Yarik Menchaca Resendiz

Stuttgart, Germany
YarikMenchacaR@gmail.com +49 151 47565395
LinkedIn Github Website

Profile

I am a final-year Ph.D. candidate specializing in Natural Language Processing (NLP), Large Language Models (LLMs), and Artificial Intelligence (AI). With a strong background in both academic and industry research. I focus on developing AI/ML methods, prompt optimization for LLMs, and deploying NLP applications. I have also taught and mentored students. I enjoy applying my skills to solve real-world problems, whether it's through research or implementing AI-driven solutions.

Skills

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

Professional Experience

06.2018 – 09.2021 – **Data Scientist – Kantar, México**
Researched and developed AI/ML methodologies for data analysis and insights to help brands make data-driven decisions and enhance their marketplace performance.
— `Built TextAI's engine`, an in-house text mining tool for Named Entity Recognition, Sentiment Analysis, Machine Translation, Topic Modeling, and Topic Naming, supporting 10 languages (English, Spanish, German, French, Portuguese, Chinese, etc.).
— `Developed domain-specific to general language machine translation models` that convert specialized language (e.g., legal, medical, technical jargon) into general language, improving accessibility and understanding for wider audiences.
— `Created a machine learning model to analyze user purchase behavior`, identifying frequently bought products, suggest suitable substitutes, and categorizing items for personalized recommendations.
— `Developed model distillation and quantization techniques` to optimize inference performance and cloud deployment.

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.**
Focused on guiding and conditioning LLMs outputs using psychological emotion theories to improve their emotional intelligence and adaptability. Developed multi-objective prompt optimization techniques for LLMs and applied reinforcement learning through LLMs prompt optimization.

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*.

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

Research Publications

1. Yarik Menchaca Resendiz and Roman Klinger. “Affective Natural Language Generation of Event Descriptions through Fine-grained Appraisal Conditions”. In: *Proceedings of the 16th International Natural Language Generation Conference*. Ed. by C. Maria Keet, Hung-Yi Lee, and Sina Zarrieß. Prague, Czechia: Association for Computational Linguistics, Sept. 2023, pp. 375–387. DOI: 10.18653/v1/2023.inlg-main.26.
2. Yarik Menchaca Resendiz and Roman Klinger. “Emotion-Conditioned Text Generation through Automatic Prompt Optimization”. In: *Proceedings of the 1st Workshop on Taming Large Language Models: Controllability in the era of Interactive Assistants!* Ed. by Devamanyu Hazarika, Xiangru Robert Tang, and Di Jin. Prague, Czech Republic: Association for Computational Linguistics, Sept. 2023, pp. 24–30. URL: https://aclanthology.org/2023.tllm-1.3.
3. Yarik Menchaca Resendiz et al. “IMS_medicALY at #SMM4H 2024: Detecting Impacts of Outdoor Spaces on Social Anxiety with Data Augmented Ensembling”. In: *Proceedings of The 9th Social Media Mining for Health Research and Applications (SMM4H 2024) Workshop and Shared Tasks*. Ed. by Dongfang Xu and Graciela Gonzalez-Hernandez. Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 83–87. URL: https://aclanthology.org/2024.smm4h-1.19.
4. Amelie Wuehrl et al. “What Makes Medical Claims (Un)Verifiable? Analyzing Entity and Relation Properties for Fact Verification”. In: *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*. Ed. by Yvette Graham and Matthew Purver. St. Julian’s, Malta: Association for Computational Linguistics, Mar. 2024, pp. 2046–2058. URL: https://aclanthology.org/2024.eacl-long.124.

Courses & Certificates

TensorFlow Developer
Introduction to Deep Learning in Python
Deep Learning in Python,
Advance Deep Learning with Keras in Python.
Learning TensorFlow with JavaScript
Building Deep Learning Applications with Keras 2.0
Introduction to TensorFlow for Artificial Intelligence
Introduction to Deep Learning with OpenCV
Sequences, Time Series and Prediction
Natural Language Processing in TensorFlow
Convolutional Neural Networks in TensorFlow
Custom Model, Layers, and Loss Functions with TensorFlow
Natural Language Processing
Machine Learning for Marketing: Essential Training
Artificial Intelligence Foundations: Thinking Machines
Introduction to Recommender Systems
Big Data Processing, exploration with Scala and Apache Spark
R Statistics Essential Training
Building Web Applications in R with Shiny Course
Data Analysis in R, the data.table Way
Statistics Foundations 1
Statistics Foundations 2
Statistics Foundations 3
Design database MYSQL
Linux administrator
Web Design Fundamentals
Fundamentals Building Apps for Wearables
Python programmer
Financial Markets

Languages

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

Miscellaneous Experience

Awards and Achievements

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