

Emilio Villa-Cueva

✉ evillacueva@gmail.com
☎ +52 477 106 49 17

villacu.github.io

EDUCATION

- **Mohamed bin Zayed University of Artificial Intelligence (MBZUAI)** Abu Dhabi, UAE
PhD in Natural Language Processing; August 2024 - Present
- **Mathematics Research Center (CIMAT)** Guanajuato, México
MSc in Computer Science and Industrial Mathematics; GPA: 9.86/10; July 2022 - July — 2024
 - Thesis: *Adaptation Techniques in Transformers for Text Classification in under-resourced settings*
- **University of Guanajuato** Guanajuato, México
Bachelor in Engineering Physics; GPA: 9.79/10 August 2017 - July 2022

RESEARCH EXPERIENCE

- **Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI)**
Research Assistant. Department of Natural Language Processing September 2023 - March 2023
 - *Focus:* Few-shot Crosslingual transfer.
 - *Supervisor:* **Dr. Tamar Solorio.**
- **Mathematics Research Center (CIMAT)**
Research Assistant. Department of Computer Science, NLP Group June 2021 - July 2024
 - *Focus:* Adversarial Training, Domain Adaptation Techniques, Question-Answering Systems, Few-Shot Classification.
 - *Supervisors:* **Dr. Adrian Pastor Lopez-Monroy** and **Dr. Fernando Sanchez Vega.**
- **University of Guanajuato**
Research Assistant. Division of Sciences and Engineering (DCI) January 2020 - February 2022
 - *Project:* Design and construction of a low-cost meteorological station in the DCI-UG campus to measure solar irradiance in the area.
 - *Supervisor:* **Dr. Modesto Sosa Aquino.**

PUBLICATIONS

1. E. Villa-Cueva, A. P. López-Monroy, F. Sanchez-Vega, and T. Solorio, “Adaptive Cross-Lingual Text Classification through In-Context One-Shot Demonstrations,” in *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics.*, Association for Computational Linguistics, June 2024 (**Accepted**)
2. E. Villa-Cueva, M. Valles-Silva, A. P. López-Monroy, F. Sanchez-Vega, and L.-S. J. Roberto, “Few Shot Profiling of Cryptocurrency Influencers using Natural Language Inference & Large Language Models,” in *CLEF 2023 Labs and Workshops, Notebook Papers*, 2023
3. E. Villa-Cueva, D. Vallejo-Aldana, F. Sanchez-Vega, and A. P. López-Monroy, “Walter Burns at SemEval-2023 Task 5: NLP-CIMAT - Leveraging Model Ensembles for Clickbait Spoiling,” in *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)*, (Online), Association for Computational Linguistics, 2023
4. E. Villa-Cueva, I. González-Franco, F. Sanchez-Vega, and A. P. López-Monroy, “NLP-CIMAT at PoliticEs 2022: PolitiBETO, a Domain-Adapted Transformer for Multi-class Political Author Profiling,” in *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2022)*, CEUR Workshop Proceedings, CEUR-WS.org, 2022
5. E. Villa-Cueva, F. Sanchez-Vega, and A. P. López-Monroy, “Bi-Ensembles of Transformer for Online Bilingual Sexism Detection,” in *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2022)*, CEUR Workshop Proceedings, CEUR-WS.org, 2022

AWARDS AND OTHER ACKNOWLEDGEMENTS

Awards

- First-place prize Award at the PAN-CLEF "Profiling Cryptocurrency Influencers with Few-shot Learning" shared task (**2023**)
Awarded by Symanto Research
- Full Scholarship for attending the "Montreal Industrial Problem Solving Workshop" at the CRM in University of Montreal. (**2023**)
Awarded by the University of Montreal
- Scholarship for Master studies at CIMAT Mexico (**2022**)
Awarded by CONAHCYT for throughout the duration of the degree (two years)
- Academic Merit Award (**2021**)
Awarded by the University of Guanajuato to the student that graduates with the highest GPA in their class.

PARTICIPATION IN WORKSHOPS AND EVENTS

- **Workshop Facilitator:** (2022)
Conducted a workshop on Domain Adaptation of Transformers at the MexLef 2022 conference in Cholula, Puebla. The session focused on techniques for adapting BERT models to specific domains.
- **Panelist, Mexican NLP Summer School 2024:** (2024)
Participated in the panel: "Getting into NLP, insights from experts and peers" at the Mexican NLP Summer School 2024, co-located with NAACL 2024.

PROJECTS

Projects from Hackathons and Others

- **Attribute inference attacks for synthetic data at Desjardins:** (Thirteenth Industrial Problem Solving Workshop, Montreal, Quebec)
Conducted attribute inference attacks to evaluate the risk of sensitive data exposure on synthetic data.
(June 2023)
- **Detecting collisions of trademarks:** (SPI 2023)
Designed a pipeline for detecting potential textual, phonetic, and semantic collisions between a query trademark and other trademarks registered in the Mexican Institute of Industrial Property.
(January 2023)
- **Object Detection and Document Binarization on scanned documents:** (RIIA "Justicia para los desaparecidos" Hackathon)
Adapted a *yolov5* architecture and an UNet to detect missing persons in low-quality scanned documents.
(September 2021)
- **Student well-being and predicting school dropout:** (UG BeeHack)
A conceptual app centered on student well-being along with a working model based on a Logistic Regression classifier to predict whether a student was at risk of dropping out based on their academic records.
(May 2021)

QUALIFICATIONS

- **Programming Languages.:** Python, R, C, C++, Matlab
- **Platforms:** Linux, Windows, MacOS Arduino, Texas Instruments, PIC Controllers