

Bridging Gaps in Hate Speech Detection

Meta-Collections and Benchmarks for Low-Resource Iberian Languages



Funded by
the European Union



UK Research
and Innovation

Usage/License



The models are available via the Hugging Face Transformers library and are licensed under Apache 2.0. Users should be aware of potential biases and possible misclassifications, as performance may vary across domains. Code is also available on GitHub under the same license.

Download



The models can be accessed and downloaded from its HuggingFace collection page: <https://huggingface.co/collections/irlab-udc/metahate/>. The code can be accessed via its GitHub repository <https://github.com/HYBRIDS-MSCA/mlhate>.

Researchers



- Paloma Piot
- José Ramon Pichel Campos
- Javier Parapar.

These researchers are affiliated with the Information Retrieval Lab at the University of A Coruña and University of Santiago de Compostela.

Summary

This research presents a standardized meta-collection of European Spanish hate speech datasets, extended through translation into European Portuguese and two distinct linguistic varieties of Galician. This aligned multilingual corpus establishes new performance benchmarks and provides a foundation for cross-lingual hate speech detection across underrepresented Iberian languages.

What We Offer

- Iberian Hate Speech Resources: Curated dataset covering Spanish, Portuguese, and Galician.
- Baseline Performance Metrics: Benchmarks for current LLMs as reference for future models.
- Cross-Lingual Analysis Framework: Tools and data to study transfer across related Iberian languages.

Key Features

- Standardized Meta-Collection: Integration of European Spanish hate speech resources into a unified format with consistent labels and metadata.
- Multilingual Aligned Corpora: Extension to European Portuguese and two Galician standards for cross-lingual research.
- Variety-Aware Design: Inclusion of multiple linguistic varieties to capture regional and dialectal differences.
- Multi-Setting Evaluation: Benchmarking of state-of-the-art LLMs across zero-shot, few-shot, and fine-tuning settings.

Collaboration Objectives

- Low-Resource Language Expansion: Extend the meta-collection to other underrepresented European languages and dialects.
- Cross-Lingual Model Optimization: Improve performance using linguistic proximity and variety-aware data.
- Benchmarking Standardization: Develop consistent evaluation protocols for multilingual hate speech detection.

