Jesus German **Ortiz Barajas**

+52 55-2883-2704 | jgermanmx@gmail.com | jgermanob.github.io

Education

Institute for Computer Science, Artificial Intelligence and Technology (INSAIT) *Sofia, Bulgaria*
PhD. in Computer Science *Nov 2024 - Present*
- Advisor: Prof. Iryna Gurevych

Universidad Nacional Autónoma de México (UNAM) *Mexico City, Mexico*
MSc. in Computer Science and Engineering *Sept 2020 - Dec 2022*
- GPA: 9.80/10.00
- Thesis: Paraphrase identification using Sentence-CROBI, a novel deep neural network architecture based on cross-encoders and bi-encoders
- Advisor: Dr. Gemma Bel-Enguix

Universidad Nacional Autónoma de México (UNAM) *Mexico City, Mexico*
Diploma in Mobile App Development for iOS *Feb 2019 - July 2019*
- GPA: 9.00/10.00

Universidad Nacional Autónoma de México (UNAM) *Mexico City, Mexico*
BEng. in Computer Engineering *Aug 2014 - Dec 2018*
- GPA: 8.54/10.00

Professional experience

Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI): Natural Language Processing Deparment *Abu Dhabi, United Arab Emirates*
Research assistant *Jan 2024 - October 2024*
- Working on parameter-efficient fine-tuning techniques for Large Language Models (LLMS).
- Contributed to CVQA, a multicultural, multilingual question-answering benchmark by collecting, labelling, and validating image-question pairs in Mexican Spanish and English.
- Mentor at the MBZUAI Undergraduate Research Internship Program (UGRIP).

Grupo de Ingeniería Lingüística: UNAM research group focusing on computational linguistics and natural language processing. *Mexico City, Mexico*
Shared Task organiser *Jan 2022 - June 2022*
- Participated in organising the PAR-MEX 2022 shared task: Paraphrase detection in Mexican Spanish at IberLEF 2022.
- Designed baseline experiments using a Spanish BERT-based model and Keras and Transformers libraries.
- Coordinated the codalab competition page and GitHub repository for the shared task.
- Analysed the results using the F1-score as a performance metric and the Maximum Possible Accuracy and the Coincident Failure Diversity as complementariness metrics.

Job Offers classification *Dec 2019 - March 2020*
- Developed system performs multi-class classification of job offers using a recurrent neural network.
- Financed by the Mexican government, its purpose is to reduce information asymmetries between job seekers and employers.
- Model building using Keras library. It consists of a Long-short term memory layer followed by a multilayer perceptron.
- Used SMOTE, Geometric-SMOTE and ADASYN algorithms to solve the high-imbalance-class problem.

Course instructor *Aug 2019 - Dec 2019*
- Semester course to explain the basics of machine learning with a theoretical-practical approach.
- An eight-student group took it with different academic backgrounds from UNAM and the private sector.
- Prepared lessons using the book *Artificial Intelligence with an introduction to machine learning* by Richard E. Neapolitan and Xia Jiang; Kaggle datasets, the sci-kit learn library and python scripts.

Aggressiveness detection on Twitter *Apr 2019 - June 2019*
- App that classifies Mexican Spanish tweets in aggressive or non-aggressive classes using machine learning and python.
- Uses multiple types of n-grams such as character n-grams, word n-grams, and aggressive words n-grams as features.
- Implements a support vector machine as a classifier with SciKit-learn and microTC framework for parameter optimization.
- Achieved an F1-score of 0.4549, 5th place of 26 competitors at MEX-A3T 2019 aggressiveness identification task.

Publications

Journals

Ortiz-Barajas, J-G., Bel-Enguix, G., Gómez-Adorno, H. 2022. Sentence-CROBI: A Simple Cross-Bi-Encoder-Based Neural Network Architecture for Paraphrase Identification. Mathematics, 10(19): 3578.

Bel-Enguix, G., Sierra, G., Gómez-Adorno, H., Torres-Moreno, J-M., **Ortiz-Barajas, J-G.**, Vásquez, J. 2022. Overview of PAR-MEX at Iberlef 2022: Paraphrase Detection in Spanish Shared Task. Procesamiento del Lenguaje Natural, 69(1): 255-263.

Conference papers

Ortiz-Barajas, J-G., Bel-Enguix, G., Gómez-Adorno, H. 2024. MBZUAI-UNAM at SemEval-2024 Task 1: Sentence-CROBI, a simple cross-bi-encoder-based neural network architecture for semantic textual relatedness. In Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024): 1060–1068

Ortiz, G., Bel-Enguix, G., Gómez-Adorno, H., Ameer, I., Sidorov, G. 2023. Job Offers Classifier using Neural Networks and Oversampling Methods. In Recent Developments and the New Directions of Research, Foundations, and Applications: Selected Papers of the 8th World Conference on Soft Computing, February 03–05, 2022, Baku, Azerbaijan, Vol. I (pp. 235-248).

Sierra, G., Bel-Enguix, G., Gómez-Adorno, H., Torres-Moreno, J-M., Hernández-García, T., Guadarrama-Olvera, JV., **Ortiz-Barajas, J-G.**, Rojas, AM., Damerau, T., Aragón Martinez, S. 2020. Enhancing Job Searches in Mexico City with Language Technologies. In Proceedings of the 1st Workshop on Language Technologies for Government and Public Administration (LT4Gov): 15-21.

Ortiz, G., Gómez-Adorno, H., Reyes-Magaña, J., Bel-Enguix, G., Sierra, G. 2019. Detection of Aggressive Tweets in Mexican Spanish Using Multiple Features with Parameter Optimization. In IberLEF@ SEPLN: 520-525.

Pre-prints

Ortiz-Barajas, J-G., Gómez-Adorno, H., Solorio, T. 2024. HyperLoader: Integrating Hypernetwork-Based LoRA and Adapter Layers into Multi-Task Transformers for Sequence Labelling.

Presentations

** presenting author*

Invited presentations

Bel-Enguix, G., Sierra, G., Gómez-Adorno, H., Torres-Moreno, J-M., **Ortiz-Barajas, J-G.***, Vásquez, J. 2022. Overview of PAR-MEX at Iberlef 2022: Paraphrase Detection in Spanish Shared Task. Oral presentation: 38th Conference of the Spanish Society for Natural Language Processing, A Coruña, Spain.

Conference presentations

Ortiz, G.* , Bel-Enguix, G., Gómez-Adorno, H., Ameer, I., Sidorov, G. 2022. Job Offers Classifier using Neural Networks and Oversampling Methods. Online presentation: 8th World Conference on Soft Computing, WConSC-2022, Baku, Azerbaijan.

Ortiz, G.*, Gómez-Adorno, H., Reyes-Magaña, J., Bel-Enguix, G., Sierra, G. 2019. Detection of Aggressive Tweets in Mexican Spanish Using Multiple Features with Parameter Optimization. Oral presentation: 4th Mexican Workshop on Plagiarism Detection and Authorship Analysis, Guanajuato, Mexico.

Teaching and Research Services

Jan 2023 - Jun 2023	**Graduate Teaching Assistant**, Text mining. Posgrado en Ciencia e Ingeniería de la Computación, UNAM.
May 2023	**Teaching**, Natural Language Processing. Colegio Científico de Datos, COCID.
Aug 2022	**Reviewer**, 21st Mexican International Conference on Artificial Intelligence.
Aug 2019 - Nov 2019	**Teaching**, Introduction to Machine learning. Instituto de Ingeniería, UNAM.

Awards & Honors

2023	**José Negrete award. Best master's thesis in an AI-related field**, Mexican Society for Artificial Intelligence (SMIA)

Skills

Languages	**Python, Swift, Java, C, PHP, Javascript.**
Libraries	**Tensorflow, Pytorch, Sci-kit learn, Pandas, Numpy, Scipy, NLTK, Spacy, Transformers.**
Tools	**Github, Latex, HTML.**