

EVARISTO VILLASECO

evaristo.va@outlook.com | +1 (929) 385-8568 | New York City, NY | [LinkedIn](#) | [Portfolio](#) | [Github](#) | [Google Scholar](#)

EDUCATION

Ph.D. in Physics, *Rutgers University* | 2020–2025 | GPA 4.0.

M.Sc. in Theoretical Chemistry and Computational Modeling, *Autonomous University of Madrid* | 2016-2018 | 1 A+ Class Honor.

B.Sc. in Chemistry, *University of Salamanca* | 2012–2016 | 8 A+ Class Honors.

SKILLS & CERTIFICATIONS

Languages, Tools & Compute: Python, SQL, Bash, Git, Flask, PySpark, HPC (Slurm, GPC-like clusters), AWS (working knowledge).

Analytics & Modeling: A/B testing, Time Series Forecasting, Tree-based Models, NLP, and GenAI, Statistics, Probability.

ML & AI: PyTorch, Scikit-learn, Statsmodels, XGBoost, Hugging Face, BeeAI (AI agentic framework).

Certificates: ColumbiaX [Corporate Finance](#) (2022), Erdős Institute [Data Science](#) & [Deep Learning](#) Boot Camps (2024).

EXPERIENCE & PROJECTS

AI Engineer | *Burnt (Startup Internship)* | 2025–Present

- Developed a GenAI-powered assistant to automate order processing for sales representatives, integrating with WhatsApp for seamless communication. Enabled real-time data validation through RAG and vector search.
 - Designed AI workflows to streamline data processing and enhance automation capabilities.
 - Collaborated cross-functionally with sales and product teams to align AI features with customer feedback and investor priorities.
- Tools: Python, SQL, SQLAlchemy, BeeAI, LangChain, Flask, Mem0, Twilio.

Research Scientist | *Rutgers University* | 2020–Present

- Developed a novel algorithm for molecular simulations that improved numerical stability and energy conservation by 87% over standard methods. Automated large-scale simulation and data processing pipelines for 10,000+ high-dimensional trajectories across 7+ molecular systems, enabling statistical analysis and benchmarking against baseline models. Integrated core contributions into open-source softwares. Results published in: J. Chem. Phys. 158 (2023) and J. Chem. Theory Comput. 19 (2023).
 - Derived a physics-informed framework and forecasting model for time-dependent molecular properties, boosting predictive accuracy by 90% over traditional modeling approaches. Designed numerical experiments to isolate the key features driving the system behavior, improving signal interpretability and exposing critical limitations in conventional approaches. Results published in: Phys. Rev. Lett. 133, 233201 (2024).
 - Led an international research collaboration, funded by the French Embassy in the US, to develop novel molecular dynamics simulation methods, resulting in 5 of 10+ peer-reviewed [publications](#) in high-impact journals and multiple presentations at international conferences.
- Tools: Python, Fortran, Bash, Git, HPC (Slurm).

Developer | *g-ctmqc & QuantumModelLib* | 2022–2025

- Contributed to open-source tools for molecular dynamics, expanding [g-ctmqc](#) simulation capabilities by 25% with new simulation schemes and adding two quantum models to [QuantumModelLib](#), increasing modeling options by 5%.
- Tools: Python, Fortran, Bash, Git.

Machine Learning Scientist | *Burnt (Startup Internship)* | 2024

- Developed a demand forecasting model for restaurant sales and menu items, projecting \$250k/year savings on food costs. Performed extensive analytics uncovering key sales trends and consumer behavior to optimize inventory and menu decisions.
 - Fine-tuned LLMs (LLaMA2, BERT, GPT-2) via QLoRA, improving product categorization across food suppliers with ~92.5% accuracy.
- Tools: Python, PyTorch, Hugging Face, Statsmodels, Scikit-learn, Pandas, Seaborn.

Data Scientist | *Independent Projects* | 2024-2025

- BiteBuddy:** Built an AI-powered nutrition assistant that provides personalized nutrition plans based on user input. Utilized prompt engineering techniques and integrated tool calling to deliver accurate, user-specific recommendations, enhancing engagement and interaction quality. Deployed on Streamlit Cloud for scalable, real-time access via a browser interface.
 - LaLiga SoccerSage:** Built a Random Forest model to predict LaLiga match outcomes, outperforming bookmaker probabilities by 24%. Integrated 10+ seasons of historical match data and engineered performance metrics for robust predictive model for sport analytics.
- Tools: Scikit-learn, XGBoost, Pandas, Seaborn, Matplotlib.