Matteo Papini

Postdoctoral researcher at UPF Barcelona

Education

2017–2021 **Ph.D., Information Technology**, *cum laude*, Politecnico di Milano, Milan, Italy.
Supervisor: Prof. Marcello Restelli
Dissertation: *Safe Policy Optimization*

2015–2017 **M.Sc., Computer Science and Engineering**, *Politecnico di Milano*, 110/110 cum laude.
Master Thesis: *Adaptive Batch Size for Safe Policy Gradient Methods*

2012–2015 **B.Sc., Computer Engineering**, *Politecnico di Milano*, 110/110 cum laude.

Research Interests

My main research interest is **Reinforcement Learning**. I study Reinforcement Learning and Contextual Bandit algorithms from a theoretical perspective to make them more applicable to real-world problems.

Research Experience

2021–present **Postdoctoral researcher**, *Universitat Pompeu Fabra*, Barcelona.

Autumn 2020 **Research Intern**, *Facebook AI Research*.

2017–2021 **Research Assistant**, *Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano*, Milan, Italy.
Research Topic: Reinforcement Learning for Industry 4.0.
Including industrial projects with PIRELLI TYRE S.P.A. *and* Scuderia Ferrari.

Publications (Conference and Journal Papers)

[1] **M. Papini**, A. Tirinzoni, M. Restelli, A. Lazaric, and M. Pirotta. Leveraging good representations in linear contextual bandits. In **ICML**, volume 139 of *Proceedings of Machine Learning Research*, pages 8371–8380. PMLR, **2021**.

[2] A. M. Metelli, **M. Papini**, P. D'Oro, and M. Restelli. Policy optimization as online learning with mediator feedback. **AAAI**, to appear, **2021**.

[3] **M. Papini**, A. Battistello, and M. Restelli. Balancing learning speed and stability in policy gradient via adaptive exploration. In **AISTATS**, volume 108 of *Proceedings of Machine Learning Research*, pages 1188–1199. PMLR, **2020**.

[4] A. M. Metelli, **M. Papini**, N. Montali, and M. Restelli. Importance sampling techniques for policy optimization. *Journal of Machine Learning Research* **(JMLR)**, 21(141):1–75, **2020**.

[5] P. D'Oro, A. M. Metelli, A. Tirinzoni, **M. Papini**, and M. Restelli. Gradient-aware model-based policy search. In **AAAI**, pages 3801–3808. AAAI Press, **2020**.

[6] L. Bisi, L. Sabbioni, E. Vittori, **M. Papini**, and M. Restelli. Risk-averse trust region optimization for reward-volatility reduction. In **IJCAI**, pages 4583–4589. ijcai.org, **2020**.

[7] **M. Papini**, A. M. Metelli, L. Lupo, and M. Restelli. Optimistic policy optimization via multiple importance sampling. In **ICML**, volume 97 of *Proceedings of Machine Learning Research*, pages 4989–4999. PMLR, **2019**.

[8] M. Beraha, A. M. Metelli, **M. Papini**, A. Tirinzoni, and M. Restelli. Feature selection via mutual information: New theoretical insights. In **IJCNN**, pages 1–9. IEEE, **2019**.

[9] **M. Papini**, D. Binaghi, G. Canonaco, M. Pirotta, and M. Restelli. Stochastic variance-reduced policy gradient. In **ICML**, volume 80 of *Proceedings of Machine Learning Research*, pages 4023–4032. PMLR, **2018**.

[10] A. M. Metelli, **M. Papini**, F. Faccio, and M. Restelli. Policy optimization via importance sampling. In **NeurIPS**, pages 5447–5459, **2018**.

[11] **M. Papini**, M. Pirotta, and M. Restelli. Adaptive batch size for safe policy gradients. In **NeurIPS**, pages 3591–3600, **2017**.

Preprints

[P2] **M. Papini**, M. Pirotta, M. Restelli. Smoothing Policies and Safe Policy Gradients. **2019**.

Workshop Papers

[W1] **M. Papini**, Andrea Tirinzoni, Aldo Pacchiano, Marcello Restelli, Alessandro Lazaric, and Matteo Pirotta. Reinforcement Learning in Linear MDPs: Constant Regret and Representation Selection. ICML-2021 Workshop on Reinforcement Learning Theory, virtual, **2021**.

[W2] **M. Papini**, A. Battistello, and M. Restelli. Safe Exploration in Gaussian Policy Gradient. NeurIPS-2019 Workshop on Safety and Robustness in Decision Making, Vancouver, Canada, **2019**.

[W3] **M. Papini**, A. Battistello, and M. Restelli. Safely Exploring Policy Gradient. In 14th European Workshop on Reinforcement Learning, Lille, France, **2018**.

Schools and Exchange Programs

Jul–Aug 2018 **Deep Learning and Reinforcement Learning Summer School**, Toronto, Canada.

Oct 2017 **ACAI Summer School on Reinforcement Learning**, Nieuwpoort, Belgium.

Autumn 2016 **Erasmus Programme**, KTH Royal Institute of Technology, Stockholm, Sweden.

Invited Talks

July 1, 2021 **Leveraging Good Representations in Linear Contextual Bandits and MDPs**, *Mathematical Statistics and Learning*, Barcelona.

November 2020 **Safe Policy Optimization**, *for the research seminar organized by Prof. Gerhard Neumann (Karlsruhe Institute of Technology)*, delivered online.

September 2019 **Optimistic Policy Optimization via Multiple Importance Sampling**, *Workshop on Markets, Algorithms, Prediction and LEarning (MAPLE 2019)*, Politecnico di Milano, Milano, Italy.

Teaching Experience

2020, 2021 **Teaching Assistant**, *Politecnico di Milano*.
Artificial Intelligence (online class), Prof. Andrea Bonarini

2018, 2019 **Teaching Assistant**, *Politecnico di Milano*.
Informatica B (introductory computer science class), Prof. Luca Cassano

Jul 2019 **Teaching Assistant**, *Reinforcement Learning Summer School, Lille, France*, organized by the SCOOL (formerly SequeL) team.

2018 **Teaching Assistant**, *Politecnico di Milano*.
Web and Internet Economics, Prof. Nicola Gatti

2017 **Lab Assistant**, *Politecnico di Milano*.
Informatica B (introductory computer science class), Prof. Luca Cassano

2016 **Lab Tutor**, *Politecnico di Milano*.
Prova Finale–Ingegneria del Software (software engineering: final project), Prof. Carlo Ghezzi

Editorial Activities

2021 Reviewer for NeurIPS, **expert reviewer** for ICML, reviewer for AISTATS, emergency reviewer for AAAI.

2020 **Top 33% reviewer** for ICML, reviewer for NeurIPS, AISTATS, AAAI, UAI, ECAI.

2019 Reviewer for ICML, NeurIPS, UAI.

Mentoring

2017–2020 **Master Students**, *I have co-supervised 9 master theses at Politecnico di Milano*.

Languages

Italian, *mother tongue*.

English, *fluent (FCE, Grade A — C1 CEFR level)*.