



## Giuseppe Colavito

📍 Via Maurizio Lettieri 62, 70024 Gravina in Puglia (BA)

📞 +39 350 547 8689

✉ giuseppecolavito999@gmail.com

LinkedIn <https://www.linkedin.com/in/peppocola/>

Born January 24th, 1999

---

### Work Experience

July 2023 – Sept 2023

#### Research Assistant

GARR - Università degli Studi di Bari Aldo Moro

Research Assistant at Collaborative Development Research Group:

- Deploying Text Categorization Models as Cloud Services
- Supervisors: Prof. Filippo Lanubile

March 2022 – May 2022

#### Master Curricular Internship

Università degli Studi di Bari Aldo Moro

Apprenticeship at Collaborative Development Research Group:

- Development, implementation and evaluation of a supervised classifier for issue label classification for automatic classification of Github issues. Use of state-of-the-art models for text classification, including pre-trained language models such as BERT.
- Supervisors: Prof. Filippo Lanubile, Prof. Nicole Novielli

October 2021 – July 2022

#### Reserch Assistant

Università degli Studi di Bari Aldo Moro

Research Assistant at Collaborative Development Research Group:

- Development of classifiers for emotion recognition through NLP
- Supervisors: Prof. Filippo Lanubile, Prof. Nicole Novielli

March 2020 – June 2020

#### Curricular Internship

Università degli Studi di Bari Aldo Moro

Apprenticeship at Collaborative Research Group:

- Development of a text classifier for documentation decluttering for Java source code. Use of NLP techniques and Supervised Learning.
- Supervisors: Prof. Nicole Novielli

---

### Education

November 2022 – Present

#### National PhD in Artificial Intelligence, AI for Society

Double affiliation, Università degli Studi di Pisa and Università degli Studi di Bari Aldo Moro

- Project: "Generative AI For Automatic Labeling in Software Engineering"
- Supervisors: Prof. Nicole Novielli, Prof. Filippo Lanubile

2020 – 2022

#### Master's Degree in Computer Science

Università degli Studi di BARI - Dipartimento di Informatica

- Master's degree (2 years)
- Computer Science - Artificial Intelligence
- Thesis title: Issue Report Classification Using BERT
- Mark: 110 cum laude
- Date: July 15th 2020
- Supervisors: Prof. Filippo Lanubile, Prof. Nicole Novielli

## 2017 – 2020 Bachelor Degree in Computer Science

Università degli Studi di BARI - Dipartimento di Informatica

- Bachelor's degree (3 years)
- Informatica
- Thesis title: Development of a text classifier for documentation decluttering
- Mark: 110 cum laude
- Date: July 20th 2022
- Supervisors: Prof. Nicole Novielli, Prof. Pierpaolo Basile

## Publications

---

- 2023 **Impact of Data Quality for Automatic Issue Classification Using Pre-trained Language Models**  
Colavito, G. and Lanubile, F. and Novielli, N., Journal of Systems and Software.
- 2023 **Few-Shot Learning for Issue Report Classification**  
Colavito, G. and Lanubile, F. and Novielli, N., Intl. Workshop on Natural Language-based Software Engineering (NLBSE 2023), co-located with ICSE 2023, Melbourne, Australia, May 20 2023.
- 2022 **Issue Report Classification Using Pre-trained Language Models**  
Colavito, G. and Lanubile, F. and Novielli, N., Intl. Workshop on Natural Language-based Software Engineering (NLBSE 2022), co-located with ICSE 2022, Pittsburgh, USA, May 8 2022. (DOI: 10.1145/3528588.3528659)
- 2022 **Lexicon Enriched Hybrid Hate Speech Detection with Human-Centered Explanations**  
M. Polignano, G. Colavito, C. Musto, M. de Gemmis, G. Semeraro. In: UMAP '22 Adjunct. Barcelona, Spain: Association for Computing Machinery, 2022, pp. 184–191. (DOI: 10.1145/3511047.3537688)
- 2020 **Leveraging Textual and Non-Textual Features for Documentation Decluttering**  
G. Colavito, P. Basile and N. Novielli, 2020 IEEE International Conference on Software Maintenance and Evolution (ICSME), 2020, pp. 862-863, doi: 10.1109/ICSME46990.2020.00113.

## Talks

---

### July 12th 2023 SATTose 2023

I presented a summary of my work on Issue Report Classification using Pre-Trained Language Models and Few-Shot Learning, during the 15th Seminar on Advanced Techniques & Tools for Software Evolution (SATTose)

### May 20th 2023 NLBSE 2023, co-located with ICSE 2023

I presented the "Few-Shot Learning for Issue Report Classification" paper during the International Workshop on Natural Language-based Software Engineering (NLBSE) 2023, co-located with ICSE 2023

### May 8th 2022 NLBSE 2022, co-located with ICSE 2022

I presented the "Issue Report Classification Using Pre-trained Language Models" paper during the International Workshop on Natural Language-based Software Engineering (NLBSE) 2022, co-located with ICSE 2022

### September 29th 2020 ICSME 2020, co-located with ICSE 2020

I presented the "Leveraging Textual and Non-Textual Features for Documentation Decluttering" paper during the DocGen2 workshop, at the IEEE International Conference on Software Maintenance and Evolution (ICSME) 2020, co-located with ICSE 2020

**Schools**

---

June 12th - 15th 2023

**16th International Summer School on Software Engineering (ISSSE)**

Organized by the University of Salerno.

June 5th - 9th 2023

**2nd AI & Society Summer School**

Organized by the National PhD Course in Artificial Intelligence.

April 3rd - 7th 2023

**9th International Spring School on Deep Learning (DeepLearn)**

Organized by the University of Bari, Universitat Rovira i Virgili, and IRDTA.

**Personal skills**

---

**Coding**

C, C++, Java, Haskell, Prolog, Python and scientific packages(Pandas, NumPy, sklearn, Tensorflow, PyTorch)

**Machine Learning**

Supervised and Unsupervised Learning, Neural Network Architectures, Natural Language Processing, Sentiment Analysis, Computer Vision.

**Databases**

Cassandra, MySQL, Neo4j, Oracle

**Developing environments**

Git, GitHub, PyCharm, Anaconda, Jupyter Notebook, Google Colab, Visual Studio Code, IntelliJ IDEA

**Spoken languages**

English (Professional)

- Level B2 certification issued by the teacher responsible for the English language courses for the Degree in Computer Science of the University of Bari.

Italian (Native)