# SUMMARY

Research Assistant with 3 years of experience in Natural Language Processing, and Machine Learning. Proficient in training LLMs, and Machine Learning (ML) algorithms. Strong communication skills through academic and professional engagements. Proven ability in 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 |
| • Created a custom conversational LLM (akin to ChatGPT) fine-tuning Llama2 with a dataset of 330 news about Venezuelan migration. The final model performs better when asked questions about the topic.  • Trained Machine Learning models (KNN, RFC) and fine-tuned NLP models (BERT) to classify frustration types.  • 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) |
| *Social Media Data Analyst (Consultant)* | 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 on the topic of Venezuelan migration in Colombia.  • 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 from Colombian entities.  • Collaborated with teams from Norway, Colombia, Greece, and the US on a $1.4M Minerva research initiative. | |
| **Universidad del Norte** | Barranquilla, Colombia |
| *Social Media Data Analyst (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 in Colombia and Greece, achieving 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 towards Venezuelans.  • Determined key actors of a Twitter conversation about migration through Social Network Analysis with Gephi. | |

# EDUCATION

|  |  |
| --- | --- |
| **Old Dominion University** | Norfolk, VA |
| M.Sc. Modeling and Simulation Engineering - GPA: 3.6 | Aug 2022 – Aug 2024 (expected) |
| **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

• 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, (2023) [Under revision].

• Joseph Martínez, Melissa Miller-Felton, Jose J. Padilla, Erika Frydenlund, and Katherine Palacio. “Behind Derogatory Terming for Venezuelan Migrants in Colombia: Xenophobia and Sexism Identification with Twitter Data and NLP." SBP-BRiMS 2023, (2023) [Poster session].

• 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

|  |  |
| --- | --- |
| Data Visualization: Tableau, Power BI, Shiny, HTML, ggplot2, plotly | Data Analysis: Python, R, SQL |
| Machine Learning: Classification, Regression, Clustering, PCA | Deep Learning: Pytorch, Keras, ANN |
| NLP: scikit-learn, spaCy, Transformers, Hugging Face, OpenAI GPT-4 | Version control: Git/GitHub |
| **Languages:** Spanish, English, German (Intermediate) |  |

# 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 |