

# 1 Appendix A. Detection Accuracy with the Two Datasets

Models	Embeddings	Accuracy (%)	F1 (SC)	F1 (PC)	F1 (SI)	Training Time (s)
Baseline1-SVM		44.39	0	0	0.61	0.97
Baseline2-Random Forest		53.94	<b>0.48</b>	0.48	0.61	1.36
CNN	Char	49.85	0.25	0.49	0.56	2.41
	Word	53.03	0.33	0.51	0.60	2.54
	Doc	44.09	0	0	0.61	4.94
	Char+Word	53.64	0.34	0.51	0.60	2.75
	Char+Doc	51.82	0.38	0.47	0.60	2.55
	Word+Doc	51.67	0.36	0.49	0.58	2.90
	Char+Word+Doc	52.27	0.39	0.48	0.59	2.72
LSTM	Char	47.12	0.27	0.39	0.57	24.75
	Word	52.27	0.26	0.51	0.60	25.34
	Doc	43.64	0	0.01	0.61	54.15
	Char+Word	54.24	0.42	0.48	0.63	24.79
	Char+Doc	49.70	0.35	0.47	0.57	25.01
	Word+Doc	51.82	0.39	0.50	0.57	24.82
	Char+Word+Doc	52.27	0.43	0.50	0.57	25.28
Bi-LSTM	Char	51.82	0.38	0.45	0.61	49.72
	Word	56.67	0.44	0.56	0.62	48.68
	Doc	43.94	0	0.06	0.60	110.52
	Char+Word	56.67	0.40	0.50	<b>0.66</b>	49.73
	Char+Doc	51.21	0.40	0.49	0.56	49.55
	Word+Doc	50.61	0.37	0.49	0.57	49.34
	Char+Word+Doc	53.94	0.41	0.49	0.62	49.43
Transformer Encoder	Char	43.78	0.26	0.37	0.54	11.57
	Word	49.85	0.31	0.47	0.57	11.71
	Doc	48.03	0.29	0.49	0.53	12.36
	Char+Word	46.06	0.25	0.46	0.51	12.02
	Char+Doc	48.48	0.36	0.42	0.57	12.18
	Word+Doc	51.52	0.37	0.49	0.58	10.56
	Char+Word+Doc	49.70	0.33	0.44	0.59	12.32
BERT		<b>57.58</b>	0.40	<b>0.57</b>	0.64	24.07
RoBERTa		55.45	0	0.37	0.52	0.63
AlBERT		47.27	0	0.04	0.48	0.55
Electra		51.36	0.29	0.49	0.60	22.20
GPT2		52.12	0	0.48	0.63	14.23

Table 1: Detection Accuracy with Suicide-related Tweets Dataset (SC: Strongly Concerning, PC: Possibly Concerning, SI: Safe to Ignore)

Models	Embeddings	Accuracy(%)	F1(NR)	F1(LR)	F1(MR)	F1(SR)	Training Time(s)
Baseline1-SVM		52.69	0	0	0	0.69	9.92
Baseline2-Random Forest		52.69	0.42	0	0	0.68	1.57
CNN	Char	52.69	0	0	0	0.69	12.26
	Word	52.69	0	0	0	0.69	11.24
	Doc	52.69	0	0	0	0.69	5.52
	Char+Word	52.15	0.1	0	0	0.69	16.47
	Char+Doc	49.46	0.47	0.08	0.12	0.65	13.76
	Word+Doc	56.99	0.58	0	0.09	0.69	13.68
	Char+ Word+Doc	52.15	0.48	0	0.11	0.67	16.05
LSTM	Char	52.69	0	0	0	0.69	848.67
	Word	52.69	0	0	0	0.69	245.07
	Doc	52.69	0	0	0	0.69	59.33
	Char+Word	52.69	0	0	0	0.69	886.14
	Char+Doc	47.31	0.38	0.09	<b>0.29</b>	0.62	892.24
	Word+Doc	55.91	0.50	0	0	0.71	885.39
	Char+ Word+Doc	55.91	0.54	0	0	0.70	895.94
Bi-LSTM	Char	52.69	0	0	0	0.69	1814.12
	Word	52.69	0	0	0	0.69	1763.65
	Doc	52.69	0	0	0	0.69	122.32
	Char+Word	53.23	0.11	0	0	0.69	1889.90
	Char+Doc	55.38	0.56	0	0.19	0.69	1920.11
	Word+Doc	51.61	0.49	<b>0.11</b>	0.11	0.66	1931.09
	Char+ Word+Doc	46.78	0.46	0.07	0.11	0.63	1919.43
Transformer Encoder	Char	52.15	0	0	0	0.69	566.00
	Word	41.40	0.05	0	0.09	0.59	513.08
	Doc	46.77	0.40	0.08	0.13	0.64	11.80
	Char+Word	52.69	0.30	0	0	0.68	626.09
	Char+Doc	58.39	0.48	0	0.16	0.61	551.60
	Word+Doc	54.30	0.45	0	0.08	0.70	565.59
	Char+ Word+Doc	40.32	0.41	0	0.12	0.57	685.88
BERT		57.52	0.61	0	0.12	0.69	85.97
RoBERTa		<b>58.60</b>	<b>0.59</b>	0	0	<b>0.72</b>	53.93
AlBERT		49.46	0.47	0	0	0.64	88.92
Electra		56.45	0.49	0	0	0.71	88.56
GPT2		51.08	0.05	0	0	0.68	43.18

Table 2: Detection Accuracy with UMD Reddit Suicidality Dataset (NR: No Risk, LR: Low Risk, MR: Moderate Risk, SR: Severe Risk)