

cases	doc_1		doc_2		decision	id
	authors	<ul style="list-style-type: none">Nimesh BhanaTerence L van Zyl	authors	<ul style="list-style-type: none">Nimesh BhanaTerence L. van Zyl	DUPLICATES	136
	title	Knowledge Graph Fusion for Language Model Fine-Tuning	title	Knowledge Graph Fusion for Language Model Fine-tuning		
	publication_date	2022-06-21 00:00:00	publication_date	2022-06-21 00:00:00		
	source	SupportedSources.SEMANTIC_SCHOLAR	source	SupportedSources.INTERNET_ARCHIVE		
	journal		journal			
	volume		volume			
	doi	10.1109/ISCM156532.2022.10068451	doi			
	urls	<ul style="list-style-type: none">https://www.semanticscholar.org/paper/eeaf05f6f3a8dbd14cc318736d6ec833f8ef3f60	urls	<ul style="list-style-type: none">https://web.archive.org/web/20220709034533/https://arxiv.org/pdf/2206.14574v1.pdf		
	id	id1771291892618029916	id	id-6743743115898506959		
	abstract	Language Models such as BERT (Bidirectional Encoder Representations from Transformers) have grown in popularity due to their ability to be pre-trained and perform robustly on a wide range of Natural Language Processing tasks. Often seen as an evolution over traditional word embedding techniques, they can produce semantic representations of text, useful for tasks such as semantic similarity. However, state-of-the-art models often have high computational requirements and lack global context or domain knowledge which is required for complete language understanding. To address these limitations, we investigate the benefits of knowledge incorporation into the fine-tuning stages of BERT. An existing K-BERT model, which enriches sentences with triplets from a Knowledge Graph, is adapted for the English language and extended to inject contextually relevant information into sentences. As a side-effect, changes made to K-BERT for accommodating the English language also extend to other word-based languages. Experiments conducted indicate that injected knowledge introduces noise. We see statistically significant improvements for knowledge-driven tasks when this noise is minimised. We show evidence that, given the appropriate task, modest injection with relevant, high-quality knowledge is most performant.	abstract	Language Models such as BERT have grown in popularity due to their ability to be pre-trained and perform robustly on a wide range of Natural Language Processing tasks. Often seen as an evolution over traditional word embedding techniques, they can produce semantic representations of text, useful for tasks such as semantic similarity. However, state-of-the-art models often have high computational requirements and lack global context or domain knowledge which is required for complete language understanding. To address these limitations, we investigate the benefits of knowledge incorporation into the fine-tuning stages of BERT. An existing K-BERT model, which enriches sentences with triplets from a Knowledge Graph, is adapted for the English language and extended to inject contextually relevant information into sentences. As a side-effect, changes made to K-BERT for accommodating the English language also extend to other word-based languages. Experiments conducted indicate that injected knowledge introduces noise. We see statistically significant improvements for knowledge-driven tasks when this noise is minimised. We show evidence that, given the appropriate task, modest injection with relevant, high-quality knowledge is most performant.		
	versions		versions			