

# Francesco Salvi

MSC STUDENT · COMPUTATIONAL SCIENCE AND ENGINEERING

✉ salvi.fnc@gmail.com | 🏠 frasalvi.github.io | 📄 frasalvi | 📺 frasalvi | 🐦 fraslv

## Education

### EPFL

Lausanne, Switzerland

MSC IN COMPUTATIONAL SCIENCE AND ENGINEERING

Sept. 2021 - Feb. 2024

- GPA: 5.72/6. Keywords: **Computational Social Science, Natural Language Processing, Network Science** and Data Science.
- **Class delegate**, representing 40+ students to the Section, the Executive Board and the General Students Association.
- Student **Teaching Assistant** for "Applied Data Analysis" (Fall 2022) and "Machine Learning for behavioral data" (Spring 2022).

### University of Bologna

Bologna, Italy

BSC IN PHYSICS

Sept. 2018 - July 2021

- Grade: **110/110 cum laude**. Final GPA: **29.96/30**, top **0.25% of all UniBo students**.
- UniBo grant for students with outstanding academic records, awarded to **< 0.5% of students** from the whole university.

## Experience

### Fondazione Bruno Kessler

Trento, Italy

VISITING FELLOW

Sept. 2023 - Present

- Developing a randomized controlled experiment (**JavaScript, React**) to study the persuasiveness of LLMs in the spread of misinformation.

### EPFL | NLP Lab

Lausanne, Switzerland

RESEARCH ASSISTANT

June 2023 - Sept. 2023

- Training medical LLMs for doctors at scale via **continued pretraining** and **supervised instruction-tuning** of **LLaMA2-70B** and **Falcon-40B**.
- Adapted efficient model-parallel (tensor, sequence, and pipeline) and multi-node infrastructure for distributed training on a large cluster.

### Bain & Company | NPSx

Munich, Germany

ASSOCIATE INTERN, DATA SCIENCE

July 2022 - Sept. 2022

- Created **predictive models for NPS and CLV**, cornerstones of a product now sold to Fortune 500 companies.
- Productionised the data pipeline by integrating **MLOps** practices, testing and automation within **Google Cloud Platform**.

## Publications

- EPFL LLM Team. *MEDITRON-70B: Scaling Medical Pretraining for Large Language Models*. <https://arxiv.org/abs/2311.16079>
- Suresh A., **Salvi F.\***, Radojevic L.\*, Magron A, Kristof V., Grossglauser M. *Studying Lobby Influence in the European Parliament*. Under review at ICWSM24, <https://arxiv.org/abs/2309.11381>

## Projects

### Topical classification of Wikipedia images

RESEARCH PROJECT IN THE DATA SCIENCE LAB

Feb. 2022 - Present

- Developing a hierarchical classifier for Wikipedia images, categorizing them only by their topic in a weak supervision framework.
- Designed a set of **linguistically motivated heuristics** to associate images with training labels within a custom taxonomy.
- Year-long collaboration with [Wikimedia Research](#), soon to be wrapped up with a **first-author publication** (*in preparation*).

### A data-driven Political Compass

APPLIED DATA ANALYSIS COURSE PROJECT

Sept. 2021 - Dec. 2021

- Analyzed a dataset of ~180 mln quotations with **Apache Spark** within Python.
- Leveraged **Topic Extraction** and **Sentiment Analysis** to infer a new Political Compass and curated a [data story](#) on US politics.

## Extracurricular Activity

### Lead The Future

MENTEE

Sept. 2020 - Present

- Selected to be part of LeadTheFuture, a **leading mentorship organization** for Italian students in STEM, with acceptance rate < 20%.
- Part of the team reviewing applications for the 2023-24 cycle.

### Italian Association of Physics Students (AISF)

Bologna, Italy

PRESIDENT OF BOLOGNA LOCAL COMMITTEE

June 2019 - June 2020

- Elected **President of Bologna Local Committee** of AISF, an association gathering together more than 1400 Italian physics students.
- Organised conferences, scientific aperitifs, LaTeX courses, visits to research centres and social events with 50+ participants.