

Luana Guedes Barros Martins

Research Scientist — Deep Reinforcement Learning

Goiânia, Brazil · luanagbmartins@gmail.com · luanagbmartins.github.io

SUMMARY

Research Scientist with 6+ years of experience at the intersection of Deep Reinforcement Learning, Robotics, and AI Alignment. Since July 2019, I have been leading R&D projects at the Center of Excellence in Artificial Intelligence (CEIA) and the Advanced Knowledge Center of Immersive Technologies (AKCIT) at the Universidade Federal de Goiás, bridging foundational research and production-grade deployments across finance, agribusiness, and tourism. My long-term research goal is to develop agents capable of general, adaptive intelligence through interaction and experience — with particular interest in RL for real-world systems, LLM alignment, and human-agent cooperation.

RESEARCH INTERESTS

- Deep Reinforcement Learning — Sequential decision-making, policy optimization, offline RL, real-world deployment
- Robotics — Learning-based control, sim-to-real transfer, embodied agents
- LLM Alignment — Reasoning, uncertainty quantification, safe and reliable language model behavior
- Recommender Systems — Fairness, personalization, and RL-based recommendation at scale

EDUCATION

M.Sc. in Computer Science · Universidade Federal de Goiás · April 2022 – May 2024

Advisor: Prof. Dr. Telma Woerle de Lima Soares · Goiânia, Brazil

Thesis: Balancing Relevance and Exposure Fairness in Personalized Recommendation Systems through Reinforcement Learning

Proposed A2Fair — an Actor-Critic RL framework that jointly optimizes recommendation relevance and group exposure fairness. The adaptive reward function weights relevance and fairness according to each user's affinity for diversity, enabling personalized and fairness-aware recommendations. Validated on public benchmarks against state-of-the-art methods.

Keywords: Recommender Systems · Reinforcement Learning · Exposure Fairness · Personalization

B.Sc. in Computer Science · Universidade Federal de Goiás · 2014 – 2020

Advisor: Prof. Dr. Telma Woerle de Lima Soares · Goiânia, Brazil

Thesis: Analysis of the Capability of Generalization of Reinforcement Learning Algorithms in the Context of Electronic Games

Evaluated the generalization limitations of PPO in the GVGAI benchmark — demonstrating overfitting tendencies in standard RL evaluation protocols and motivating held-out level evaluation as a more rigorous methodology.

Keywords: Reinforcement Learning · Neural Networks · Intelligent Agents · PPO · GVGAI

EXPERIENCE

Research Scientist · Centro de Excelência em Inteligência Artificial (CEIA) & Centro de Competência EMBRAPAII (AKCIT) · July 2019 – Present

Remote

Active participation in R&D projects at Brazil's Center of Excellence in AI, collaborating with industry partners to develop innovative AI solutions — from initial research through production deployment.

Research

- Deep Reinforcement Learning (Jul 2019 – present): Research on RL with focus on offline, multi-task, and transfer learning scenarios. Developed agents capable of generalizing across distinct tasks, learning robust representations that distinguish in-distribution data from out-of-distribution samples.
- Recommender Systems (Mar 2021 – May 2024): Research on recommendation systems for marketplaces, with focus on item/product exposure fairness. Developed algorithms ensuring fair and balanced recommendation distributions so all items receive adequate visibility.
- Reinforcement Learning from Human Feedback (Jan 2024 – present): Research integrating human evaluations into the LLM training process. Investigated the viability of using LLMs as judges in place of human annotators to reduce cost and improve scalability.

R&D Projects (Aug 2020 – Present)

- Debt Collection Personalization: Applied RL to orchestrate debtor contact strategies, optimizing timing, channel, and approach to maximize recovered value on outstanding debts.
- Agribusiness Marketplace Personalization: Built recommendation systems to personalize offers across agricultural products — land, livestock, job listings, and specialized services.
- Tourism Offer Personalization: Applied recommendation techniques to personalize travel offers including flights, hotels, restaurants, and tours.
- Intelligent Vehicle Inspection: Used recommendation systems to suggest parts in automotive repair budgets.
- Credit Recovery Algorithms: Applied RL to improve conversion rates on newly negotiated debts, learning adaptive follow-up strategies from individual payment behaviors.
- Agribusiness Marketplace Search & Recommendation Validation: Validated and customized search engine and recommendation components for an agribusiness marketplace platform.
- Large Language Model for the Energy Sector: Developed a large-scale NLP model applied to multiple solutions in the electric power industry.
- Autonomous Agents: Developing digital agents capable of planning, acting, cooperating, and learning in complex environments — exploring model-based RL, hierarchical planning, multi-agent intention inference, offline policy initialization, and meta-learning for task-conditioned policies.
- Digital Twins for Mobile Robotics: Building high-fidelity digital twins and RL pipelines for humanoids and quadrupeds — covering sim-to-real transfer, motion capture retargeting, offline2online RL, multi-agent coordination, and policy decomposition for edge deployment.

Robotics Club — Pequi Mecânico · Universidade Federal de Goiás · *Undergraduate*

- Designed navigation and control pipelines for the IEEE SEK autonomous robotics team → 1st Place, LARC 2017 (Latin American Robotics Competition)
- Built navigation and control systems for the IEEE VSSS robot soccer team
- Organized community workshops introducing Arduino and robotics fundamentals

Technology Teacher · Labschool-BR · *February 2019 – November 2019 · Part-time*

Goiânia, Brazil

Taught technology classes to high school students, covering programming and robotics fundamentals. Curriculum included building Android apps with MIT App Inventor, creating web/mobile/desktop games with Stencyl, and preparing students for the Brazilian Robotics Olympiad (OBR).

PUBLICATIONS

[2026] Partial Reasoning in Language Models: Search and Refinement Guided by Uncertainty

M da Luz, B Brandão, L Martins, G Oliveira, B de Oliveira, L Melo, TWL Soares

arXiv preprint arXiv:2601.12040 · 1 citation

[2025] Learning Without Critics? Revisiting GRPO in Classical Reinforcement Learning Environments

BLM de Oliveira, FV Frujeri, MPCM Queiroz, LGB Martins, TWL Soares, L Melo

arXiv preprint arXiv:2511.03527 · 3 citations

[2025] InfoQuest: Evaluating Multi-Turn Dialogue Agents for Open-Ended Conversations with Hidden Context

BLM de Oliveira, LGB Martins, B Brandão, LC Melo

arXiv preprint arXiv:2502.12257 · 3 citations

[2025] Reinforcement Learning for Debt Pricing: A Case Study in Financial Services

B Brandão, LGB Martins, BLM de Oliveira, ML da Luz, E Garcia, LC Melo, et al.

RLC 2025 Workshop on Practical Insights into Reinforcement Learning for Real-World Systems

[2025] Personalizing Fairness: Adaptive RL with User Diversity Preference for Recommender Systems

LGB Martins, BLM de Oliveira, B Brandão, M Santana, TWL Soares

RLC 2025 Workshop on Practical Insights into Reinforcement Learning for Real-World Systems

[2024] Sliding Puzzles Gym: A Scalable Benchmark for State Representation in Visual Reinforcement Learning

BLM de Oliveira, LGB Martins, B Brandão, ML da Luz, TWL Soares, L Melo

arXiv preprint arXiv:2410.14038 · 6 citations

[2021] PulseRL: Enabling Offline Reinforcement Learning for Digital Marketing Systems via Conservative Q-Learning

L Melo, L Martins*, B Oliveira*, B Brandao*, DW Soares, T Lima (*equal contribution)*

2nd Offline RL Workshop @ NeurIPS 2021 · 3 citations

TECHNICAL SKILLS

Reinforcement Learning: PPO, Actor-Critic, GRPO, Conservative Q-Learning (CQL), Offline RL, Policy Optimization

Machine Learning: Deep Learning, Representation Learning, Neural Networks

NLP / LLM: Alignment, Reasoning under Uncertainty, Dialogue Evaluation

Recommender Systems: RL-based Recommendation, Exposure Fairness, Personalization

Robotics: Navigation, Control Systems, Sim-to-Real Transfer, Embedded Systems

LANGUAGES

- Portuguese — Native
- English — Professional proficiency
- Spanish — Intermediate