

# Juan Lara

## Computer Scientist & Applied Mathematician

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

## Professional Summary

Computer Scientist and Mathematician trained at the Universidad Nacional de Colombia, with specialized expertise in Machine Learning. Currently collaborating on applied research at Harvard Business School, where I merge theoretical rigor with computational solutions to study the structure of firms. Throughout my career, I have developed advanced mathematical models, implemented numerical simulations, and applied machine learning techniques to transform data into strategic insights. My interdisciplinary approach enables me to turn complex theories into practical and efficient tools that drive real-world impact.

Proactive, creative, and problem-solving-oriented, I excel at collaborating and communicating complex ideas clearly. I am particularly interested in research projects and consulting engagements that integrate mathematical foundations with computational solutions to address challenging organizational and strategic problems.

## Professional Experience

### Research Assistant Harvard Business School

Sep 2022 - Present  
Boston, USA (Remote)

Conduct advanced research in organizational economics with Professor Jorge Tamayo, focusing on two key areas:

- **Theoretical Modeling** – Formalized organizational hierarchies and resource optimization through economic frameworks based on established literature (Garicano, Caliendo & Rossi-Hansberg), creating rigorous mathematical proofs to analyze policy impacts on productivity.
- **Computational Implementation** – ~~Translated theoretical concepts into actionable insights~~ by developing ML models (XGBoost, clustering), implementing NLP pipelines for technical document analysis, and building modular data architectures that integrate statistical analysis with visualization techniques (UMAP, SVD).

Supervisor: Jorge Tamayo

### Data Scientist Ipsos

Feb 2024 - Jan 2025  
Bogota, D.C., Colombia (Hybrid)

Transformed market research operations through strategic data science implementations:

- **Field Operations Innovation** – Developed a mobile R Shiny application with embedded machine learning models for pharmacy segmentation, integrating geospatial mapping (Leaflet) with cloud storage, resulting in 85% reduction in manual classification work.
- **NLP-Powered Analytics** – Created *TextInsight*, a comprehensive Python library using transformer models (BERT) and network visualization (NetworkX) for multilingual survey analysis, reducing processing time by 60% while delivering more nuanced insights.
- **Interactive Reporting** – Designed dynamic dashboards and visualization tools that streamlined decision-making processes across Latin American operations, earning recognition as Total Ops Star Employee for LATAM.

Supervisor: Sandra Pastrán

## Education

### B.S. in Computer Science

Universidad Nacional de Colombia, Bogotá D.C.

Emphasis on Machine Learning

Director: Omar Duque Gomez

Feb 2019 - Nov 2023

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

Feb 2018 - Jun 2022

GPA: 4.7/5.0

Detailed List of Exams

### Technical Baccalaureate in Business Administration

Centro Educativo los Andes, Bogotá D.C.

Emphasis on Entrepreneurship and Investigation

Director: Cristian Santiesteban

Feb 2015 - Nov 2017

GPA: 4.5

Detailed List of Exams

### Technician in Maintenance of Computer Equipment

Servicio Nacional de Aprendizaje - SENA, Bogotá D.C.

Emphasis on Corrective Software

Director: Carlos Wilches

Nov 2015 - Dec 2016

GPA: 4.6/5.0

Detailed List of Exams

## Technical Skills

### Core Competencies

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



#### Machine Learning & AI

Design and implementation of ML models, NLP pipelines, and integration of large language models for practical business applications



#### Data Science

Implementing statistical analysis, predictive modeling, and advanced visualizations to transform complex data into actionable insights



#### Natural Language Processing

Building specialized NLP solutions with transformer architectures, semantic analysis, and text classification for multilingual contexts



#### Mathematical Modeling

Development of analytical frameworks and theoretical models for organizational economics and strategic decision-making



#### Interactive Applications

Creating responsive web applications and dashboards with R Shiny and Python frameworks for data-driven decision support



#### Research & Problem Solving

Bridging theoretical concepts with practical implementations to solve complex organizational and business challenges

### Core Technical Expertise

Key areas of knowledge and methodological approaches

Mathematical Foundations

Mathematical Modeling

Graph Theory

Probability

Optimization

Numerical Simulation

Machine Learning & AI

NLP Pipelines

LLM Integration

Predictive Modeling

Transformer Models

Embeddings

## Data Science

Interactive Visualization

Statistical Analysis

Text Mining

Dimensionality Reduction

## Application Development

Shiny

Web Development

Cloud Integration

API Implementation

Responsive Design

## Research & DevOps

Academic Research

Containerization

CI/CD Pipelines

Model Deployment

Technical Documentation

## Technical Proficiency

Languages, frameworks and tools where I have the most experience

### Programming Languages

Python



Advanced <sup>5</sup> 4+ years

R, SQL <sup>+</sup>



Intermediate-Advanced

Other Languages



Intermediate (C++, JavaScript)

### Data Science & ML

GenAI <sup>~</sup>



Advanced (OpenAI GPT, Llama, Claude)

Data Visualization



Advanced (Matplotlib, NetworkX, PyVis, Leaflet)

NLP <sup>+</sup>



Intermediate-Advanced (BERT, Transformers)

ML Frameworks <sup>←</sup>



Intermediate (Scikit-learn, TensorFlow & PyTorch)

### Web & Application Development

Interactive Apps



Advanced (Streamlit, R Shiny)

AI Frameworks



Advanced (Langchain, LlamaIndex, Gradio)

Web Development



Intermediate (React, FastAPI, Tailwind CSS)

### DevOps & Tools

Version Control



Advanced (Git, GitHub)

Deployment



Intermediate (Docker, MLflow)

Cloud Services



Basic-Intermediate (GCP, AWS, Azure)

CI/CD <sup>+</sup>

## Key Projects

### Cunservicios Platform

Full Stack Developer

2025

#### Technologies:

React

FastAPI

SQLAlchemy

Tailwind CSS

JWT

Formik

Developed a web platform for water and sewage utility services management with a modern microservices architecture. The system features user-facing modules for bill consultation and payment, service requests, PQR (petitions, complaints, claims) management, and administrative tools. Implemented responsive design, form validation, API integration, and database modeling using industry best practices. The platform streamlines customer service operations while providing a seamless user experience across devices.

## TextInsight - Python Library

2023-2024

Lead Developer

**Technologies:** Python Transformers OpenAI NetworkX PyVis

Built a multilingual text analysis library combining BERT for sentiment analysis, GPT-3.5 for text correction and topic generation, and embeddings for graph visualization. Features include batch processing for large datasets, flexible API, and interactive visualizations. Deployed at Ipsos to analyze survey data, cutting analysis time by 60% while delivering deeper insights.

## Pharmacy Segmentation Application

2024

Lead Developer

**Technologies:** R Shiny Google Cloud Storage Random Forest Reactive UI Leaflet

Developed a responsive mobile field application for pharmacy segmentation using R Shiny and integrated with Google Cloud Storage. The system combines survey data collection with real-time machine learning classification to categorize pharmacies into strategic segments (NUEVA FARMA, MEGA FARMA, TRADI FARMA, FARMA ESENCIAL). Features include geolocation mapping with Leaflet, route management for field agents, real-time data validation, and secure cloud integration. Implemented for Familia/Ipsos to optimize product distribution strategies, reducing manual segmentation effort by 85% while improving market targeting precision.

## Research Interests

### AI Agents & Multi-agent Systems

Exploring the development and optimization of autonomous AI agents and their interactions within multi-agent environments, with applications in organizational decision-making and strategic simulations.

### Computational Organizational Theory

Investigating mathematical models of organizational structure, knowledge flows, and decision hierarchies using computational simulations and empirical validation.

## 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.

## Additional Training

+ SQL

### AI Agents Fundamentals

Hugging Face

Credential ID: juanlara

February 2025

Certificate

### Artificial Intelligence Expert Certificate (CAIEC)

Certiprof

November 2024

Certificate

Credential ID: TLZVDQTVTGG-XWHHHQPTQ-RDJFLDLRK

**Artificial Intelligence Bootcamp**

Talento Tech Cymetria

Credential ID: 2518458921

May-October 2024

159 hours

**DevOps Certification**

Platzi

Credential ID: cc4cfe8a-d78a-4883-8a75-ca90931151f6

October 2024

Comprehensive Program

**Algorithmic Toolbox**

Coursera

Credential ID: 8GR62BCT499V

2023

Advanced Course

**Linux and Bash for Data Engineering**

Coursera

Credential ID: CAZUJPW6D4BP

2023

Technical Course

**Python and Pandas for Data Engineering**

Coursera

Credential ID: 72QS5JSBC67L

2023

Technical Course

**Languages**

Spanish  
English

Native  
Advanced