

# Lucas Mendes Barbosa

Data Scientist | NLP, LLMs, OCR & Cloud Solutions

☎ +55 14 99708-8943 | ✉ [lucas.mendestech@gmail.com](mailto:lucas.mendestech@gmail.com) | 🌐 [musicmoodai.com.br](https://musicmoodai.com.br)  
🔗 [linkedin.com/in/lucas-mendes-barbosa](https://linkedin.com/in/lucas-mendes-barbosa)

## Professional Summary

Data Scientist with 5+ years of experience in applied AI and analytics, specializing in LLMs, NLP, OCR, and AWS-based automation to deliver actionable insights.

## Professional Experience

### • CloudWalk, Inc.

**Operations Intelligence Analyst | Jan 2026 – Present | São Paulo, Brazil**

- Design and maintain analytical dashboards to monitor product performance, user journeys, and business health.
- Conduct deep-dive analyses across products and operations to identify inefficiencies and growth opportunities.
- Partner with product, risk, growth, and engineering teams to deliver actionable insights.
- Develop AI-powered assistants and automated reporting pipelines for daily metric monitoring.
- Apply experimentation, data enrichment, and modeling to support scalable, data-informed decisions.
- **Key Projects:**
  - ◆ Currently onboarding at CloudWalk as an Operations Intelligence Analyst, focusing on understanding product workflows, core KPIs, and the data ecosystem.

### • Lee, Brock, Camargo Advogados

**Mid-Level Data Scientist | Mar 2025 – Jan 2026 | São Paulo, Brazil**

- Led the design and deployment of AI-driven solutions for OCR, NER, and document classification using GenAI (Gemini, Claude, and ChatGPT), optimizing legal operations and research workflows.
- Built and maintained scalable serverless architectures on AWS (Lambda, SQS, SNS) and leveraged AWS Bedrock for document summarization and text generation, optimizing automated data processing and knowledge management.
- Implemented and enforced software quality standards through Clean Code, PEP 8 (Python), SOLID design principles, and CI/CD-integrated automated testing.
- Partnered with legal and business stakeholders to align AI initiatives with strategic firm objectives.
- **Key Projects:**

- ◆ Automated judicial expense documents, turning a fully manual process into an AI-powered pipeline with 98% accuracy (OCR + LLM).
- ◆ 1st place in an internal hackathon with a custom RAG solution for the legal sector.

## ● Liga Ventures

### Associate Data Analyst | Jan 2023 – Oct 2024 | São Paulo, Brazil

- Designed and automated scalable data pipelines using Airflow, PySpark, Pandas (Python), and Selenium to streamline data ingestion and transformation.
- Developed ML-based recommendation systems to enhance startup–corporate matchmaking and internal intelligence processes.
- Built interactive dashboards in Looker Studio and Power BI, delivering actionable insights for business and innovation teams.
- Promoted a data-driven culture through internal workshops and hands-on training sessions.
- **Key Projects:**
  - ◆ Delivered a data-driven innovation talk for Ford Latam, showcasing how advanced analytics can enhance customer understanding and engagement.

### Junior Data Analyst | Jan 2022 – Jan 2023 | São Paulo, Brazil

- Managed and maintained the company's data lake, ensuring structured and reliable data storage for analytics and machine learning pipelines.
- Designed and automated ETL workflows using Airflow, Pandas, and Python scripting, streamlining data ingestion, transformation, and recurring processes while reducing manual effort and processing time.
- **Key Projects:**
  - ◆ Maintained the CI/CD pipeline for the company's data lake project, ensuring reliable deployment and data flow automation.

### Data Intern | May 2021 – Jan 2022 | São Paulo, Brazil

- Contributed to ETL scripts using Python and SQL/NoSQL to support data integration and reporting processes.
- Assisted in the design and implementation of the company's data lake, ensuring structured data storage and accessibility.
- Helped develop metric dashboards to provide visibility and support decision-making.
- **Key Projects:**
  - ◆ Participated in the creation of the company's first data lake, establishing the foundation for scalable analytics.

## Education

### ● FIAP – Postgraduate in Machine Learning Engineering

May 2025 – Present

- MLOps: AWS SageMaker, CI/CD pipelines, experiment tracking, Docker, monitoring.
- GenAI: prompt engineering, fine-tuning, and deployment of generative AI systems.
- Machine & Deep Learning: CNNs, RNNs, BERT, GPT, GANs, computer vision, supervised/unsupervised learning, feature engineering, and model evaluation.
- **Key Projects:**
  - ◆ Recognized as the 2nd-place startup at FIAP's Tech For Change Hackathon for developing an AI-driven ecosystem to innovate the future of work and employability.

## ● **UNISAGRADO – Bachelor of Science in Computer Science**

**Jan 2021 – Dec 2024**

- Solid foundation in data structures, algorithms, OOP, networking, AI, and databases.
- **Key Projects:**
  - ◆ 1st place at the 26th Scientific and Technological Initiation Symposium (FATEC-SP) and the 31st Scientific, Technological, and Innovation Initiation Forum (UNISAGRADO) for research on AI applied to emotion recognition in Brazilian music.