 año 6 meses)

Machine learning trainee — Becario — Lic. & Mg. en Física estadística Bariloche, Río Negro, Argentina

- Implementación de una simulación basada en agentes con 300 000+ agentes. El pipeline de datos incluye: producción automática de config files y log files, scripts de análisis de salida, testing y benchmarking.
- Modelado matemático: sistemas dinámicos, modelos estocásticos, teoría de juegos, modelos de salud y tecnología.
- AI: Machine Learning, Deep Learning, LSTM, NLP, procesamiento de imágenes, encoding, feature-engineering y clustering.
- Manejo de datos: preprocesado de datos de campo, producción y análisis de datos de simulación, data mining.
- Software: Data analysis en Python. Experiencia en proyectos en salud, física y biología. CUDA orientado a simulaciones. Estructura de datos en C; árboles, pilas, colas.
- Presentación de los insights de proyectos con stakeholders.
- Simulación del sistema de control de orientación y corrección de trayectoria de un modelo de satélite con filtro Kalman.
- Cursos de imágenes médicas: Uso de OpenCV, ImageJ, Matlab y Python.
- Proyectos Data-driven con trasfondo estadístico.
- Charlas de comunicación de resultados experimentales y de conocimiento técnico adquirido.

Proyectos y habilidades técnicas

Simulación demográfica de agentes: Reproducción y recursos | Python, Bash, Pandas, Matplotlib Abril 2021

- Prototipo en OOP Python e implementación final en Pandas que permitió reducir a la mitad los tiempos de ejecución.
- Presentaciones públicas de resultados y visualizaciones comparando con modelos clásicos de dinámica poblacional.

Predicción demográfica | Sklearn, Keras, Matplotlib, Python Agosto 2020

- Exploración y análisis de datos para la etapa de feature engineering.
- Implementados clasificadores binarios para diferenciar viviendas habitadas/deshabitadas a partir de datos satelitales.
- Estimación de subpoblaciones a partir de una distribución de personas por viviendas.
- Estimación a partir de datasets pequeños conociendo solamente información satelital y datos de microcensos.

Lenguajes de Programación : Python, C, BASH, SQL

Development tools : Regex, Jupyter, IPython, git, postgresSQL, BigQuery, Dataproc, Streamsets, PySpark, REST APIS

Technologies/Frameworks : Linux, Docker, Pandas, Sci-kit learn, Keras, Numpy, Matplotlib

Documentation : Latex, Markdown, pandoc, Notion, Joplin, Mermaid, Hugo, Inkscape, VS code

Idiomas : Inglés (ESOL B2), Español (Nativo)