# Paper draft (old)



Exploratory analysis of Venezuelan migration twitter data through manual and automated approaches

Jose Padilla1, Joseph Mart ́ınez1\*, Katherine Palacio2, Erika Frydenlund1  
1 Virginia Modeling Analysis and Simulation Center, Old Dominion University, Suffolk,  
Virginia, United States  
2 Industrial Engineering Department, Universidad del Norte, Barranquilla, Atlántico,  
Colombia  
\* [jmart130@odu.edu](mailto:jmart130@odu.edu)

**Useful links:**

[**https://journals.plos.org/plosone/s/latex**](https://journals.plos.org/plosone/s/latex)

[Hyperparameter Search with Transformers and Ray Tune (huggingface.co)](https://huggingface.co/blog/ray-tune)

[hyperparameters paper]: Tune: A Research Platform for Distributed Model Selection and Training

[3.3. Metrics and scoring: quantifying the quality of predictions — scikit-learn 1.1.2 documentation](https://scikit-learn.org/stable/modules/model_evaluation.html#multiclass-and-multilabel-classification)

[Comprehensive Guide on Multiclass Classification Metrics | Towards Data Science](https://towardsdatascience.com/comprehensive-guide-on-multiclass-classification-metrics-af94cfb83fbd?gi=e5a19c16a368)

[nlptown/bert-base-multilingual-uncased-sentiment · Hugging Face](https://huggingface.co/nlptown/bert-base-multilingual-uncased-sentiment?text=I+like+you.+I+love+you)

<https://huggingface.co/dccuchile/bert-base-spanish-wwm-cased>

# Report 10/20/2022

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | **Node** | | user | | tweet | | hashtag | | url | | media | | |  |  |  | | --- | --- | --- | | **Edge** | | | | **Source node** | **Target node** | **label** | | tweet | hashtag | Hash\_hashtag | | tweet | link | Has\_link | | tweet | media | Has\_media | | tweet | user | Has\_mention | | user | tweet | quoted | | user | user | Quotes\_from | | user | tweet | tweets | | user | tweet | retweets | | user | tweet | replied\_to | |

Only in the 7M sample

Graphical user interface, text

Description automatically generated



**TODO:**

* EXTRACT ALL THE TWEETS FROM 2014 TO 2022-6 ♦
* FILTER BY THE TEXT OF THE TWEET, NOT ALL OF THEM ARE VALID