

# Erratum: Cross-Lingual Metaphor Detection for Low- to High-Resource Languages

**Anna Hülsing**

University of Hildesheim  
anna.huelsing@uni-hildesheim.de

**Sabine Schulte im Walde**

University of Stuttgart  
schulte@ims.uni-stuttgart.de

## 1 Error and Corrected Results

In the paper “Cross-Lingual Metaphor Detection for Low- to High-Resource Languages”, the F1-scores describe the models’ performance in detecting non-metaphorical expressions instead of describing the performance in detecting metaphorical expressions, which we actually intended. Hence, we repeated the experiments to see how the models perform in terms of the intended metric. The results for using the basic training dataset with default hyperparameters are shown in Table 1. The main insight we drew from the original experiments still holds: neural cross-lingual methods in general outperform the non-neural model for languages with relatively large (German) and small amounts of pretraining data (Latin). For Russian, zero-shot mBERT and the Random Forest classifier perform similarly. The results for Latin, however, are generally poorer than in the original experiments, which highlights the difficulty of dealing with this kind of low-resource language. Please contact us if you have questions concerning the corrected results of other experiments presented in the paper.

	basic training dataset		
	ru	ge	la
<b>baseline</b>	66.7	66.7	66.7
<b>mB0</b>	83.2 $\pm$ 1.2	<b>77.1</b> $\pm$ 2.0	65.5 $\pm$ 3.8
<b>mB20</b>	68.1 $\pm$ 31.0	26.4 $\pm$ 13.1	0.0 $\pm$ 0.0
<b>MAD-X</b>	81.6 $\pm$ 3.4	75.4 $\pm$ 1.9	<b>67.4</b> $\pm$ 5.3
<b>RF</b>	<b>83.3</b> $\pm$ 0.2	65.0 $\pm$ 0.3	54.0 $\pm$ 1.5

Table 1: Mean F1-scores for verbal MD across three runs with different seeds ( $\pm$  SD) for default hyperparameters with the basic training dataset and across our target languages Russian (ru), German (ge) and Latin (la).