

Enrique David Guzmán Ramírez

+1 (437) 973-8963 | david.guzman@alumni.utoronto.ca | davidguzmanr.github.io | github.com/davidguzmanr
linkedin.com/in/enrique-david-guzman-ramirez | David Guzmán

Work Experience

- J.D. Power

Mexico City - Remote

Data Engineer

March 2024 - Present

 - Developed pipelines in PySpark to prepare ingest, transform and load the data of multiple clients to different projects and teams.
 - Implemented data quality and validation checks to ensure accuracy and integrity of data throughout the ETL process, reducing data-related issues by 30%.
- National Research Council Canada

Toronto, Canada

Machine Learning Research Intern

May 2023 - December 2023

 - Collaborated in the *Speech Generation for Indigenous Language Education* project to develop state-of-the-art speech synthesis systems and techniques with Deep Learning for Indigenous languages in Canada.
 - Contributed to the development of *EveryVoice TTS Toolkit*, a tool for creating text-to-speech systems focused on low-resource languages, achieving robust performance with up to 90% less data.
- Grupo Salinas

Mexico City

Data Scientist

November 2021 - July 2022

 - Developed a collaborative filtering recommender system for the homepage of the E-commerce area with PyTorch, which increased precision at 10 by 50%.
 - Developed a content-based recommender system and association rules for the product detail pages of the E-commerce area with scikit-learn and MLxtend, automating the original 2 week manual process.
 - Deployed and monitored ML models as APIs for various use cases with AWS SageMaker.

Education

- MSc in Applied Computing

3.93/4 GPA

University of Toronto, Department of Computer Science

Sep 2022 – June 2024
- BSc in Data Science

9.60/10 GPA

National Autonomous University of Mexico, Research Institute in Applied Mathematics

August 2019 - June 2021
- BSc in Physics

9.54/10 GPA

National Autonomous University of Mexico, Faculty of Science

August 2015 - June 2020

Technical skills

Certifications	AWS Certified Machine Learning - Specialty (MLS-C01)
Programming Languages	Python, Julia, R, Fortran, SQL.
Machine and Deep Learning	Scikit-learn, XGBoost, PyTorch, Keras, TensorFlow, PySpark MLlib.
Data Analysis and Manipulation	Numpy, SciPy, Pandas, Matplotlib, Seaborn, Bokeh, Plotly.
NLP	NLTK, spaCy, Hugging Face, LangChain.
Cloud services	AWS, Azure.
Databases	PostgreSQL, Redshift, MongoDB, Redis.
Miscellaneous	Linux, Git, LaTeX, Slurm, REST APIs, Scrum.
Languages	Spanish (Native), English (Proficient), French (Intermediate)

Publications

- Steve Bakos, Félix Gaschi, **David Guzmán**, Riddhi More, Kelly Chutong Li, En-Shiun Annie Lee “ALIGNFREEZE: Navigating the Impact of Realignment on the Layers of Multilingual Models Across Diverse Languages.” Presented at the 2025 Annual Conference of the North American Chapter of the Association for Computational Linguistics. 2025. Available at ACL Anthology: <https://aclanthology.org/2025.naacl-short.48>.
- Aidan Pine, Erica Cooper, **David Guzmán**, Eric Joanis, Anna Kazantseva, Ross Krekoski, Roland Kuhn, Samuel Larkin, Patrick Littell, Delaney Lothian, Akwiratékhá’ Martin, Korin Richmond, Marc Tessier, Cassia Valentini-Botinhao, Dan

Wells, Junichi Yamagishi. “*Speech Generation for Indigenous Language Education*.” 2024. Computer Speech & Language. Available at: <https://doi.org/10.1016/j.csl.2024.101723>.

- Tong Su, Xin Peng, Sarubi Thillainathan, **David Guzmán**, Surangika Ranathunga, En-Shiun Annie Lee. “*Unlocking Parameter-Efficient Fine-Tuning for Low-Resource Language Translation*.” Presented at the 2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics. 2024. Available at ACL Anthology: <https://aclanthology.org/2024.findings-naacl.263>.

Awards

2022 **Vector Scholarship in Artificial Intelligence Recipient 2022-2023.**

2019 **Support Program for Research and Technological Innovation Projects (PAPIIT, Mexico).**