# DATASUMMIT AT MULTIPRIDE 2026:

## CONTEXT-AWARE MULTILINGUAL DETECTION OF SLUR RECLAMATION IN LGBTQ+ CONTEXTS

Federico Dingeo, Marco Viviani

Università degli Studi di Milano-Bicocca, Dipartimento di Informatica, Sistemistica e Comunicazione (DISCo)

## Introduction

- **Context:** LGBTQIA+ semantic reclamation - historically offensive slurs as terms of empowerment and identity.
- **Problem:** Standard NLP systems often misclassify reclaimed slurs as hate speech, failing to balance toxicity detection with freedom of expression.
- **Examples:**
  - Reclaimed: In quanto disabile e frocia questi sono i miei PrideMonths.
  - Not reclaimed: Ecco, adesso pensate all'utero in affitto ed al male che fate al bambino branco di finocchi arcobaleno.

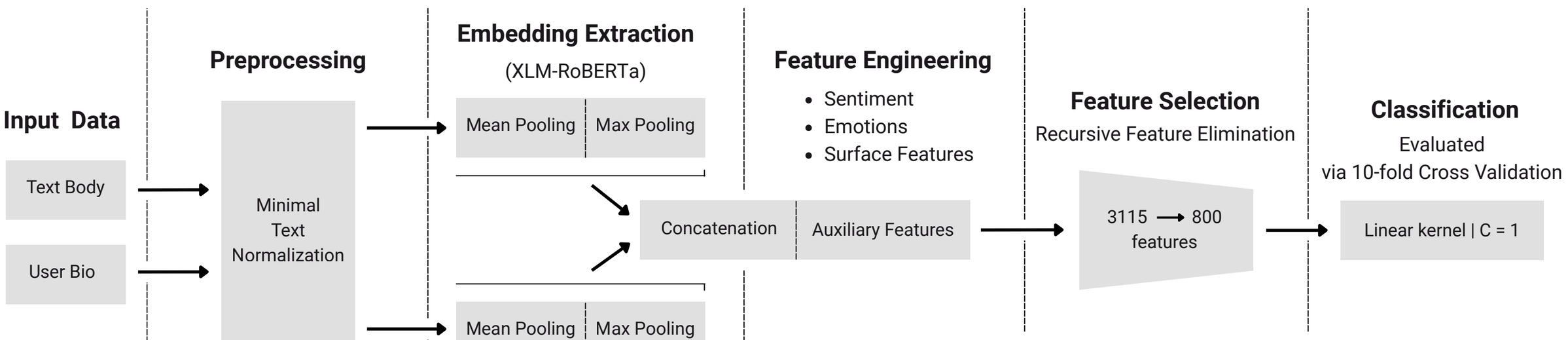
## Multipride Tasks

The MultiPRIDE shared task proposes two binary classification tracks:

- **Task A (Textual Content):** Classification based solely on the message text.
- **Task B (Contextual Content):** Classification integrating the text with user-level metadata, specifically the author's biography.

**Our Contribution:** We specifically focused on **Task B**, developing a **multilingual system** that leverages texts, user biographies, and engineered features to better contextualize linguistic usage.

## Methodology




**Technical Details**

- For Sentiment Analysis: *Twitter-RoBERTa Sentiment*
- For Emotion Analysis: *DistilRoBERTa Emotion* & *RoBERTa-Base GoEmotions*
- Surface Features:
  - Text & User Bio token count (via *TweetTokenizer*)
  - Presence of LGBTQIA+ emojis in the bio (via *TweetTokenizer*)
  - Dominant pronoun within the text (via *Stanza POS tagging)*

## Results

Final classification results obtained on the official test set.

Italian

| Category | Precision | Recall | F1-Score |
|---|---|---|---|
| Class 0 | 0.9117 | 0.9163 | 0.9140 |
| Class 1 | 0.6397 | 0.6258 | 0.6327 |
| Macro Avg | 0.7757 | 0.7711 | 0.7733 |

Spanish

| Category | Precision | Recall | F1-Score |
|---|---|---|---|
| Class 0 | 0.9050 | 0.8249 | 0.8631 |
| Class 1 | 0.3409 | 0.5113 | 0.4090 |
| Macro Avg | 0.6229 | 0.6681 | 0.6361 |

**Language Variance:** The system achieved solid performance on Italian texts, whereas Spanish ones showed a substantial decline, primarily due to the model's difficulty in recognizing the reclaimed minority class (Class 1).

**Generalization Gap:** Across both languages, test set results were significantly lower than internal validation estimates, indicating that the model overfitted the training data.

## Limitations & Future Works

**Limitations:**

- **Optimization Bias:** Focusing too much on maximizing the Macro F1 score during training increased the risk of overfitting.
- **Contextual Misalignment:** Adding English texts, which had no user biographies available, introduced noise and negatively affected the linear classifier.

**Future Works:**

- **Alternative Architectures:** Test other multilingual models to find embeddings that work more consistently across all languages.
- **Zero-Shot NLP:** Use prompt-based approaches to overcome the lack of data and classify texts without needing many labeled examples.

[1] C. Ferrando, L. Draetta, M. Madeddu, M. Sosto, V. Patti, P. Rosso, C. Bosco, J. Mata, E. Gualda, Multipride at Evalita 2026: Overview of the multilingual automatic detection of slur reclamation in the lgbtq+ context task, in: Proceedings of the Ninth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2026), CEUR.org, Bari, Italy, 2026.

2] F. Cutugno, A. Miaschi, A. P. Aprosio, G. Rambelli, L. Siciliani, M. A. Stranisci, Evalita 2026: Overview of the 9th evaluation campaign of natural language processing and speech tools for Italian, in: Proceedings of the Ninth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2026), CEUR.org, Bari, Italy, 2026.