



# Leon Doungala

AI Engineer Experienced

Portfolio : [github.com/LeonDoungala22](https://github.com/LeonDoungala22) | <https://leondoungala22.github.io>

Viale Marconi, Tradate, Italy

+39 3293670086

doungala.leon@gmail.com

[linkedin.com/in/leon-doungala](https://www.linkedin.com/in/leon-doungala)

## SKILLS

- Generative AI & LLMs
- Retrieval-Augmented Generation (RAG)
- Chat Search Systems & Autonomous Agents IA Integration
- Machine Learning & Deep Learning
- Python & Predictive Modeling & Data Science
- Research & Development

## EDUCATION

- Rome Business School, Rome, Italy** - **University of Valencia, Valencia, Spain** Dec 2022  
Double Master's in Data Science
- eCampus University, Come, Italy** Apr 2024  
Master of Science in Computer Engineering and Automation, Artificial Intelligence
- Ferrara University, Ferrara, Italy** Sept 2020  
Bachelor's of Science in Electronic and Computer Engineering

## PROFESSIONAL EXPERIENCE

### AI Engineer / Data Scientist , Experienced Consultant

Jun 2025 - Present

Amaris Consulting , Milan, Italy

- Developed and collaborated with teams on an AI agent assistant for smart devices and EV charging support for client support service using **LLM, RAG ( Langchain )** .
- Processed **64M+** telemetry sessions applying clustering, **Isolation Forest**, and **anomaly detection** techniques.
- Built advanced AI/ML pipelines and an RAG agent in R&D to enhance predictive analytics and insights.

### AI / ML and Data Science Engineer

Mar 2024 – Jun 2025

People Lab , Milan, Italy

- Developed AI solutions for **300+ clinics** and companies, including a RAG-based Clinical Semantic Search Engine and an **ATS** system for HR recruiters supporting **200+ companies**.
- Engineered scalable **multimodal AI systems for semantic search**, autonomous agents, and workflow automation, driving innovation in healthcare and industry.
- Designed AI projects with **Generative AI, RAG, Transfer Learning, and LLMs**, building predictive systems and scalable solutions using ML, DL, MongoDB, MySQL, Azure Cosmos, Python, and Vector Databases, **Langchain , Flask , Fast api..**

### Software Developer | Data Scientist

Apr 2022 – Dec 2023

Italian Stock Exchange (Euronext Group) , Milan, Italy

- Developed and maintained software applications, optimized SQL Data Marts (Workbench, Athena), and automated processes with Python.
- Collaborated within agile teams, managed AWS data ingestion (Glue, Python Spark), and enhanced machine learning models.
- Applied data science techniques to support decision-making and drive actionable insights.

### Developer Analyst Salesforce

Jun 2020 – Dec 2021

Lutech group , Milan, Italy

- Salesforce: Skilled in development, automation, deployment, error analysis, Apex, VisualForce, and data management.

### HCM Technical Consultant - Intern

Sep 2019 – Feb 2020

Oracle , Milan, Italy

- Gained hands-on experience in Oracle SQL, Fusion Reporting, BI, integrations, HCM cloud (BI Publisher, OTBI, HCM extracts), web services, and Taleo Client

## LANGUAGES

- French - Native ( C2 level)
- Italian - Proficiency ( C1 level)
- English - Intermediate ( B2 level)

## TECHNICAL PROFICIENCIES

- Programming Languages:** Python, Java, SQL, Shell
- Databases:** MongoDB, MySQL, Postgresql
- Libraries & Frameworks:** Scikit-learn ,TensorFlow, LangChain, Pandas, Flask
- Cloud :** AWS, Azure , Salesforce , Elasticsearch
- APIs & AI Models:** OpenAI API , Anthropic API , LLM integrations , REST API's
- Tools:** Git, Docker, Tableau, Jupyter Notebook, GitHub
- Web Development:** HTML, CSS, JavaScript, Flask
- AI Expertise:** LLMs, AI Agents , RAG, Predictive modeling

## Highlighted Projects

**OneclickMedical** : Multimodal Virtual Doctor AI Assistant API , Swagger , Documentation and web app .

- **Description**: The AI Health Assistant is an AI system designed to assist individuals by providing helpful, reassuring, and “medically” sound guidance (screening) based on their symptoms , appointments by ai assistant .
- **Key Features**: Dynamic medical data retrieval using RAG langchain , Saas API oriented , conversational flows, and task automation.
- **API's** : [Swagger](#) , **Web app (test)** : <https://virtualdoctoronclickai.onrender.com/>

## **Advanced Human Resources Analytics**

- **Description**: Applied PCA and K-Means clustering for employee segmentation and HR data analytics.
- **Algorithm**: PCA, K-Means
- **Repository Link**: [View Project](#)

## **M.Sc. Thesis** : Artificial Intelligence and Machine Learning Techniques for the Characterization and Prediction of Diabetes

- **Focus**:  
Advanced AI and ML techniques, including **RAG**, **semantic search**, and **generative AI**, to address diabetes prediction and characterization challenges. Developed supervised (e.g., XGBoost, Random Forest) and unsupervised (e.g., K-Means, PCA) models for actionable medical insights.

### **Key Points**:

- Designed integrative dashboards and **predictive analytics** for healthcare.
- Developed advanced **predictive models** for diabetes risk and characterization.
- Implemented LangChain and **Vector Databases** for **RAG workflows** and **semantic search**.
- Leveraged **Generative AI** for data augmentation and handling imbalanced datasets.

**Repository (Private repo, access on request)** : [Explore Thesis Repository](#)

## **Exploring Customer Segmentation**

- **Description**: Segmented customer data using clustering techniques for market analysis.
- **Algorithm**: K-Means
- **Repository Link**: [View Project on GitHub](#)

## **Heart Disease Prediction**

- **Description**: Built predictive models to analyze heart disease risks using ML and NLP techniques.
- **Algorithm**: SVM, NLP
- **Repository Link**: [View Project on GitHub](#)

## **ATS System for CV Filtering and Job Matching**

- **Description**: Developed an AI-powered ATS system to align candidates with job descriptions or assist HR with advanced semantic search. Utilized RAG, LangChain, Anthropic APIs, OpenAI, and Azure Cosmos DB for precision matching and intelligent filtering.
- **Algorithm**: Retrieval-Augmented Generation (RAG) with semantic search and conversational AI.
- **Repository Link**: [View Project on GitHub](#)

## **Collaboration and Articles**:

- **Collaboration**: **Elyon – Advanced Diabetes Prediction Platform**  
*Collaborating with Professors **Pratizia Vizza** and **Pierangelo Veltri** (University of Calabria) on **Elyon**, an advanced **diabetes prediction platform** leveraging **Generative AI**, **Machine Learning**, and **Data Science**. Co-authoring a **scientific article** on the platform's methodology, results, and clinical applications.*
- **Publication**: **Advancing Credit Card Fraud Detection** : Author of “[Advancing Credit Card Fraud Detection through Machine Learning and Data Science Techniques](#)” – **Affidaty Blog** (2025). The article details a **reproducible ML pipeline** for fraud detection, from **data preprocessing** and **imbalance handling** to **model explainability**, with an [interactive notebook](#).

## ADDITIONAL INFORMATION

---

- **Driving License**: Category B (with personal vehicle)
- **Soft Skills**: Problem-solving, communication, adaptability, teamwork, proactivity
- **Interests**: AI for Healthcare, bass guitar, running
- **Recommendations**: Available upon request
- **Data Privacy**: I consent to the processing of my personal data included in this CV in compliance with the GDPR (EU Regulation 2016/679).