

## Resumen Artículo

### ¿Language Models as Knowledge Bases?

Recientemente, los modelos de lenguaje pre-entrenados de alta capacidad como ELMO y BERT se han hecho muy importantes en el NLP (procesamiento de lenguaje natural). Esto es debido a que están muy optimizados y se usan para resolver la siguiente tarea en una secuencia o algunas palabras ocultas o alguna parte de una secuencia dada (por ejemplo: "Dante nació en [oculto] en el año 1265"). Estos parámetros parecen almacenar grandes cantidades de conocimiento lingüístico útil para tareas downstream (tareas que dependen de otro para completarse).

En contraste, las bases de conocimiento son soluciones efectivas para acceder a datos estableciendo queries como (Dante, born-in, X). Sin embargo, en la práctica a veces necesitamos extraer datos relacionados de texto u otras modalidades para poblar estas bases del conocimiento. Esto genera muchos líneas de trabajo complejas, lo que provoca que estos procesos puedan propagarse fácilmente y acumularse a través de este tipo de trabajo.

Dados las cualidades de los modelos de lenguaje como representaciones potenciales de conocimiento relacional, estamos interesados en el conocimiento relacional y presentes en los modelos pre-entrenados como ELMO y BERT que intentos de los relacionales aprenden como si fueran sin una configuración compleja en comparación con conocimiento simbólico externo alto



máximo de texto?

Para el propósito de responder las preguntas de arriba, utilizaremos la prueba LAMA (Análisis de Modelo de Lenguaje) la cual consiste en un conjunto de preguntas de conocimiento, cada una compuesta de un conjunto de hechos. Definimos que un modelo de lenguaje preentrenado conoce un hecho (sujeto-relación-objeto) como (Dante-Nacido en-Florencia) si este puede predecir satisfactoriamente los objetos ocultos en una oración como (Dante, nacido en \_\_\_\_\_).

Nuestra investigación revela que el modelo BERT más grande de Devlin et al. captura conocimiento relacionado al conocimiento con un extractor de relaciones y un enlazador de entidades basado en reglas. Además, maximización de hechos puede ser recuperado sorprendentemente bien de lenguajes preentrenados. Sin embargo, algunos relaciones (en particular las relaciones N-a-M) provocan un rendimiento muy pobre. Por último, BERT logra resultados muy razonables para aseguramiento de la calidad abierta al dominio, alcanzando un 57.1 por ciento de precisión comparado al 63.5 por ciento de una base de conocimiento construida usando un sistema de extracción de relaciones con reglas específicas.

Norma