# Juan Lara

## Computer Scientist & Mathematician

✉ larajuand@outlook.com    🌐 Web Page    📍 Bogotá, Colombia
📞 +57 315 512 8464    in julara    ⌗ JuanLara18

## Professional Summary

Interdisciplinary Machine Learning Specialist with a solid foundation in Computer Science and Mathematics, complemented by applied research experience at Harvard Business School. Applies theoretical insights to design scalable predictive models, integrate state-of-the-art generative architectures, and optimize end-to-end data pipelines. Transforms complex datasets into actionable strategic insights, driving impactful organizational decision-making.

## Professional Experience

**Research Assistant**                                    **Sep 2022 - Present**
*Harvard Business School*                                *Boston, USA (Remote)*

- **Integrated advanced ML pipelines** into research workflows, applying clustering algorithms, XGBoost predictive models, and NLP techniques to analyze large-scale organizational datasets, revealing key insights on firm learning strategies.
- **Generate data-driven insights** to explain firm behavior under different training programs and technology shocks, translating quantitative research into actionable recommendations for reskilling strategy.
- **Design mathematical models** of organizational hierarchies, contract dynamics, and technology shocks—providing formal proofs and simulation-based validation that underpin an upcoming academic publication.

**Supervisor:** Jorge Tamayo

**Data Scientist**                                        **Feb 2024 - Jan 2025**
*Ipsos*                                        *Bogota, D.C., Colombia (Hybrid)*

- **Engineered scalable segmentation applications** using ML models and robust data-processing pipelines on Google Cloud Platform, enhancing operational efficiency and accuracy across multiple data sources.
- **Streamlined analytical workflows** through automated Python pipelines, significantly reducing manual processing efforts, enabling dynamic real-time reporting, and achieving recognition as the Total Ops Star Employee in LATAM.
- **Led development of TextInsight**, a Python-based NLP library leveraging generative AI and advanced network analysis (NetworkX), cutting text-analysis processing time from several hours to less than one hour, enhancing decision-making across multiple business units in Latin America.

**Supervisor:** Sandra Pastrán

## Education

**B.S. in Computer Science**

*Universidad Nacional de Colombia*, Bogotá D.C.

Emphasis on Machine Learning

Director: Omar Duque Gomez

GPA: 4.7/5.0

Detailed List of Exams

**B.S. in Mathematics**

*Universidad Nacional de Colombia*, Bogotá D.C.

Emphasis on Applied Mathematics

Director: Omar Duque Gomez

GPA: 4.7/5.0

Detailed List of Exams

## Technical Skills

### Core Competencies

*Areas of technical expertise where I have the most depth and experience*

**Machine Learning Systems**

*End-to-end design and deployment of scalable machine learning models, including NLP pipelines and integration of LLM. Experienced in building adaptive generative architectures with high interpretability and real-time capabilities for business applications.*

**ML App Development**

*Expertise in creating responsive, interactive web applications and dashboards using Python frameworks (Streamlit, FastAPI). Strong background in designing full-stack architectures, building robust APIs, configuring CI/CD pipelines, and developing reusable components.*

**Mathematical Reasoning**

*Advanced analytical and theoretical modeling skills, specializing in abstraction, pattern generalization, and decomposition of complex organizational and strategic problems to develop scalable solutions.*

**Applied Data Science**

*Proficient in statistical analysis, predictive modeling, and data visualization. Skilled in data extraction, result interpretation, statistical validation, and automated metrics deployment to support data-driven decision-making.*

### Technical Proficiency

*Languages, frameworks, and tools with proven experience and impact*

**ML & AI Frameworks**

| | | |
|---|---|---|
| Generative AI | | *OpenAI GPT, LlaMa, Gemma* |
| NLP | | *Huggin Face Transformers, NLTK, SpaCy v3+* |
| ML Frameworks | | *Scikit-learn, TensorFlow, PyTorch, XGBoost* |

### Automation & Agentic AI

Agent Frameworks — *n8n, Langchain*

Cloud — *GCP-Vertex, AWS-SageMaker*

### DevOps & Tools

Version Control — *Git, GitHub*

Deployment — *Docker, MLflow*

CI/CD — *GitHub Actions*

### Visualization & Web Dev

Data Visualization — *Matplotlib, Plotly, NetworkX*

Interactive Applications — *Streamlit, FastAPI*

Web Development — *React, Tailwind CSS*

## Additional Training

**Bases de datos SQL** — **April 2025**
*Platzi* — *Certificate*
**Credential ID:** 53984442-3b5e-43e9-ae00-d68331327f26

**AI Agents Fundamentals** — **February 2025**
*Hugging Face* — *Certificate*
**Credential ID:** juanlara

**Artificial Intelligence Expert Certificate (CAIEC)** — **November 2024**
*Certiprof* — *Certificate*
**Credential ID:** TLZVDQTVTGG-XWHHHQPTQ-RDJFLDLRK

**Artificial Intelligence Bootcamp** — **May-October 2024**
*Talento Tech Cymetria* — *159 hours*
**Credential ID:** 2518458921

**DevOps Certification** — **October 2024**
*Platzi* — *Comprehensive Program*
**Credential ID:** cc4cfe8a-d78a-4883-8a75-ca90931151f6

**Algorithmic Toolbox** — **2023**
*Coursera* — *Advanced Course*
**Credential ID:** 8GR62BCT499V

**Linux and Bash for Data Engineering** — **2023**
*Coursera* — *Technical Course*
**Credential ID:** CAZUJPW6D4BP

## Distinctions & Awards

**Total Ops Star Employee - LATAM**                                **April 2024**
*Ipsos*

Recognized for developing TextInsight, demonstrating exceptional initiative, technical expertise, and commitment to operational excellence across Latin America operations.

**Best Averages Scholarship**                                       **2018-2023**
*Universidad Nacional de Colombia*

Awarded for 10 consecutive semesters to the top 15 students with highest academic performance in the program, maintaining excellence throughout entire academic career.

## Research Interests

🔍 **Continual Learning & Lifelong AI**
*Studying methods that allow models to learn continuously from streaming data without catastrophic forgetting, with applications in dynamic environments and evolving tasks.*

🔍 **Explainable AI (XAI)**
*Developing interpretable machine learning models and post-hoc explanation techniques to improve transparency, accountability, and trust in AI systems.*

🔍 **AI for Systems Optimization**
*Applying ML and optimization techniques to improve system-level performance in scheduling, resource allocation, and operations research.*

## Languages

**Spanish**          Native
**English**          Advanced