Yarik Menchaca Resendiz

Stuttgart, Germany

YarikMenchacaR@gmail.com +49 151 47565395

LinkedIn Github Website

Profile

Final-year Ph.D. candidate (thesis submitted) specializing in Machine Learning, Natural Language Processing, and Large Language Models (LLMs). Proven track record in both academic and industry research, with practical experience in training, fine-tuning, optimizing, quantizing, prompting, and deploying ML models – particularly LLMs – in production environments. Skilled in teaching and mentoring, and motivated by solving real-world problems through research and applied solutions.

Skills

Coding	–	Python (8 yrs), SQL (4 yrs), Java (3 yrs), C++ (4 yrs), C# (3 yrs), Matlab (4 yrs), R (2 yrs), ...
ML Models	–	Training from scratch/fine-tuning/API: Gemini, GPT, LLaMA , T5, BART, Mistral, BERT, KNN, SVM, ...
Libraries	–	TensorFlow, Keras, PyTorch, Transformers, Scikit-learn, SciPy, NLTK, Spacy, Numpy, Pandas, ...
Visualization	–	Matplotlib, Seaborn, Pyplot, Graphviz, Shiny, Gephi, Power Bi, Jupyter Notebook.
Deployment	–	MLOPs practices, Docker, Microsoft Azure, Amazon Web Services, Git.

Professional Experience

02.2025 – Present – **Researcher – ZIPD, University of Trier, Germany**
– Led the design and development of a **Retrieval-Augmented Generation** (RAG) system with a user-friendly interface to summarize scientific texts for non-expert users.
– Defined a 37-point evaluation framework and collaborated across a multidisciplinary team.

04.2022 – 01.2025 – **Doctoral Researcher – University of Stuttgart & University of Bamberg, Germany**
– Researched **Conditional Text Generation** by using psychological emotion theories to guide **LLMs** outputs to improve their emotional intelligence and adaptability.
– Investigated **Multi-Objective Prompt Optimization** methods for **LLMs**.
– Conducted research prompt optimization methods for language models, incorporating **Reinforcement Learning** approaches.
– Mentored and supervised bachelor's and master's students in NLP courses and thesis.

06.2018 – 09.2021 – **Data Scientist – Kantar, México**
– Coordinated with commercial teams to align ML solutions with client needs, enhancing brand strategy and market impact.
– Built TextAI's engine a multilingual NLP system for **NER**, **Sentiment Analysis**, **Machine Translation** and **Topic Modeling**, then **led a team** to scale TextAI to 10+ languages.
– Developed **Machine Translation** models to convert specialized language (e.g., legal, medical) into user-friendly language for specific audiences.
– Developed a **Recommender System** to analyze purchase behavior, suggest relevant products, recommend substitutes, and organize items into categories.
– Developed **Model distillation** and **Quantization Techniques** to optimize inference performance and reduce cloud deployment costs.

02.2017 – 08.2017 – **Research Intern – University of Wolverhampton, UK**
– Developed a Machine Learning **Sentiment Analysis** application for Spanish texts.

Education

2021 – 2024 – **Ph.D. in Computational Linguistics – University of Stuttgart, Germany.**
Thesis title: *Controlling Affective Variables in Conditional Natural Language Generation*.

2015 – 2017 – **M.Sc. Computer Science – Instituto Politécnico Nacional, México.**
Final Grade: 9.8/10 (Graduated with Honors).
Thesis title: *Stress and Relaxation Strength Detection of Spanish Tweets*.

2011 – 2015 – **B.Eng. Computer Science – Instituto Politécnico Nacional, México.**
Final Grade: 8.3/10.
Thesis title: *Intelligent and Adaptive Alert System for Gas Leak Detection in Stoves with Real-Time Notifications to Email and Mobile Devices*.

Education (continued)

2008 – 2011 – **Technician in Computer Systems – CBTis 03, México.**
Final Grade: 9.0/10.
Hardware maintenance, network configuration, software development, system management, database.

Selected Research Publications

[1] Yarik Menchaca Resendiz and Roman Klinger. "Affective Natural Language Generation of Event Descriptions through Fine-grained Appraisal Conditions". In: *International Natural Language Generation Conference*. 2023. URL: https://aclanthology.org/2023.inlg-main.26.

[2] Yarik Menchaca Resendiz and Roman Klinger. "Emotion-Conditioned Text Generation through Automatic Prompt Optimization". In: *1st Workshop on Taming Large Language Models: Controllability in the era of Interactive Assistants! at INLG*. 2023. URL: https://aclanthology.org/2023.tllm-1.3.

[3] Yarik Menchaca Resendiz and Roman Klinger. "MOPO: Multi-Objective Prompt Optimization for Affective Text Generation". In: *Proceedings of the 31st International Conference on Computational Linguistics*. Ed. by Owen Rambow et al. Abu Dhabi, UAE: Association for Computational Linguistics, Jan. 2025, pp. 5588–5606. URL: https://aclanthology.org/2025.coling-main.375/.

[4] Yarik Menchaca Resendiz et al. "IMS_medicALY at #SMM4H 2024: Detecting Impacts of Outdoor Spaces on Social Anxiety with Data Augmented Ensembling". In: *Social Media Mining for Health Research and Applications (SMM4H 2024) Workshop and Shared Tasks at ACL*. 2024. URL: https://aclanthology.org/2024.smm4h-1.19.

[5] Yarik Menchaca Resendiz and Roman Klinger. *LLM-based Affective Text Generation Quality Based on Different Quantization Values*. 2025. arXiv: 2501.19317 [cs.CL]. URL: https://arxiv.org/abs/2501.19317.

[6] Johannes Schäfer et al. "Which Demographics do LLMs Default to During Annotation?" In: *Proceedings of the 63nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 2025. URL: https://arxiv.org/abs/2410.08820.

[7] Amelie Wuehrl and Yarik Menchaca Resendiz et al. "What Makes Medical Claims (Un)Verifiable? Analyzing Entity and Relation Properties for Fact Verification". In: *Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. 2024. URL: https://aclanthology.org/2024.eacl-long.124.

Selected Courses & Certificates

TensorFlow Developer
Introduction to Deep Learning in Python
Deep Learning in Python,
Advance Deep Learning with Keras in Python.
Building Deep Learning Applications with Keras 2.0
Introduction to TensorFlow for Artificial Intelligence
Natural Language Processing in TensorFlow
Convolutional Neural Networks in TensorFlow
Custom Model, Layers, and Loss Functions with TensorFlow
Artificial Intelligence Foundations: Thinking Machines
Big Data Processing, exploration with Scala and Apache Spark
Introduction to Recommender Systems
R Statistics Essential Training
Building Web Applications in R with Shiny Course
Data Analysis in R, the data.table Way
Sequences, Time Series and Prediction
Statistics Foundations
Design database MYSQL
Linux administrator
Web Design Fundamentals
Fundamentals Building Apps for Wearables
Python programmer

Languages

Spanish (Native), **English** (Bilingual), **French** (Advanced), **German** (Intermediate), **Italian** (Basic)

Miscellaneous Experience

Awards and Achievements

2021 – CONAHCyT Scholarship – Full funding for Ph.D in Computational Linguistics.

2015 – CONACyT Scholarship – Full funding for M.Sc. in Computer Science.

Thesis Supervision

2023 – **Master Thesis**, Plug and Play Domain Adaptation for Neural Machine Translation
Emīls Kadiķis