

Results from various Embeddings

Codegraph

August 2020

1 USE Embeddings

Docstrings > 300 characters

Results	Model	Pr@5	Pr@10	LPr@5	LPr@10
All masking	Universal Sentence Encoder	0.28	0.41	0.77	0.85
	roberta-base-nli-stsb-mean-tokens	0.08	0.12	0.41	0.55
	roberta-large-nli-stsb-mean-tokens	0.07	0.10	0.38	0.52
	bert-base-nli-stsb-mean-tokens	0.08	0.12	0.46	0.61
	bert-large-nli-stsb-mean-tokens	0.07	0.10	0.40	0.56
	distilbert-base-nli-stsb-mean-tokens	0.06	0.10	0.34	0.49
Without masking	Universal Sentence Encoder	0.34	0.47	0.83	0.88
	roberta-base-nli-stsb-mean-tokens	0.10	0.15	0.51	0.64
	roberta-large-nli-stsb-mean-tokens	0.10	0.15	0.52	0.64
	bert-base-nli-stsb-mean-tokens	0.12	0.16	0.58	0.71
	bert-large-nli-stsb-mean-tokens	0.11	0.16	0.56	0.68
	distilbert-base-nli-stsb-mean-tokens	0.09	0.14	0.50	0.63
With one masking	Universal Sentence Encoder	0.28	0.4	0.78	0.85
	roberta-base-nli-stsb-mean-tokens	0.07	0.11	0.41	0.55
	roberta-large-nli-stsb-mean-tokens	0.07	0.10	0.38	0.51
	bert-base-nli-stsb-mean-tokens	0.08	0.12	0.46	0.61
	bert-large-nli-stsb-mean-tokens	0.07	0.10	0.40	0.56
	distilbert-base-nli-stsb-mean-tokens	0.06	0.10	0.35	0.50

All Docstrings

Results	Model	Pr@5	Pr@10	LPr@5	LPr@10
All masking	Universal Sentence Encoder	0.21	0.30	0.67	0.76
	roberta-base-nli-stsb-mean-tokens	0.05	0.09	0.39	0.52
	roberta-large-nli-stsb-mean-tokens	0.05	0.08	0.34	0.48
	bert-base-nli-stsb-mean-tokens	0.06	0.09	0.41	0.56
	bert-large-nli-stsb-mean-tokens	0.05	0.07	0.35	0.49
Without masking	distilbert-base-nli-stsb-mean-tokens	0.04	0.07	0.31	0.45
	Universal Sentence Encoder	0.26	0.34	0.72	0.79
	roberta-base-nli-stsb-mean-tokens	0.07	0.11	0.48	0.60
	roberta-large-nli-stsb-mean-tokens	0.07	0.10	0.46	0.58
	bert-base-nli-stsb-mean-tokens	0.08	0.12	0.48	0.64
With one masking	bert-large-nli-stsb-mean-tokens	0.07	0.11	0.48	0.60
	distilbert-base-nli-stsb-mean-tokens	0.06	0.10	0.44	0.56
	Universal Sentence Encoder	0.2	0.29	0.67	0.76
	roberta-base-nli-stsb-mean-tokens	0.05	0.08	0.39	0.52
	roberta-large-nli-stsb-mean-tokens	0.04	0.08	0.34	0.48
	bert-base-nli-stsb-mean-tokens	0.05	0.09	0.41	0.56
	bert-large-nli-stsb-mean-tokens	0.05	0.07	0.35	0.50
	distilbert-base-nli-stsb-mean-tokens	0.04	0.07	0.31	0.45