Cesar Macias Sanchez

+52 773 180 8616 |maciass.cesar.96@gmail.com
LinkedIn |GitHub (Personal) |GitHub (Organization) |Google Scholar

Professional Summary

AI Engineer, Data Scientist, NLP Specialist, & Research Scientist with 5+ years of experience in deep learning, multilingual NLP, and machine learning research. Proven expertise in dataset creation, model development, and AI competition leadership. Skilled in Python, PyTorch, TensorFlow, Transformers, and MLOps workflows. Published multiple peer-reviewed AI papers and experienced in full-stack development for real-world applications.

Technical Skills

Programming Languages: Python, JavaScript, Dart, LaTeX

Machine Learning & Data Science: PyTorch, Keras, TensorFlow, Transformers, Scikit-Learn, Pandas, Numpy, Scipy, Seaborn, Matplotlib, Spacy, Regex, Stanza

MLOps: GitHub Actions

Databases: SQL

Developer Tools: Git, Linux

Frameworks & Platforms: Flutter

Domain-Specific Software: AutoCAD, Staad Pro

Professional Experience

Full-Stack Developer — *IDCO de Hidalgo SA de CV* **Feb 2025 – Jul 2025**

- Spearheaded the development of a full-stack integrated management platform for the construction sector using Flutter.
- Engineered secure authentication and full CRUD functionality for users, projects, and staff allocation.
- Automated the generation of staff credentials from dynamic templates, significantly reducing manual processing time and eliminating data entry errors.
- Coordinated with stakeholders to define formal requirements, reducing rework due to ambiguity by 60%.
- Streamlined development by implementing bi-weekly Agile sprints, managing backlog refinement, user stories, and validation demos.

Structural Engineer — *IDCO de Hidalgo SA de CV* **2020-2021**

Structural Engineer — *Julio García Carbajal* **2019-2020**

Relevant Projects

Task Proposer Competition Organizer, HOMO-MEX @ IberLEF 2024

- Led the design of three complex hate speech detection subtasks: binary classification, multi-class classification, and token-level span identification.
- Developed a 20,000-text dataset and established strong baseline models using fine-tuned BETO and various BERT-based ensembles.
- Engineered and deployed the official Codabench competition platform, supporting approximately 60 international teams.

Task Proposer Competition Organizer, MiSonGyny @ IberLEF 2025

- Led a novel shared task to detect metaphorical and hidden misogyny in song lyrics, managing the project from March to May 2025.
- Developed the data pipeline, scraping and preprocessing raw text to create a final, annotated dataset of ˜2,500 Spanish song lyrics.
- Established baseline performance metrics using classical ML models (SVMs) and deployed the competition on the Codabench platform.

Education

PhD in Computer Science — *Instituto Politécnico Nacional (IPN), Computing Research Center (CIC)* **2023-2027**

Master in Computer Science — *Instituto Politécnico Nacional (IPN), Computing Research Center (CIC)* **2021-2023**

Bachelor of Civil Engineering — *Instituto Politécnico Nacional (IPN), ESIA Zacatenco* **2014-2019**

Publications Research

- Macias C, Soto M, Cardoso-Moreno MA, Calvo H. Modeling the depressive mind: An artificial intelligence approach to deciphering Beck's cognitive triad. *Journal of Intelligent & Fuzzy Systems.* 2024.
- Macias, C., Calvo, H., & Gambino, O. J. (2024). Dataset of controversial news posts in Spanish from the reader's perspective. *Data in Brief, 53.*
- Gómez-Adorno, H., Bel-Enguix, G., Calvo, H., et al. (2024). Overview of HOMO-MEX at IberLEF 2024: Hate Speech Detection Towards the Mexican Spanish speaking LGBT+ Population. *Procesamiento Del Lenguaje Natural, 73*, 393-405.

- *For a complete list of publications, please visit my Google Scholar profile.*

Leadership Professional Development

- **Founder Lead Organizer,** **Summer School in Computational Cognitive Sciences NLP** (2023-Present)
  - Proposed, founded, and organized this annual summer school, now in its third edition.
  - Designed and instructed a hands-on workshop on Natural Language Processing for all attendees.
  - Led the technical support team, resolving software and connectivity issues for attendees.
- **Lecturer**, "From Neuroscience to Computation"
- **Academic Committee**, Congreso CORE 2023
- **Conference Volunteer**, COMIA 2023, MICAI 2023, IEEE SSCI 2023, and SIMBig Conference
- **Teaching**, Natural Language Processing for Begginers (2023-present), Machine Learning (2024), From Neuroscience to computing (2024), Artificial Intelligence and Python (2025)

Certifications Diplomas

- **Diploma in Principles of Neuroscience** (200 hours), Consejo Mexicano de Neurociencias, March 2024
- **Certificate:** From Data to Insights with Google Cloud, *Coursera* (Sept 2021)
- **Certificate:** What is Data Science, *Coursera* (2020)

Languages

Spanish: Native

English: C1