

# 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 (23 meses)**

*Data Scientist*

*Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina*

- 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.
- Implementar una feature de entrenado con máscara para nuestros modelos PyTorch de CV, reduciendo los tiempos que le toman al equipo de Data Science agregar una nueva condición de moderación de múltiples semanas a 0.
- Participación en la Hackathon interna de 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, PyTorch, Python, modelos & BI tools in-house.

### Intellignos, Havas

**Mar. 2022 – Dic. 2022 (10 meses)**

*Data Scientist*

*Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina*

- Mejorar la eficiencia y precisión de las métricas de evaluación para un modelo de atribución, convirtiendo una costosa métrica de  $O(ne^n)$  a una de  $O(n)$  con  $n$  el número de features consideradas en la métrica.
- Optimizado de la ETL, el soporte, la documentación, la automatización y el diseño de archivos de configuración al fusionar los ETL de dos clientes en un solo sistema, reduciendo el tamaño del equipo y el tiempo de soporte en un 50%.
- Replica local vía Matplotlib del dashboard en producción de los clientes reduciendo a la mitad el tiempo de QA.
- 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 de microcenso identificando las 5 características más relevantes para la estimación demográfica.
- Clasificación binaria del estado de ocupación en un grupo de viviendas con aproximadamente 100,000 residentes.
- 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 300k+ 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: ETL, data mining de datos de simulación con scripts Bash y Python.
- Software: Data Science en Python. Experiencia en proyectos en salud, física y biología. CUDA orientado a simulaciones. Estructura de datos en C; árboles, pilas, colas.
- 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.
- 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, PostgreSQL, BigQuery, Dataproc, Streamsets, PySpark, REST APIS

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

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

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