PhD in Information Technology and Electrical Engineering
Università degli Studi di Napoli Federico II

PhD Student: Valerio La Gatta

Cycle: XXXVI

Training and Research Activities Report

Year: First

Tutor: prof. Vincenzo Moscato   _____

Co-Tutor:

Date: October, 2021

1. Information:

- ➢ **PhD student: Valerio La Gatta**
- ➢ **DR number: DR995141**
- ➢ **Date of birth: 15/01/1996**
- ➢ **Master Science degree: Computer Engineering   University: Università degli Studi di Napoli Federico II**
- ➢ **Doctoral Cycle: XXXVI**
- ➢ **Scholarship type: UNINA**
- ➢ **Tutor: Prof. Moscato Vincenzo**
- ➢ **Co-tutor:**

2. Study and training activities:

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
Digital Project Management: Practices, processes, techniques, tools and scientific approach	Seminar	1	0.2	18/11/2020	Prof. Flora Amato, Prof. Giuseppe Longo	Y
#andràtuttobene: Images, Texts, Emojis & Geodata in a Sentiment Analysis Pipeline	Seminar	1.5	0.3	25/11/2020	Prof. Flora Amato, Prof. Giuseppe Longo	Y
At the Nexus of Big Data, Machine Intelligence, and Human Cognition	Seminar	1	0.2	02/12/2020	Prof. Flora Amato, Prof. Giuseppe Longo	Y
Exploiting Deep Learning and Probabilistic Modeling for Behavior Analytics	Seminar	1	0.2	09/12/2020	Prof. Flora Amato, Prof. Giuseppe Longo	Y
Data Driven Transformation in WINDTRE through Managers voice	Seminar	2	0.4	16/12/2020	Prof. Flora Amato, Prof. Giuseppe Longo	Y
GDPR basics for computer scientists	Seminar	1.5	0.3	10/12/2020	DIETI	Y
Explainable Artificial Intelligence and Fuzzy Systems	Seminar	1	0.2	17/12/2020	The IEEE Italy Section Computatio	Y

					nal Intelligence Society	
Artificial Intelligence Between Research and Industry	Seminar	2	0.4	07/12/2020	The Consulate General of Italy for Scotland and Norther Ireland, the Italian Institute of Culture in Edinburgh and the Scotland Office of the Italian Chamber of Commerce and Industry for the UK	Y
AIRO PhD School 2021 and 5th AIRO-Young Workshop	Doctoral School	16	3.6	08/02/2021 - 10/02/2021	Prof. Antonio Sforza, Prof. Maurizio Boccia, Prof. Claudio Sterle, Dr. Adriano Masone	Y
From Photometric Redshifts to Improved Weather Forecasts: an interdisciplinary view on machine learning	Seminar	1	0.2	13/01/2021	Prof. Flora Amato, Prof. Giuseppe Longo	Y
Cybercrime and electronic evidence: the internation legal framework for an effective criminal justice response	Seminar	1	0.2	20/01/2021	Prof. Flora Amato, Prof. Giuseppe Longo	Y

AI LEGAL: Artificial Intelligence for notary's sector: a case study, Salvatore Falange	Seminar	1	0.2	27/01/2021	Prof. Flora Amato, Prof. Giuseppe Longo	Y
The era of Industry 4.0: new frontiers in business model innovation	Seminar	1	0.2	03/02/2021	Prof. Flora Amato, Prof. Giuseppe Longo	Y
Machine Learning: causality lost in translation	Seminar	1.5	0.3	10/02/2021	Prof. Flora Amato, Prof. Giuseppe Longo	Y
Approaches to Graph Machine Learning	Seminar	1	0.2	17/02/2021	Prof. Flora Amato, Prof. Giuseppe Longo	Y
Scientific Programming and Visualization with Python	Course	18	2	08/03/2021 - 10/03/2021	DIST, Prof. Alessio Botta	Y
Statistical data analysis for science and engineering research	Course	12	4	17/02/2021 - 04/03/2021	DIETI, Prof. Roberto Pietrantuono	Y
Data science for patient records analysis	Course	10	2.5	10/02/2021 - 17/03/2021	ITEE, ICTH, Prof. Marcello Cinque	Y
Big Data and Computational Linguistics	Seminar	2	0.4	10/03/2021	Prof. Flora Amato, Prof. Giuseppe Longo	Y
Sensoria Health	Seminar	1	0.2	17/03/2021	Prof. Flora Amato, Prof. Giuseppe Longo	Y
Distributional Semantics Methods: How Linguistic features can improve the semantic	Seminar	2	0.4	28/04/2021	Prof. Flora Amato, Prof. Giuseppe Longo	Y

representation						
Robo Ludens: A game design taxonomy for human-robot interaction	Seminar	1	0.2	05/03/2021	Prof. Silvia Rossi	Y
IEEE Authorship and Open Access Symposium: Best Practices to Get Published to Increase the Exposure and Impact of Your Research	Seminar	2	0.4	21/04/2021	IEEE	Y
Natural Language Processing	Course	48	6	10/03/2021 - 11/06/2021	Prof. Francesco Cutugno	Y
Introduction to Underwater Robotics	Seminar	2	0.4	18/05/2021	Dr. Fabio Ruggiero	Y
5G: l'architettura, le applicazioni e la rete di accesso radio,	Seminar	2	0.4	08/06/2021	Prof. Nicola Pasquino	Y
Sadas Engine, an innovative DBMS for the Data Warehouse, great performance in the VLDB environment	Seminar	1	0.2	23/06/2021	Prof. Flora Amato, Prof. Giuseppe Longo	N
Teaching activities regarding practical lectures/seminars during the courses of "Hardware and Software for Big Data" and "Machine Learning and Big Data per la salute"	Tutorship	40	3.6	01/05/2021 - 30/06/2021	Prof. Vincenzo Moscato	N
Strategic Orientation for STEM research & writing	Course	18	4	15/06/2021 - 14/10/2021	DIETI, Ms Chie Shin Fraser	Y

1)   Courses, Seminar, Doctoral School, Research, Tutorship
2)   Choose: Y or N

2.1. Study and training activities - credits earned

	Courses	Seminars	Research	Tutorship	Total

Bimonth 1	**0**	**2.2**	**7.8**	**0**	**10**
Bimonth 2	**3.6**	**1.3**	**4.8**	**0**	**9.7**
Bimonth 3	**8.5**	**1.6**	**2**	**0**	**12.1**
Bimonth 4	**6**	**1**	**1.4**	**1.6**	**10**
Bimonth 5	**0**	**0**	**5**	**0**	**5**
Bimonth 6	**4**	**0**	**13.2**	**0**	**17.2**
Total	**22.1**	**6.1**	**34.2**	**1.6**	**64**
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	

3. Research activity:

During my first year of PhD, I carried out two research activities within my research field, namely *automatic fact-checking and fake news detection* and *eXplainable Artitificial Intelligence*.

- **Automatic fact-checking and fake news detection**
  Although fake news is not a new phenomenon, since the last decade it has become one of the major threats to democracy, journalism, and freedom of expression. The rise of social media has been playing a key-role since those platforms enable the creation, the publication and the consumption of news online faster and cheaper [1]. As a result, huge amount of false information which spreads across the population affects our life. The most reliable solution to this problem has consisted in the increasing number of fact-checking organizations which are arising to debunk false information though domain experts' analyses. However, the quantity of information is much more than the one people can effectively check and thus, it has become essential to speed up and/or (semi-)automate the verification process.
  In the first year of PhD, I studied the whole fact-checking pipeline, composed by four stages [2]: (1) *claim detection*, i.e. estimating check-worthiness of a given text; (2) *verified claim retrieval*, i.e. detecting previously fact-checked information; (3) *evidence retrieval*, i.e. selecting possible evidences supporting or refuting the input statement; (4) *claim verification*, i.e. assessing claim's truthfulness.
  In particular, I focused on the second step since, based on the evidence that the same viral claim is often reposted by thousands of people in a short time-frame or also after a while in a different context, detecting previously fact-checked information can ease the manual fact-checkers' work, increasing their productivity and thus their effectiveness, but also could improve the automatic verification process since the veracity prediction of an input claim could be based on a set of already verified information. The verified claim retrieval problem has been considered only recently by [3] which formulates the information retrieval task of ranking a list of verified documents according to the relevance with an input claim. From this perspective, I considered the standard retriever-reranker architecture [4] adopted in modern information retrieval and question-answering tasks, and benchmark the state-of-the-art models especially considering the recent advances that neural ranking models and transformer-based systems have brought to both stages of the architecture. This work, in collaboration with prof. Tanmoy Chakraborty from IIT Delhi, has led to the paper "Information retrieval algorithms and neural

ranking models to detect previously fact-checked information", which has been submitted to the journal Knowledge Based Systems.

Currently, I am working on a new model capable of exploiting multimodal data, i.e. texts and images of the input claim and of the verified documents, to improve both the retriever and the re-ranker performance. Further future works include the analysis of multimodal input in the other stages of the fact-checking pipeline, the role of time within the detection process and the connection between fake news detection and hate speech diffusion within social network.

- **eXplanaible Artificial Intelligence (XAI)**
  Nowadays, sophisticated machine learning models are assisting humans in the most disparate domains, but they usually are so complex and intricate that no one - neither their designers - is able to understand the "cognitive" processes leading to decisions. The main reason for the current focus on the XAI topic is the increasing number of AI applications, especially in critical domains (e.g. medicine, military, finance) where the interpretability of decisions is becoming a mandatory requirement. According to DARPA [2], while producing more explainable models, it is necessary to maintain a high level of learning performance (e.g. prediction accuracy) and to enable human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners. In the quest of reaching these goals, XAI research is divided into two branches: (1) *model-specific* techniques, which deal with the inner workings of specific models, and (2) *model-agnostic* techniques, which aim to design new systems capable of explaining predictions of any existing models. In addition, XAI methodologies differ on the scope of interpretability, which can be *local*, if the explanation refers to an individual instance, or *global*, if it refers to the overall behavior of the model [3].
  In the first year of PhD, I studied the state-of-the-art XAI techniques (e.g. LIME [4], SHAP [5]) and, in particular, those relying on *feature importance explanations*, i.e. the importance of predictors over the model's decisions. Their major limitation depends on the fact that, while being locally accurate, it is not clear how those explanations would apply to new instances. To this end, I designed a novel technique, namely PASTLE, which aims to enrich the above-mentioned feature-based explanations with information about how the prediction would change if feature changes were applied. This work has led to the paper "PASTLE: Pivot-Aided Space Transformation for Local Explanations", published in the journal Pattern Recognition Letters.
  In addition, I extended PASTLE framework to provide rule-based explanations. Specifically, clustering techniques have being employed to find sets of instances homogeneously classified by the predictive model. This work has led to the paper "CASTLE: Cluster-Aided Space Transformation for Local Explanations", published in the journal Expert System with Applications. While the current PASTLE/CASTLE frameworks deal with tabular data, the extension to provide explanations for image classification tasks is under development.
  In addition, I also contributed to the paper "Interpretability in AI-based Behavioral Malware Detection Systems", which has been submitted to IEEE Transactions on Dependable and Secure Computing, where we benchmark a wide range of state-of-the-art XAI techniques in the context of malware detection.

Further future works include the analysis of XAI requirements within the fact-checking pipeline (e.g. the fairness of the claim detection stage, the interpretability of the claim verification stage, the most useful explanation form w.r.t. the final end-user).

In parallel to the above-mentioned activities, a collaboration with the Campania region, regarding Covid-19 diffusion and the effects of government countermeasures, has led to the paper "An Epidemiological Neural network exploiting Dynamic Graph Structured Data applied to the COVID-19 outbreak", published in the journal IEEE Transactions on Big Data.

[1] X. Zhou and R. Zafarani, "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," *ACM Computing Surveys,* vol. 53, no. 5, pp. 1-40, 2020.

[2] D. Gunning and D. W. Aha, "DARPA's Explainable Artificial Intelligence (XAI) Program," *Ai Magazine,* vol. 40, no. 2, pp. 44-58, 2019.

[3] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access,* vol. 6, pp. 52138-52160, 2018.

[4] M. T. Ribeiro, S. Singh and C. Guestrin, ""Why Should I Trust You?": Explaining the Predictions of Any Classifier," in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, 2016.

[5] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *NIPS'17 Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017.

4. Research products:

- **V. La Gatta**, V. Moscato, M. Postiglione, G. Sperlì; "An Epidemiological Neural network exploiting Dynamic Graph Structured Data applied to the COVID-19 outbreak"; IEEE Transactions on Big Data, IEEE TBD; published; 2020
- **V. La Gatta**, V. Moscato, M. Postiglione, G. Sperlì; "CASTLE: Cluster-Aided Space Transformation for Local Explanations"; Expert Systems with Applications, ESWA; published; 2021
- **V. La Gatta**, V. Moscato, M. Postiglione, G. Sperlì; "PASTLE: Pivot-Aided Space Transformation for Local Explanations"; Pattern Recognition Letters, PRL; published; 2021
- T. Chakraborty, **V. La Gatta**, V. Moscato, G. Sperli; "Information retrieval algorithms and neural ranking models to detect previously fact-checked information"; Knowledge-Based Systems, KBS; submitted
- **V. La Gatta**, V. Moscato, M. Pennone, M. Postiglione, G. Sperlì; "Music Recommendation via Hypergraph Embedding"; IEEE Transactions on Neural Networks and Learning Systems, IEEE TNNLS; submitted

- **V. La Gatta**, V. Moscato, M. Postiglione, G. Sperlì; "Few-shot Named Entity Recognition with Cloze Questions"; NeurIPS Data-Centric AI Workshop; submitted
- A. Galli, **V. La Gatta**, V. Moscato, M. Postiglione, G. Sperlì; "Interpretability in AI-based Behavioral Malware Detection Systems", IEEE Transactions on Dependable and Secure Computing, IEEE TDSC; submitted

5. Conferences and seminars attended

- 5th AIRO-Young Workshop, 11-12/02/2020, organized by the Operations Research Group, Department of Electrical Engineering and Information Technology (DIETI), University "Federico II" of Naples, online workshop

6. Activity abroad:

7. Tutorship

- Co-supervisor of five master theses in Computer Engineering
- Weekly two hours of teaching activities regarding practical lectures/seminars during the course "Hardware and Software for Big Data", Master Degree in Data Science
- Weekly two hours of teaching activities regarding practical lectures/seminars during the course "Machine Learning and Big Data per la salute", Master Degree in Biomedical Engineering