Caterina Motti

Milan, Italy caterina.motti@mail.polimi.it mttcrn.github.io motticaterina
mttcrn

Education

Politecnico di Milano *2023 – 2026*
MS in Computer Science Engineering - 108/110 (expected)

- Relevant AI courses: **Machine Learning**, Artificial Neural Networks & Deep Learning, Natural Language Processing, Game Theory, Uncertainty in AI, **Model Identification & Data Analysis**.
- Other relevant courses: Advanced Computer Architectures (**RISC-V**), Advanced Operating Systems, **Computer Security**, Offensive & Defensive Cybersecurity, **Advanced Research Topics in Cybersecurity**, Formal Languages and Compilers.
- **Athens Network Program**: Selected on academic merit (130 awardees university-wide) for intensive **Deep Learning** training at Telecom Paris.

Politecnico di Milano *2020 – 2023*
BS in Computer Science Engineering - 102/110

- Relevant courses: Fundamentals of C programming, Algorithms & Data Structures, Software Engineering, Computer Architectures (**MISP**) & Operating Systems.

Universidad de Alicante *2022*
Erasmus exchange in Computer Science Engineering

Research Experience

Master Thesis - Reinforcement Learning Researcher *Sept 2025 – present*
Politecnico di Milano - Advisor: Prof. Marcello Restelli

- Conducting research on unsupervised **Drift Detection** in Reinforcement Learning, supported by **MLcube** (Politecnico di Milano Spin-off).
- Designed a **Particle Filter** (Sequential Monte Carlo) to reconstruct **latent variables**, enabling the detection of hidden changes in environment dynamics without ground-truth labels.
- Evaluating the system's ability to identify concept drift in non-stationary environments without ground-truth labels.

Skills

- Languages: **Python**, **C**, Java, C++, MATLAB, VHDL, SQL
- AI Frameworks: **PyTorch**, **TensorFlow**, Keras, Gymnasium, Scikit-learn, Pandas, NumPy, LangChain
- Tools & Platforms: Git, UNIX/Linux, GDB, Valgrind, Ghidra
- Spoken Languages: English (C1, TOEIC: 960/990), Italian (Native)

Projects

Kaggle challenges: Deep Learning for Image Analysis. *view on GitHub*

- 1st challenge: Developed and fine-tuned **state-of-the-art models** for blood cell classification, achieving 91% accuracy using **transfer learning** and **fine-tuning** techniques.
- 2nd challenge: Designed a **custom U-Net** architecture from scratch to tackle **semantic segmentation** of Mars terrain, optimizing model performance without relying on pre-trained networks, reaching 66% mIoU on the final test set.
- Tools: Python, **Tensorflow**, **Keras**, Jupyter notebooks

NLP-project: Processing of RecipeNLG dataset *view on GitHub*

- Built a complete NLP pipeline on RecipeNLG, processing **2M+ recipes** and conducting comprehensive **EDA**.
- Developed a RAG-based recipe assistant using Mistral-7B as a backbone, implementing sparse retrieval algorithms (**BM25**, **DFRee**) to ground generation.
- Designed a **multi-agent workflow** using **LangGraph** for iterative recipe refinement and integrated **text-to-image** and **text-to-speech** models for multimodal output generation.
- Tools: Python, scikit-learn, Gensim, PyTorch, HuggingFace, LangGraph

CTF challenges: Offensive & Defensive cybersecurity. *view on GitHub*

- The challenges cover a wide range of topics on **binary exploitation** (shellcode, **reverse engineering**, mitigation bypass, symbolic execution and fuzzing, **return-oriented programming**, heap exploitation, race conditions, packing) and kernel exploitation.
- Performed low-level debugging of binary executables on Linux using GDB and Ghidra, analyzing stack frames, heap memory layout, and CPU registers
- Tools: GDB, pwndgb, pwntools, IDA, Ghidra, z3, angr, Python, C.

Distributed Multiplayer Game: Software Engineering project. *view on GitHub*

- Engineered a distributed system using Java RMI, Sockets and the MVC pattern, enabling a single server to manage scalable, concurrent matches for multiple clients.
- Implemented **comprehensive Unit Tests** to ensure high code coverage and system reliability.
- Coordinated a 4-person team using Git, adopting **Agile**-like practices such as **pair programming** and **code reviews**.
- Tools: **Java**, JavaFx, JUnit, JSON, Git

High-Performance Route Optimizer: Algorithms & Data Structures Project. *view on GitHub*

- Developed a highly efficient **BFS-based** (Breadth-First Search) algorithm for optimal route planning, optimizing performance under strict **memory** and **time constraints** (< 88 MiB RAM and 4s execution).
- Implemented custom binary search trees (BST) and queues from scratch using **manual memory management** (`malloc`/`free`) to minimize heap fragmentation.
- Tools: **C**, **GDB**, Valgrind.

FPGA-based Memory Access Controller: Logical Network Project *view on GitHub*

- Designed and implemented a hardware component in VHDL for FPGA-based memory access, allowing selective data retrieval and display without overwriting previous outputs.
- Tools: VHDL, Vivado.

Work Experience

Laboratory Tutor (Fundamentals of C Programming) *2024 – present*
Politecnico di Milano

- Assisted undergraduate students with programming exercises and debugging.

Peer-to-Peer Tutor *2022 – present*
Politecnico di Milano

- Mentored students in Mathematics, Physics, Logic, C programming and Machine Learning, helping improve study habits and exam preparation.