UNIVERSIDADE DE ÉVORA

CITIUS — Centro Singular de Investigación en Tecnologías Inteligentes

USC — UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

CT-BERT-PRCT

CT-BERT fine-tuned for Population Replacement Conspiracy Theory Detection

Usage/License

MIT License

Download

The model can be accessed and downloaded from its Hugging Face page:

Researchers

- Erik Bran Marino
- Renata Vieira
- Davide Bassi
- Ana Sofia Ribeiro
- Suso Baleato

Summary

A specialized BERT model fine-tuned to detect Population Replacement Conspiracy Theory (PRCT) content across social media platforms. The model demonstrates good performance in identifying both explicit and implicit PRCT narratives, with decent cross-platform and multilingual generalization capabilities.

Key Features

- Cross-platform detection capabilities
- Multilingual support
- Good performance on implicit content
- Balanced precision and recall metrics
- Efficient computational performance
- Easy integration with existing content moderation systems

What We Offer

A ready-to-use model for detecting PRCT content, trained on a dataset of Youtube English comments and validated across different platforms and languages. The model achieves 83.8% accuracy on YouTube data and maintains 71.9% accuracy in Telegram cross-platform and multilingual scenario.

Collaboration Objectives

- Integration with content moderation systems
- Further fine-tuning for additional languages, narratives and platforms
- Development of hybrid approaches combining BERT and other detection methods
- Collaboration with researchers studying online extremism and conspiracy theories

hybrids

HYBRIDS Website

🌐 www.hybridsproject.eu

Get in Touch

✉ erik.marino@uevora.pt