

APPENDIX

Leveraging Machine Translation for cross-lingual fine-grained cyberbullying classification among pre-adolescents

Kanishk Verma^a, Maja Popović^b, Alex Poulis^c, Yelena Cherkasova^d, Cathal Ó hÓbáin^e, Angela Mazzone^f, Tijana Milosevic^g, and Brian Davis^h

ADAPT SFI Research Centre, Dublin City University ^{a-b-e-h}
Dublin, Ireland

{kanishk.verma, maja.popovic, cathalohobain, brian.davis}@adaptcentre.ie

DCU Anti-Bullying Centre, Dublin City University ^{a-f-g}

Dublin, Ireland

{angela.mazzone, tijana.milosevc}@dcu.ie

TransPerfect DataForce ^c

apoulis@translations.com

G3 Translate, Strategic Partner of Transperfect ^d

ycherkasova@g3translate.com

Abstract

1. HT Dataset Binary Classification Results

Refer table 1 and 2 for binary classification results on raw-original annotations and scenario-C annotations by domain experts respectively.

2. MT Dataset Binary Classification Results

Refer table 3 and 4 for binary classification results on raw-original annotations and scenario-C annotations by domain experts respectively.

Table 1. HT Dataset Binary Classification Results with original labels

Model	Emojis	Embeddings	Test-Set		
			F1-hate	F1-no hate	F1-macro avg
SVM	Yes	TF-IDF	0.71	0.61	0.66
		FastText	0.75	0.62	0.68
	No	TF-IDF	0.71	0.61	0.66
		FastText	0.75	0.62	0.68
Decision Tree	Yes	TF-IDF	0.72	0.58	0.65
		FastText	0.65	0.52	0.59
	No	TF-IDF	0.72	0.58	0.66
		FastText	0.68	0.57	0.62
Random Forest	Yes	TF-IDF	0.74	0.57	0.66
		FastText	0.78	0.65	0.71
	No	TF-IDF	0.76	0.6	0.68
		FastText	0.78	0.63	0.7
XGBOOST	Yes	TF-IDF	0.74	0.59	0.67
		FastText	0.76	0.65	0.71
	No	TF-IDF	0.74	0.54	0.66
		FastText	0.74	0.63	0.69
GRU	Yes	FastText	0.77	0.69	0.73
	No	FastText	0.75	0.66	0.7
Bi-LSTM	Yes	FastText	0.75	0.64	0.7
	No	FastText	0.78	0.63	0.7
CNN	Yes	FastText	0.74	0.61	0.67
	No	FastText	0.66	0.62	0.64
BERT	Yes	bert-base-uncased	0.84	0.81	0.83
	No	bert-base-uncased	0.79	0.72	0.75

Table 2. HT Dataset Binary Classification Results with additional labels

Model	Emojis	Embeddings	Test-Set		
			F1-hate	F1-no hate	F1-macro avg
SVM	Yes	TF-IDF	0.74	0.64	0.7
		FastText	0.74	0.64	0.69
	No	TF-IDF	0.74	0.64	0.7
		FastText	0.73	0.63	0.68
Decision Tree	Yes	TF-IDF	0.66	0.54	0.60
		FastText	0.69	0.55	0.62
	No	TF-IDF	0.72	0.59	0.65
		FastText	0.61	0.48	0.54
Random Forest	Yes	TF-IDF	0.75	0.59	0.67
		FastText	0.8	0.65	0.74
	No	TF-IDF	0.74	0.61	0.67
		FastText	0.81	0.68	0.74
XGBOOST	Yes	TF-IDF	0.7	0.59	0.65
		FastText	0.74	0.67	0.71
	No	TF-IDF	0.7	0.59	0.65
		FastText	0.82	0.77	0.79
GRU	Yes	FastText	0.83	0.79	0.81
	No		0.84	0.75	0.79
Bi-LSTM	Yes	FastText	0.85	0.74	0.79
	No		0.79	0.75	0.77
CNN	Yes	FastText	0.82	0.77	0.79
	No		0.81	0.76	0.79
BERT	Yes	bert-base-uncased	0.86	0.79	0.82
	No		0.83	0.79	0.81

Table 3. MT Dataset Binary Classification Results with original labels

Model	Emojis	Embeddings	Test-Set		
			F1-hate	F1-no hate	F1-macro avg
SVM	Yes	TF-IDF	0.71	0.6	0.66
		FastText	0.72	0.57	0.65
	No	TF-IDF	0.7	0.59	0.65
		FastText	0.73	0.58	0.65
Decision Tree	Yes	TF-IDF	0.68	0.56	0.62
		FastText	0.65	0.55	0.6
	No	TF-IDF	0.67	0.52	0.6
		FastText	0.62	0.51	0.57
Random Forest	Yes	TF-IDF	0.75	0.56	0.65
		FastText	0.73	0.56	0.64
	No	TF-IDF	0.73	0.54	0.63
		FastText	0.78	0.63	0.7
XGBOOST	Yes	TF-IDF	0.69	0.53	0.61
		FastText	0.72	0.58	0.65
	No	TF-IDF	0.71	0.58	0.65
		FastText	0.73	0.61	0.67
GRU	Yes	FastText	0.76	0.68	0.72
	No	FastText	0.73	0.68	0.7
Bi-LSTM	Yes	FastText	0.72	0.55	0.64
	No	FastText	0.64	0.7	0.67
CNN	Yes	FastText	0.7	0.67	0.68
	No	FastText	0.7	0.57	0.63
BERT	Yes	bert-base-uncased	0.78	0.74	0.76
	No	bert-base-uncased	0.75	0.67	0.71

Table 4. MT Dataset Binary Classification Results with additional labels

Model	Emojis	Embeddings	Test-Set		
			F1-hate	F1-no hate	F1-macro avg
SVM	Yes	TF-IDF	0.68	0.51	0.6
		FastText	0.72	0.61	0.66
	No	TF-IDF	0.68	0.51	0.6
		FastText	0.68	0.57	0.62
Decision Tree	Yes	TF-IDF	0.69	0.54	0.62
		FastText	0.55	0.45	0.51
	No	TF-IDF	0.66	0.45	0.55
		FastText	0.65	0.42	0.53
Random Forest	Yes	TF-IDF	0.71	0.49	0.6
		FastText	0.75	0.63	0.69
	No	TF-IDF	0.66	0.44	0.55
		FastText	0.81	0.68	0.74
XGBOOST	Yes	TF-IDF	0.67	0.47	0.57
		FastText	0.69	0.57	0.63
	No	TF-IDF	0.7	0.57	0.64
		FastText	0.69	0.66	0.67
GRU	Yes	FastText	0.77	0.64	0.705
	No		0.74	0.65	0.7
Bi-LSTM	Yes	FastText	0.79	0.7	0.74
	No		0.58	0.63	0.61
CNN	Yes	FastText	0.77	0.67	0.72
	No		0.76	0.67	0.71
BERT	Yes	bert-base-uncased	0.76	0.74	0.75
	No		0.71	0.66	0.68