



dcc dccuchile/**bert-base-spanish-wwm-uncased**

like 13

- Fill-Mask
- PyTorch
- TensorFlow
- JAX
- Transformers
- Spanish
- arxiv:1904.09077
- arxiv:1906.01502
- arxi
- bert
- masked-lm
- AutoTrain Compatible

- Model card
- Files and versions
- Community

Edit model card

BETO: Spanish BERT

BETO is a BERT model trained on a big Spanish corpus. BETO is of size similar to a BERT-Base and was trained with the Whole Word Masking technique. Below you find Tensorflow and Pytorch checkpoints for the uncased and cased versions, as well as some results for Spanish benchmarks comparing BETO with Multilingual BERT as well as other (not BERT-based) models.

- Spanish Wikis: Wich include Wikipedia, Wikinews, Wikiquotes and more. These were first processed with wikiextractor (<https://github.com/josecannete/wikiextractorforBERT>) using the wikis dump of 20/04/2019.
- ParaCrawl: Spanish portion of ParaCrawl (<http://opus.nlpl.eu/ParaCrawl.php>)
- EUBookshop: Spanish portion of EUBookshop (<http://opus.nlpl.eu/EUbookshop.php>)
- MultiUN: Spanish portion of MultiUN (<http://opus.nlpl.eu/MultiUN.php>)
- OpenSubtitles: Spanish portion of OpenSubtitles2018 (<http://opus.nlpl.eu/OpenSubtitles-v2018.php>)
- DGC: Spanish portion of DGT (<http://opus.nlpl.eu/DGT.php>)

¿Cómo se crea un word embedding?

1. Tomamos **grandes cantidades** de texto.
2. Descomponemos el texto en substrings de tamaño **n**.

En mi viaje a Tailandia comí duriáns por primera vez.

¿Cómo se crea un word embedding?

1. Tomamos **grandes cantidades** de texto.
2. Descomponemos el texto en substrings de tamaño n .

En mi viaje a Tailandia comí duriáns por primera vez.

¿Cómo se crea un word embedding?

1. Tomamos **grandes cantidades** de texto.
2. Descomponemos el texto en substrings de tamaño n .

En mi viaje a Tailandia comí duriáns por primera vez.

¿Cómo se crea un word embedding?

1. Tomamos **grandes cantidades** de texto.
2. Descomponemos el texto en substrings de tamaño n .

En mi viaje a Tailandia comí duriáns por primera vez.

¿Cómo se crea un word embedding?

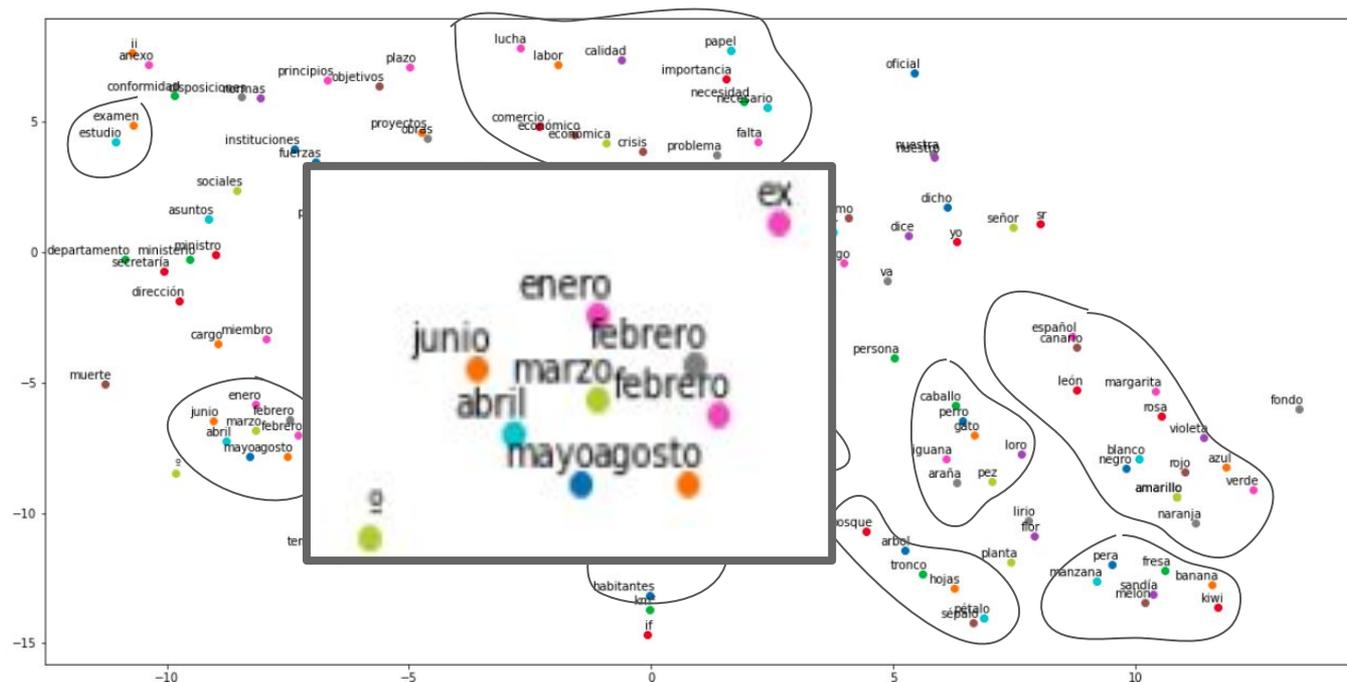
1. Tomamos **grandes cantidades** de texto.
2. Descomponemos el texto en substrings de tamaño **n**.
3. Cada substring es un ejemplo de entrenamiento para una red neuronal que trata de **predecir una palabra dada su contexto**.

¿Cómo se crea un word embedding?

1. Tomamos **grandes cantidades** de texto.
2. Descomponemos el texto en substrings de tamaño n .
3. Cada substring es un ejemplo de entrenamiento para una red neuronal que trata de **predecir una palabra dada su contexto**.
4. El embedding es la penúltima capa de la red neuronal, donde cada dimensión es una neurona de esta red.

Word Embeddings

Espacio vectorial donde palabras con **significados parecidos** tienden a tener **vectores similares**.



Proyección 2D del espacio 300D del embedding de FastText

