PERSONAL INFORMATION

Debora Nozza

Gender Female | Date of birth 15/11/1989 | Nationality Italian

EDUCATION

Feb 2020 - Feb 2021 **Research Fellow**, *Bocconi University*, Milan.

Supervisor: Prof. Dirk Hovy.

My research interests regard the detection and counter-acting of hate speech in text and bias in Machine Learning models which have become fundamental tasks for fighting against several types of discrimination (misogyny, xenophobia, racism, etc.) that subsist in Social Media.

Jan 2018 - Jan 2020 **Research Fellow in Computer Science**, *University of Milano-Bicocca*, Milan.

Title: "Deep Learning Representation for Large Attributed Graphs".

Supervisors: Prof. Enza Messina and Prof. Elisabetta Fersini.

My research activity focused on the exploitation and improvements of data representation obtained by Deep Learning techniques in order to address Natural Language Processing tasks, i.e. Sentiment Analysis, Named Entity Recognition and Linking, Hate Speech detection.

Dec 2014 - Oct 2018 **PhD in Computer Science**, *University of Milano-Bicocca*, Milan.

Thesis Title: "Deep Learning for Feature Representation in Natural Language Processing".

Supervisors: Prof. Enza Messina and Prof. Elisabetta Fersini.

Feb 2017 - May 2017 **Visiting PhD Student**, *INRIA*, Sophia Antipolis – Méditerranée.

Visiting PhD student at INRIA in the WIMMICS research group under the supervision of Dr. Fabien Gandon.

2011 - 2014 **Master's Degree in Computer Science**, *University of Milano-Bicocca*, Milan.

110/110 summa cum laude.

Fundamental Courses: Machine Learning, Data Analytics, Data and Text Mining, Probabilistic Models for Decision Making.

Thesis Title: "Learning representation for sentiment classification".

Supervisors: Prof. Enza Messina and Prof. Patrick Gallinari (University Pierre and Marie Curie, Paris).

July 2013 - Feb 2014 **Visiting Master Student**, *University Pierre and Marie Curie*, Paris.

Internship Student at Laboratoire d'Informatique de Paris 6 (LIP6) in the Machine Learning and Information Access research group, under the supervision of Prof. Patrick Gallinari.

2008 - 2011 **Bachelor's Degree in Computer Science**, *University of Milano-Bicocca*, Milan.

Fundamental Courses: Data Mining, Business Intelligence, Statistics, Computational Methods, Computational Models, Analysis of Financial Systems

Title of the Thesis: "Experimenting techniques of constrained inference for Information Extraction".

Supervisors: Prof. Enza Messina and Prof. Elisabetta Fersini.

RESEARCH ACTIVITIES

Talks and Posters

- Talk "Towards a Constrained deep Attributed Graph Embedding" at the *49th Annual Conference of the Italian Operations Research Society* (Genoa, Italy, September 5, 2019).
- Talk "Word Embeddings for Unsupervised Named Entity Linking" at the *12th International Conference on Knowledge Science, Engineering and Management* (Athens, Greece, August 30, 2019).
- Talk "A study on biased words for Hate Speech against women" for the Milan Women in Machine Learning and Data Science meetup event (Milan, Italy, 2019).
- Talk "Dal Machine Learning al Deep Learning, come le macchine interpretano il linguaggio" for the Milan Digital Week event (Milan, Italy, 2019).
- Talk "Adapting Named Entity Types to New Ontologies in a Microblogging Environment" at the *31st International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems* (Montreal, Canada, 2018).
- Seminar on "Mentions and Synsets embeddings induced from a heterogeneous graph structure" at INRIA Sophia Antipolis - Méditerranée (Nice, France, 2017).
- Poster "Towards Adaptation of Named Entity Classification" at the *32nd ACM Symposium on Applied Computing* (Marrakesh, Morocco, 2017).
- Poster "UNIMIB@NEEL-IT: Named Entity Recognition and Linking of Italian Tweets" at the *5th International Workshop EVALITA* (Naples, Italy, 2016).
- Invited talk "Deep learning in natural language processing" at the conference of *AI, Machine Learning and Sentiment Analysis Applied to Finance* (London, UK, 2016).

Editorial Activities

- Reviewer for Computers & Operations Research - Elsevier, Expert Systems with Applications - Elsevier, Journal of Web Semantics - Elsevier and Information - MDPI.
- Reviewer for ACL 2020.
- Program Committee member of the Widening NLP Workshop 2020 at ACL 2020.
- Technical Program Committee member of the Ninth International Conference on Social Media Technologies, Communication, and Informatics, SOTICS 2020.
- Technical Program Committee member of the Fifteenth International Conference on Internet and Web Applications and Services, ICIW 2020.
- Area Chair at the 14th Women in Machine Learning Workshop (WiML) at NeurIPS 2019.
- Session Chair at the 12th International Conference on Knowledge Science, Engineering and Management, 2019.
- Reviewer for Widening NLP Workshop at ACL 2019.
- Local Committee member of the Second International Conference on Data Science and Social Research, 2019.
- Review committee member for The Eighth Joint Conference on Lexical and Computational Semantics, 2019.
- Program Committee member for the Social Media track in The 2019 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies, 2019.
- Program Committee member of NL4AI 2019, the 3nd Workshop on Natural Language for Artificial Intelligence, 2018 - 2019.
- Technical Program Committee member of the Ninth International Conference on Social Media Technologies, Communication, and Informatics, SOTICS 2019.
- Session Chair at the 31st International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems, 2018.
- Program Committee member of Italian Conference on Computational Linguistics, 2017.

Teaching experiences	*University of Milano-Bicocca*, Milan.

— Teaching Assistant for the course of "*Data Analytics*" for the master degree in Computer Science. Overall hours: 24. Period: Feb 2020 - Jul 2020.
— Teaching Assistant for the course of "*Data Analytics*" for the master degree in Computer Science. Overall hours: 24. Period: Feb 2019 - Jul 2019.
— Teaching Assistant for the course of "*Data Analytics*" for the master degree in Computer Science. Overall hours: 12. Period: Feb 2018 - Jul 2018.
— Teaching Assistant for the course of "*Data Analytics*" for the master degree in Computer Science. Overall hours: 12. Period: Feb 2017 - Jul 2017.
— Teaching Assistant for the course of "*Probabilistic Models for Decision Making*" for the master degree in Computer Science. Overall hours: 30. Period: Sep 2017 - Jan 2018.
— Teaching Assistant for the course of "*Laboratory of programming in Java*" for the bachelor degree in Mathematics. Overall hours: 30. Period: Sep 2016 - Jan 2017.
— Teaching Assistant for the course of "*Laboratory of programming in Java*" for the bachelor degree in Mathematics. Overall hours: 30. Period: Sep 2015 - Jan 2016.
— Co-Supervisor of 15 bachelor's and 3 master's degree thesis in Computer Science.

Industrial Research Projects

— Collaboration with Clinical Insights Ltd for the project "Patient profiling for drug complaince" (Jan 2018 - Feb 2018).
— Collaboration with Consorzio Milano Ricerche for the project regarding the development, exploration and writing of "Proof of Concept: Social Media Sentiment Analytics (Brexit)" (Nov 2016 - Mar 2017).
— Collaboration with Consorzio Milano Ricerche for the project regarding the analysis and presentation of "Methods for the analysis of longitudinal prescription data" (Nov 2014 - Dec 2014).

International Research Collaboration

— Natural Language Engineering Lab at Pattern Recognition and Human Language Technologies (PRHLT) Research Center, Universitat Politècnica de València, València.
— WIMMICS research group, INRIA Sophia Antipolis – Méditerranée, Nice.
— Centre for the Analysis of Risk and Optimisation Modelling Applications (CARISMA), Brunel University, London.
— Machine Learning and Information Access research group, Laboratoire d'Informatique de Paris 6 (LIP6), Paris.

Awards and Scholarships

— *Runner up for the Best Paper Award* at the 2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2019).
— *Best Paper Award* at the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (2016).
— PhD Scholarship funded by the Italian Ministry of Research and Education.
— In 2013, I received two scholarship provided by the University of Milano-Bicocca for a research period at the research laboratory LIP6 of the University Pierre and Marie Curie (Paris, France): one for the program "Erasmus Placement" and one for the program "EXTRA plus international mobility" (with the contribution of Assolombardo and co-financed by "Fondazione Cariplo").

PERSONAL SKILLS

Mother tongue	Italian

Other languages	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken interaction	Spoken production	
English	C1	C1	C1	C1	C1
French	B1	B1	B1	B1	A2

Levels: A1/A2: Basic user - B1/B2: Independent user - C1/C2: Proficient user
Common European Framework of Reference (CEF) level

Computer skills
- Excellent use of the most widely used Operating Systems (Linux, Windows, Macintosh).
- Experience with distributed version control systems (git, SVN).
- Good knowledge of the database querying language SQL and the database management system MySQL.
- Good knowledge of programming languages and software: C++, Cytoscape, Knime, Weka.
- Advanced knowledge of JAVA, Matlab, R, Python.
- Proficiency with several statical tools: Pandas, scikit-learn, TensorFlow, Keras, NLTK.

PUBLICATIONS

— **D. Nozza**, E. Fersini, E. Messina. "*CAGE: Constrained deep Attributed Graph Embedding*". Information Sciences. 2019.

— **D. Nozza**, C. Volpetti, E. Fersini. "*Unintended Bias in Misogyny Detection*". In Proceedings of the 2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2019). Thessaloniki, Greece, October 14-17, 2019. (**Runner up for the Best Paper**)

— **D. Nozza**, C. Sas, E. Fersini, E. Messina. "*Word Embeddings for Unsupervised Named Entity Linking*". In Proceedings of the 12th International Conference on Knowledge Science, Engineering and Management (KSEM 2019). Athens, Greece, August 28-30, 2019.

— V. Basile, C. Bosco, E. Fersini, **D. Nozza**, V. Patti, F. Rangel, P. Rosso, M. Sanguinetti. "*SemEval-2019 Task 5: Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter*". In Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval-2019). Minneapolis, Minnesota, June 6-7, 2019.

— E. Fersini, **D. Nozza**, P. Rosso. "*Overview of the Evalita 2018 Task on Automatic Misogyny Identification (AMI)*". In Proceedings of the 6th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA'18). Turin, Italy, December 10-12, 2018.

— V. Basile, E. Cabrio, F. Gandon and **D. Nozza**. "*Mapping Natural Language Labels to Structured Web Resources.*" In Proceedings of the NL4AI Workshop, co-located with the International Conference of the Italian Association for Artificial Intelligence (AI* IA). Trento, Italy, November 22-23, 2018.

— F. Bianchi, M. Palmonari, **D. Nozza**. "*Towards Encoding Time in Text-Based Entity Embeddings*". In Proceedings of the 17th International Semantic Web Conference. Monterey, California, USA, October 8-12, 2018.

— E. Fersini, P. Manchanda, E. Messina, **D. Nozza**, M. Palmonari. "*Adapting Named Entity Types to New Ontologies in a Microblogging Environment*". In Proceedings of the 31st International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. Montreal, Canada, June 25-28, 2018.

— **D. Nozza**, F. Ristagno, M. Palmonari, E. Fersini, P. Manchanda, E. Messina. "*TWINE: A real-time system for TWeet analysis via INformation Extraction*". In Proceedings of the Demonstrations at the 15th Conference of the European Chapter of the Association for Computational Linguistics. Valencia, Spain, April 3-6, 2017.

— **D. Nozza**, E. Fersini, E. Messina. "*A Multi-View Sentiment Corpus*". In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. Valencia, Spain, April 3-6, 2017.

— P. Manchanda, E. Fersini, **D. Nozza**, E. Messina, M. Palmonari. "*Towards Adaptation of Named Entity Classification*". In Proceedings of the 32nd ACM Symposium on Applied Computing. Marrakesh, Morocco, April 3-6, 2017.

- F. M. Cecchini, E. Fersini, P. Manchanda, E. Messina, **D. Nozza**, M. Palmonari, C. Sas. "UNIMIB@NEEL-IT: Named Entity Recognition and Linking of Italian Tweets". In Proceedings of the 5th International Workshop EVALITA. Naples, December 7th, 2016.

- **D. Nozza**, E. Fersini, E. Messina. "*Unsupervised Irony Detection: A Probabilistic Model with Word Embeddings*". In Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. Porto, November 10th, 2016. (**Best Paper Award**)

- **D. Nozza**, E. Fersini, E. Messina. "*Deep Learning and Ensemble Methods for Domain Adaptation*". In Proceedings of the 28th IEEE International Conference on Tools with Artificial Intelligence. San Jose, CA, November 6th, 2016.

- **D. Nozza**, D. Maccagnola, V. Guigue, E. Messina and P. Gallinari. "*A latent representation model for sentiment analysis in heterogeneous social networks*". In Proceedings of the 3rd International on Modelling and Knowledge Management applications. Grenoble, September 2nd, 2014.

I hereby authorize the use of my personal details solely for circulation within the company according to the Italian Legislative Decree n. 196/2003.