Flor Miriam Plaza del Arco

Curriculum Vitae

BIO

The last year Ph.D. student at the SINAI group at the University of Jaén. My research interests lie in the area of natural language processing, particularly in hate speech detection, sentiment and emotion modeling, social media mining and deep learning. I am experienced in creating corpora to promote NLP research in emotion analysis and offensive language detection and developing and enhancing state-of-the-art deep learning models to address different NLP tasks. I am known for being an autonomous, self-motivated, and well-organized team member. I love to create solutions to new problems and collaborate with others to share and expand my knowledge.

Education

2018–Present **PhD, Computer Science**, *SINAI research group, University of Jaén*, Spain.
Offensive language detection in text, hate speech detection, emotion modelling, mental disorders identification, natural language processing, social media mining, machine learning, deep learning.

2016–2018 : **M.S., Computer Science**, *University of Jaén*, Spain, Thesis title: Computational Analysis of Emotions in Social media.

2012–2016 : **B.S., Computer Science**, *University of Jaén*, Spain, Final project title: Patient Opinion Analysis in Web 2.0.

Research Experience

May,2022 – July,2022 **Intern/Researcher, Natural Language Processing (HLT-NLP), Fondazione Bruno Kessler, Trento, Italy**.

Advisor : Dr. Carlo Strapparava.

May,2021 – July,2021 **Intern/Researcher, Theoretical Computational Linguistics (TCL), Institute for Natural Language Processing, University of Stuttgart, Germany**.

Advisor : Dr. Roman Klinger.

April,2019 – June,2019 **Intern/Researcher, Natural Language Processing (HLT-NLP), Fondazione Bruno Kessler, Trento, Italy**.

Advisor : Dr. Carlo Strapparava.

Nov,2019 **Visiting Researcher, Research Institute in Information and Language Processing, University of Wolverhampton, UK**.

Advisor : Dr. Ruslan Mitkov.

Sep,2016 – May,2017 **Andalusian Public Foundation for Biosanitary Research, Jaén, Spain**.

Advisor : Prof. Victor Manuel Rivas Santos.

Fellowships, Honors, Recognitions & Outstanding Achievements

2019-Present **PhD Scholarship** funded by the Spanish Ministry of Science, Innovation, and Universities.

Sep,2016 – Present **Fellow** of the Society for Natural Language Processing, Spain.

2021 **Ranked 1st** among 31 teams in the DETOXIS Shared Task on toxicity detection in Spanish (IberLEF, SEPLN).

2021 **Ranked 2nd** among 72 submitted systems in the EXIST Shared Task on sexism identification in social networks (IberLEF, SEPLN).

2021 **First prize** in the "Neologisms and technical terms in the field of energy and environmental energy and environment" Hackathon with the POEMA project. The Royal Spanish Academy and the Endesa Foundation.

2021 **Provincial winner** in the general category of the "II Edition of the Contest Your Doctoral Thesis in a Twitter Thread: #HiloTesis". Spanish Foundation for Science and Technology of the Spanish Ministry of Science and Innovation.

2020 **First prize** for the best introductory research work (IV Ada Lovelace Awards). University of Jaén.

2018 **2nd Team Rank** for the application "Geographical Disease Dispersion Monitor". II Hackathon of Language Technologies, Mobile World Congress, Barcelona, Spain.

2018 **5th Best academic record** of in the M.S., Computer Science.

2016 **4th Best academic record** of in the B.S., Computer Science.

Program, Scientific & Organization Committees

2021 37th Conference of the Spanish Society for Natural Language Processing. Organizing Committee.

2020 36th Conference of the Spanish Society for Natural Language Processing. Organizing Committee.

2020 The 12th Language Resources and Evaluation Conference. Scientific Comittee.

2018 The 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. Program Committee.

Organization of International Shared Tasks

2021 MeOffendEs: Offensive Language Detection in Spanish Variants at the Iberian Languages Evaluation Forum (IberLEF) co-located with the 37th International Conference of the Spanish Society for Natural Language Processing (SEPLN'21).

2021 EmoEvalES: Emotion Detection for Spanish at the Iberian Languages Evaluation Forum (IberLEF) co-located with the 37th International Conference of the Spanish Society for Natural Language Processing (SEPLN'21).

2020 TASS 2020: Introducing Emotion Detection at the Iberian Languages Evaluation Forum (IberLEF) co-located with the 36th International Conference of the Spanish Society for Natural Language Processing (SEPLN'20).

Journal & Conference Reviewer

Journals

Information Processing & Management

Expert Systems with Applications

Artificial Intelligence Review

Natural Language Engineering

Procesamiento de Lenguaje Natural Journal, Journal for the Spanish Society of Natural Language Processing

SN Social Sciences

Conferences

Empirical Methods in Natural Language Processing (EMNLP)

Association for Computational Linguistics (ACL)

European chapter of the Association for Computational Linguistics (EACL)

North American Chapter of the Association for Computational Linguistics (NAACL)

Computational Linguistics (COLING)

European Conference on Artificial Intelligence (ECAI)

Recent Advances in Natural Language Processing (RANLP)

Language Resources and Evaluation Conference (LREC)

International Conference on Natural Language & Information Systems (NLDB)

Spanish Society for Natural Language Processing (SEPLN)

Selected Publications

Journal Articles

2021 **Flor Miriam Plaza-del-Arco**, M. Dolores Molina-González, L. Alfonso Ureña-López, and María Teresa Martín-Valdivia. A Multi-Task Learning Approach to Hate Speech Detection Leveraging Sentiment Analysis. *IEEE Access*, volume 9, pages 112478–112489, 2021.

2021 **Flor Miriam Plaza-del-Arco**, Salud María Jiménez-Zafra, Arturo Montejo-Ráez, M. Dolores Molina-González, L. Alfonso Ureña-López, and M. Teresa Martín-Valdivia. Comparing pre-trained language models for Spanish hate speech detection. *Expert Systems with Applications*, volume 166, pages 155–161, 2021.

2020 **Flor Miriam Plaza-del-Arco**, M. Dolores Molina-González, L. Alfonso Ureña-López, and M. Teresa Martín-Valdivia. Detecting Misogyny and Xenophobia in Spanish Tweets Using Language Technologies. *ACM Transactions on Internet Technology*, volume 20, 2020.

2020 **Flor Miriam Plaza-del-Arco**, M. Teresa Martín-Valdivia, L. Alfonso Ureña-López, and Ruslan Mitkov. Improved emotion recognition in Spanish social media through incorporation of lexical knowledge. *Future Generation Computer Systems*, volume 110, pages 1000–1008, 2020.

In Conference Proceedings

2022 **Flor Miriam Plaza-del-Arco**, Ana Belén Parras Portillo, Pilar López-Úbeda, Beatriz Gil, and María-Teresa Martín-Valdivia. SHARE: A Lexicon of Harmful Expressions by Spanish Speakers. In *Proceedings of the 13th Language Resources and Evaluation Conference*, page 1307–1316, 2022.

2022 **Flor Miriam Plaza-del-Arco**, María-Teresa Martín-Valdivia, and Roman Klinger. Natural language inference prompts for zero-shot emotion classification in text across corpora. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 6805–6817, Gyeongju, Republic of Korea, October 2022. International Committee on Computational Linguistics.

2021 **Flor Miriam Plaza-del-Arco**, Arturo Montejo-Ráez, L. Alfonso Ureña-López, and M. Teresa Martín-Valdivia. OffendES: A New Corpus in Spanish for Offensive Language Research. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing*, pages 1096–1108, 2021.

2021 **Flor Miriam Plaza-del-Arco**, Sercan Halat, Sebastian Padó, and Roman Klinger. Multi-Task Learning with Sentiment, Emotion, and Target Detection to Recognize Hate Speech and Offensive Language. In *Proceedings of Hate Speech and Offensive Content Identification in Indo-European Languages (HASOC) at FIRE*, 2021.

2020   **Flor Miriam Plaza-del-Arco**, Carlo Strapparava, L. Alfonso Ureña López, and M. Teresa Martín-Valdivia. EmoEvent: A Multilingual Emotion Corpus based on different Events. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 1492–1498, 2020.

2019   Pilar López Úbeda, **Flor Miriam Plaza-del-Arco**, Manuel Carlos Díaz Galiano, L. Alfonso Ureña-López, and M. Teresa Martín-Valdivia. Detecting Anorexia in Spanish Tweets. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing*, pages 655–663, 2019.

Invited Talks

November, 2022   *"El procesamiento del lenguaje natural desde la perspectiva de los especialistas en lingüística y en ingeniería"* (Natural language processing from the perspective of linguistics and computer science specialists). Polytechnic University of Madrid, Spain

July, 2022   *"Avanzando en la detección del lenguaje ofensivo en español"* (Advancing in the detection of offensive language in Spanish). NLP Spain community. Online: `https://bit.ly/3BrcLgC`

June, 2022   *"OffendES: un nuevo recurso lingüístico en español para combatir el discurso de odio en las RRSS"* (OffendES: a new linguistic resource in Spanish to combat hate speech in social networks). Round table "Combating hate speech using HLT". Thematic network of excellence funded by the Ministry of Science, Innovation and Universities of Spain. Online: `https://bit.ly/3PKSE16`

November, 2021   *"Multi-task learning with sentiment, emotion, and target detection to recognize hate speech and offensive language"*. Institute for Natural Language Processing, University of Stuttgart. Online.

Computer skills

Operating Systems   Linux, Windows

Programming Languages   Python, C++, Java

Deep Learning   PyTorch, Keras, Hugging Face

Python NLP libraries   NLTK, Pandas, SciPy, Scikit-learn, SpaCy, SciPy, Matplotlib

Typesetting: LaTeX, Vim.

Languages

Spanish   Mother tongue

English   Professional proficiency

French & Italian   Basics

Teaching Assistantship

2022   **Human–Computer Interaction. Bachelor in Computer Science**, University of Jaén.

2021   **Natural Language Processing. Bachelor in Computer Science**, University of Jaén.

2021   **Information Retrieval. Bachelor in Computer Science**, University of Jaén.

2020   **Operating Systems. Bachelor in Computer Science**, University of Jaén.

2020   **Programming. Bachelor in Industrial Organization**, University of Jaén.

Social Media & Websites

@florplaza22

in https://bit.ly/3MUUfiL
Google Scholar: https://bit.ly/3MNmWhq
ORCID: https://orcid.org/0000-0002-3020-5512
Personal Website: https://fmplaza.github.io/

References

References available upon request