Emilio Villa Cueva

emilio.villa@cimat.mx +52 477 106 49 17

Education

- **Mathematics Research Center (CIMAT)** Guanajuato, México
  *Master of Science in **Computer Science and Industrial Mathematics**;* *July 2022 - Present*

  Relevant Coursework:

  - Natural Language Processing (Grade 10/10)
  - Machine Learning I (Grade 10/10)
  - Machine Learning II (Grade 9.5/10)
  - Optimization I (Grade 10/10)
  - Statistical Learning (Grade 10/10)
  - Data Analysis (Grade 9.5/10)
  - Numerical Methods (Grade 10/10)
  - Programming and Algorithms (Grade 9.5/10)

- **University of Guanajuato** Guanajuato, México
  *Bachelor in **Engineering Physics**; GPA: **9.79/10*** *August 2017 - July 2021*

Research Experience

- **Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI)**
  Research Assistant. *Department of Natural Language Processing* *September 2023 - March 2023*
  Worked on Cross-lingual Few-Shot transfer under the supervision of **Dr. Thamar Solorio**.

- **Mathematics Research Center (CIMAT)**
  Research Assistant. *Department of Computer Science* *June 2021 - September 2023*
  Worked on deep learning language models for a variety of tasks:

  - Adversarial Training
  - Domain Adaptation Techniques
  - Question-Answering Systems
  - Few-Shot Classification

  Supervised by **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*
  Worked on the design and construction of a low-cost meteorological station in the DCI-UG campus to measure solar irradiance in the area. This project was funded by the National Council of Science and Technology (*CONACyT*) and was carried out under the supervision of 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.

Workshops

- **Workshop Facilitator**: (2022, Cholula, Puebla)
  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.

Projects

Projects from Hackathons and Others

- **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)
  Proposed a solution that used a neural network based on the *yolov5* architecture to perform Object Detection and other based on an UNet architecture for document binarization 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 an Logistic Regression classifier to predict whether a student was at risk of dropping out based on their academic records.
  (May 2021)

Relevant School Projects

- **Forecasting short-term Solar Irradiance in the city of León**:
  Implementation of deep learning architectures such as LSTM and transformers for the task of Solar Irradiance Forecasting in the city of León, México.
  (September 2020 - December 2020)
- **Simulating confined random walks under different conditions**:
  For colloidal science interest: confinement due to optical tweezers, unidirectional external fields, harmonic and Brownian cages. As part of the Soft-Matter Laboratory in DCI-UG under Dr. Erick Sarmiento Gomez at the University of Guanajuato.
  (June 2020 - January 2021)

Qualifications

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