

Diego Antognini, PhD

RESEARCH SCIENTIST IN MACHINE LEARNING & NATURAL LANGUAGE PROCESSING

✉ diegoantognini@gmail.com | 💻 www.diegoantognini.com

🏠 Diego Antognini | 🌐 Diego999 2.8k* | 🌐 diegoantognini | 🇨🇭 Swiss 🇪🇸 Spanish 🇮🇹 citizen



Summary

6+ years of research experience in natural language processing (NLP), machine learning (ML), and multi-objective recommendation systems. Currently focusing on aligning large language models (LLMs) with user preferences and developing efficient models for resource-constrained training and inference settings. Experienced in designing interpretable models that generate personalized and actionable textual explanations. Supervised 70+ B./M.Sc. projects.

Experience

Research Scientist

May 2022 - present

- Working on efficient machine learning for NLP: alternatives to transformer-based language models, cost-effective pre-training methods via self-supervision, or better leveraging data with active learning. I build models with similar performance and higher throughput than large models.
- Designing methods to adapt and personalize LLMs to users, using parameter-efficient fine-tuning methods such as prompt-tuning and adapters.
- Deployed models in the order of 1MB size and 1ms latency used in *IBM Deep Search* that analyzes in real time scientific publications and patents.
- Produced 4 patents that are filed, 4 papers published in AI & ML leading venues, and 1 paper under review at NeurIPS.

IBM Research AI

Zürich, Switzerland

Lecturer, Module head, and Supervisor for M.Sc. theses in NLP

Apr. 2022 - present

- Designing and teaching the course of computational language technologies and deep learning for NLP to 50 M.Sc. students.
- Supervised 10 M.Sc. theses in NLP with companies in the areas of medicine, law, insurances, banks, media, and data visualization.

HSLU - Lucerne University of Applied Sciences

Luzern, Switzerland

Consultant and Expert for B.Sc. and M.Eng. theses

June 2015 - present

- Gave talks on multiple technologies and offered consulting services for applied research in industrial projects in Prof. Hatem Ghorbel's group.
- Assessed 30+ B.Sc./M.Eng. theses and was an expert for the AI and GPU/parallel programming exams (Prof. Hatem Ghorbel, Prof. Cédric Bilat).

HES-SO - HE-ARC – University of Applied Sciences

Neuchâtel, Switzerland

Visiting Researcher in Prof. Julian McAuley's group

Jul. 2021 - Oct. 2021

- Created an unsupervised critiquing method and a generative language model to help users rewrite cooking recipes to satisfy dietary restrictions.

UCSD – University of California San Diego

San Diego, CA, U.S.A.

Research and Teaching Assistant

May 2017 - Mar. 2022

- Helped in teaching the courses: intelligent agents (M.Sc.), introduction to natural language processing (M.Sc., 2x), artificial intelligence (B.Sc., 3x).
- Supervised 30+ B./M.Sc. semester projects & theses. Worked with the data analytics & AI research team in Swisscom (led by Dr. Claudiu Musat).

EPFL – École Polytechnique Fédérale de Lausanne

Lausanne, Switzerland

Machine Learning and Data Mining M.Sc. thesis Intern

Sep. 2016 - Mars. 2017

- Deployed a system to extract concepts from large corpora and built interactive knowledge graphs for invention developers to gain new insights.

Iprova GmbH

EPFL Innovation Park, Lausanne, Switzerland

Education

Ph.D. in Computer Science

May 2017 - Mar. 2022

- Publications:** 15 papers in AI & ML leading venues: 8 conference, 6 workshop, 1 demo. **Advisor:** Professor Boi Faltings, head of the AI laboratory.
- Award:** Won \$9,750 in the *IARPA Geopolitical Forecasting Challenge 2018*. Press coverage: *EPFL News (English)* and *24 Heures (French)*.
- Thesis** 📄: Textual Explanations and Critiques in Recommendation Systems. It solves two challenges: explanation generation and making them actionable (critiquing). It focuses on text understanding, explainable AI, and conversational recommendation. Fastest to graduate in the AI lab.
- Implemented the 1st PyTorch graph attention network and its sparse variant. Starred and forked on Github 3.1k+, and still 8k+ views per month.

EPFL – École Polytechnique Fédérale de Lausanne

Lausanne, Switzerland

M.Sc. in Computer Science

Sep. 2014 - Apr. 2017

- Focus on NLP, AI, ML, and distributed systems. It includes a mandatory extra year of 62 ECTS credits to be accepted in the program. **GPA:** 5.5/6.0.
- Thesis:** From Relation Extraction to Knowledge Graphs (6.0/6.0). Developed a model that extracts concepts from large corpora and classifies the semantic relationship between them. It outperformed state-of-the-art models by 0.9 F1-score in the relation-classification task of SemEval-2010.

EPFL – École Polytechnique Fédérale de Lausanne

Lausanne, Switzerland

B.Sc. in Computer Science

Sep. 2011 - Aug. 2014

- Major software engineering. **Awards:** 1) Excellent thesis and 2nd best GPA; 2) Deserving student in 2nd year. **GPA:** 5.6/6.0.
- Thesis:** Computing Brain Neuronal Maps. Developed a multi-GPUs algorithm to compute an accurate 3D real-time rendering of the brain's electromagnetic activities. A speedup of ~100,000x has been achieved, reducing the computation time from 20h to 700ms.

HES-SO - HE-ARC – University of Applied Sciences









Neuchâtel, Switzerland

Publications (selected)

Assistive Recipe Editing through Critiquing 📄

Diego Antognini, Shuyang Li, Boi Faltings, Julian McAuley

EACL 2023

pNLP-Mixer: an Efficient all-MLP Architecture for Language 	ACL 2023
Francesco Fusco, Damian Pascual, Peter Staar, Diego Antognini	
Extracting Text Representations for Terms and Phrases in Technical Domains 	ACL 2023
Francesco Fusco* and Diego Antognini * (equal contribution)	
Unsupervised Term Extraction for Highly Technical Domains 	EMNLP 2022
Francesco Fusco, Peter Staar, Diego Antognini	
Fast Critiquing with Self-Supervision for VAE-based Recommender Systems 	RecSys 2021
Diego Antognini and Boi Faltings	
Interacting with Explanations through Critiquing 	IJCAI 2021
Diego Antognini , Claudiu Musat, Boi Faltings	
Rationalization through Concepts 	ACL 2021
Diego Antognini and Boi Faltings	
Multi-Dimensional Explanation of Target Variables from Documents 	AAAI 2021
Diego Antognini , Claudiu Musat, Boi Faltings	
Addressing Fairness in Classification with a Model-Agnostic Multi-Objective Algorithm 	UAI 2021
Kirtan Padh, Diego Antognini , Emma L. Glaude, Boi Faltings, Claudiu Musat	

Talks (selected)

Efficient Machine Learning in Low-Resource and Highly-Specific Domains	2023
<ul style="list-style-type: none"> MIT-IBM Watson, Cambridge, MA, U.S.A. Host: Dr. Leonid Karlinsky. Swiss Text Analytics Conference 2023, Neuchâtel, Switzerland. Keynote. 	
Textual Explanations and Critiques in Recommendation Systems	2022
<ul style="list-style-type: none"> EPFL, Lausanne, Switzerland. 	
Interacting with Explanations through Critiquing	2021
<ul style="list-style-type: none"> University of Toronto, Online. Host: Pr. Scott Sanner. Swisscom AI, Lausanne, Switzerland. Host: Dr. Claudiu Musat. IJCAI 2021, Online. 	
Fast Critiquing with Self-Supervision for VAE-based Recommender Systems	
<ul style="list-style-type: none"> RecSys 2021, Online. 	
Transformers for Time-Series Predictions	
<ul style="list-style-type: none"> University of Applied Sciences, Neuchâtel, Switzerland. Host: Pr. Hatem Ghorbel. 	
T-RECS: a Recommender Generating Explanations and Integrating Critiquing	2020
<ul style="list-style-type: none"> ECAI 2020, Online. 	
Multi-Dimensional Explanation of Ratings from Reviews (Multi-Dimensional Rationalization)	
<ul style="list-style-type: none"> University of Zürich & NLP Meetup, Zürich, Switzerland. Host: Dr. Kornelia Papp. Swisscom AI, Lausanne, Switzerland. Host: Dr. Claudiu Musat. AAAI 2021, Online. 	
Learning to Create Sentence Semantic Relation Graphs for Multi-Document Summarization	2019
<ul style="list-style-type: none"> EMNLP 2019, Hong-Kong. 	
From Relation Extraction to Knowledge Graphs	2017
<ul style="list-style-type: none"> University of Applied Sciences, Neuchâtel, Switzerland. Host: Pr. Hatem Ghorbel. NLP course at EPFL, Lausanne, Switzerland. Host: Dr. Jean-Cédric Chappelier. NLP Meetup, Zürich, Switzerland. 	

Skills

Program Committee	NeurIPS, ICML, ACL, EMNLP, NAACL, EACL, SIGIR, RecSys. Journals : ACL Rolling Review, ACM Computing Surveys.
Software	Efficient : Python, PyTorch, Tensorflow, transformers, ONNX, Spark, Bash. Prior Experience : C++, CUDA, Java, SQL.
Miscellaneous	GNU/Linux, OSX, Git, Poetry, Docker, Kubernetes, Openshift, API design, PyCharm, Astah, Balsamiq Mockups, Eclipse.

Interests

Languages	French (native), English (fluent), German (conversational), Spanish (conversational), Italian (conversational).
Sport	I ride motorbikes, paddle on lakes, or surf on the coast. In-between, I go four times per week to the gym to stay fit.
Interests	I read self-development books and question myself to become a better person. I have traveled to 26 countries and 6 continents; I would like to keep going and discover new cultures. Finally, I love dancing salsa and bachata.

References

Dr. Cristiano Malossi Principal Research Manager IBM Research AI, Switzerland Previous manager ✉ ACM@zurich.ibm.com	Dr. Claudiu Musat Research Manager Google, Switzerland Close collaborator ✉ cmusat@google.com	Prof. Boi Faltings Full CS Professor EPFL, Switzerland Thesis advisor ✉ boi.faltings@epfl.ch	Prof. Julian McAuley Full CS Professor University of California, San Diego Research stay host ✉ jmcauley@ucsd.eng.edu
--	---	---	---