# SUMMARY

Graduate Research Assistant with 3 years of experience in Data Science, Machine Learning, and Natural Language Processing. Proficient in training LLMs, and Machine Learning algorithms. Strong communication skills through professional engagements. Proven ability in programming and large-scale data analytics to derive valuable insights. Collaborative team member with a track record of successful collaboration on international research projects.

# RESEARCH EXPERIENCE

|  |  |
| --- | --- |
| **Virginia Modeling, Analysis, & Simulation Center (VMASC)** | Suffolk, VA |
| *Graduate Research Assistant* | Sep 2022 – Present |
| • Developed a conversational Chatbot expert in a custom domain enhancing answer fidelity by 90% by fine-tuning the Large Language Model (LLM) Llama2 7B on a dataset of 330 news articles.  • Trained and fine-tuned Machine Learning classification models (RFC, ANN, and BERT) to identify types of frustration in textual data, achieving an average accuracy of 70%.  • Designed and maintained detailed documentation of models, algorithms, and reports on GitHub repositories.  • Presented poster at a research conference in Pittsburgh, and attended others in Arlington, VA, and Oxford, OH. | |
| **Old Dominion University** | Norfolk, VA (Remote) |
| *Project Consultant (Social Media Data Analyst)* | Jul 2021 – Jan 2022 |
| • Geolocated, filtered, and estimated the tone of 7.3 million Tweets through fine-tuning fifteen NLP models (spaCy, BERT, RoBERTa) and Statistical Analysis. Models were highly reliable, with a mean accuracy of 72%.  • Extracted, cleaned, and pre-processed 15.5 million Tweets in multiple languages with Twitter’s API and Python.  • Created dashboards with Tableau and Google Data Studio to represent migration data.  • Collaborated with teams from Norway, Colombia, Greece, and USA on a $1.4M USD Minerva research initiative. | |
| **Universidad del Norte** | Barranquilla, Colombia |
| *Undergraduate Research Assistant* | Nov 2020 – Jun 2021 |
| • Trained a Natural Language Understanding (NLU) model with IBM Watson to semi-automate the identification of actors, factors, and relationships from news articles, achieving a final accuracy of 70%.  • Gathered, filtered, and cleaned geo-tagged 4,680 Tweets on migration using Twitter's API and Python.  • Conducted a lexicon-based Sentiment Analysis of Tweets to spot xenophobic trends. | |

# EDUCATION

|  |  |
| --- | --- |
| **Old Dominion University** | Norfolk, VA |
| M.Sc. Modeling and Simulation Engineering - GPA: 3.7 | Aug 2022 – Dec 2024 |
| **Correlation One** | New York, NY (Remote) |
| Data Science for All (Bootcamp) | Mar 2022 – Jul 2022 |
| **Universidad del Norte** | Barranquilla, Colombia |
| B.Sc. Industrial Engineering - GPA: 3.6 | Jan 2017 – Mar 2022 |

# PUBLICATIONS

• Joseph Martínez, Brian Llinas, Jhon G. Botello, Jose J. Padilla, and Erika Frydenlund. “Enhancing GPT-3.5’s Proficiency in NetLogo through Few-Shot Prompting and Retrieval-Augmented Generation” 2024 Winter Simulation Conference (WSC) IEEE (2024) [Under Revision].

• Erika Frydenlund, Joseph Martínez, Jose J. Padilla, Katherine Palacio, and David Shuttleworth. “Modeler in a Box: How Can Large Language Models Aid in the Simulation Modeling Process?” SIMULATION (2024).

• Lynch, Christopher J., Erik J. Jensen, Madison H. Munro, Virginia Zamponi, Joseph Martínez, Kevin O'Brien, Brandon Feldhaus, Katherine Smith, Ann Marie Reinhold, and Ross Gore. “GPT-4 Generated Narratives of Life Events using a Structured Narrative Prompt: A Validation Study.” (2024)

• Joseph Martínez, Melissa Miller-Felton, Jose J. Padilla, Erika Frydenlund. “Behind Derogatory Terming for Venezuelan Migrants in Colombia: Xenophobia and Sexism Identification with Twitter Data and NLP” (2023)

• Lauren Guerra, Joseph Martínez, and Francisco Pautt. “Design of a User Interface based on Information Engineering and Web Programming for the FUVADIS customer service process.” Universidad del Norte (2021).

# SKILLS

|  |  |
| --- | --- |
| NLP: scikit-learn, Transformers, spaCy, Hugging Face, OpenAI GPT-4 | Deep Learning: Pytorch, Keras, ANN |
| Machine Learning: Classification, Regression, Clustering, PCA, NumPy | Data Analysis: Python, Pandas, R, SQL |
| Data Visualization: Tableau, Power BI, R Shiny, HTML, Excel | DevOps: Git, GitHub, Docker |

# AWARDS AND ACHIEVEMENTS

|  |  |
| --- | --- |
| **SBP-BRIMS 2023 conference scholarship** |  |
| SBP-BRiMS Conference chair | Sep 2022 |
| **Outstanding results in the Saber Pro 2021 test** |  |
| Universidad del Norte  · Best overall institutional score  · Best overall score in Industrial Engineering  · Perfect score in Generic Competencies | Nov 2022 |
| **Graduate Research Assistantship** |  |
| Virginia Modeling, Analysis and Simulation Center (VMASC) | Jun 2022 |
| **Undergraduate Honorable Mention** |  |
| Universidad del Norte | Mar 2022 |
| **Andrés Bello Distinction** |  |
| Colombian Ministry of Education  · Distinction to outstanding students in the ICFES Saber 11 State Exam for high school | Nov 2016 |

# CERTIFICATIONS

|  |  |
| --- | --- |
| **Responsible Conduct of Research for Engineers** |  |
| CITI Program | Jun 2023 |
| **Social and Behavioral Research** |  |
| CITI Program | Aug 2022 |
| **DS4A / Colombia 6.0** |  |
| Correlation One | Jul 2022 |
| **Excel (Advanced level)** |  |
| Coursera | Jul 2020 |