



Embeddings Aplicado a Texto



Word Embeddings



Word Embeddings

- Palabras individuales en vectores, pero no en cualquier vector, sino en vectores que capturen su significado o contexto semántico.
- Por ejemplo: Sinónimos deberían tener vectores cercanos, antónimos vectores en sentidos opuestos, etc



Word Embeddings

- Palabras individuales en vectores, pero no en cualquier vector, sino en vectores que capturen su significado o contexto semántico.
- Por ejemplo: Sinónimos deberían tener vectores cercanos, antónimos vectores en sentidos opuestos, etc
- “Sueño” desde 1960



Word Embeddings

- Palabras individuales en vectores, pero no en cualquier vector, sino en vectores que capturen su significado o contexto semántico.
- Por ejemplo: Sinónimos deberían tener vectores cercanos, antónimos vectores en sentidos opuestos, etc
- “Sueño” desde 1960
- Realidad en 2013: "Efficient Estimation of Word Representations in Vector Space"



Word Embeddings

- Palabras individuales en vectores, pero no en cualquier vector, sino en vectores que capturen su significado o contexto semántico.
- Por ejemplo: Sinónimos deberían tener vectores cercanos, antónimos vectores en sentidos opuestos, etc
- "Sueño" desde 1960
- Realidad en 2013: "Efficient Estimation of Word Representations in Vector Space"

- Modelo Word2Vec
 - A partir de otro modelo DL
- Este modelo predice palabras que acompañan a una dada por delante y detrás



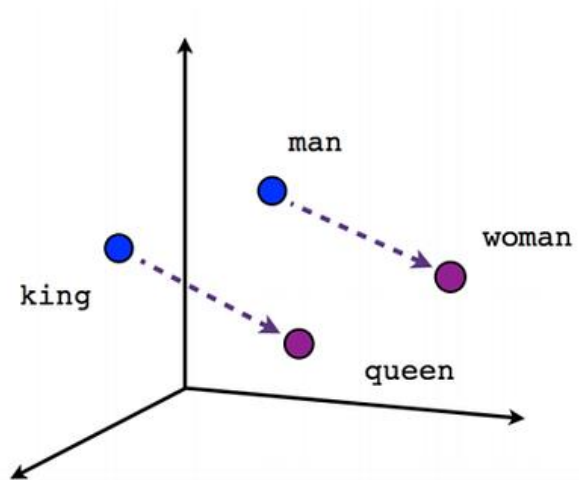
Word Embeddings

- Lo importante: propiedades semánticas se transforman en propiedades “espaciales”

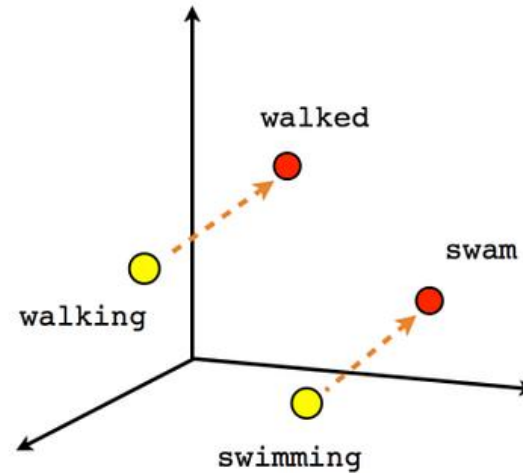


Word Embeddings

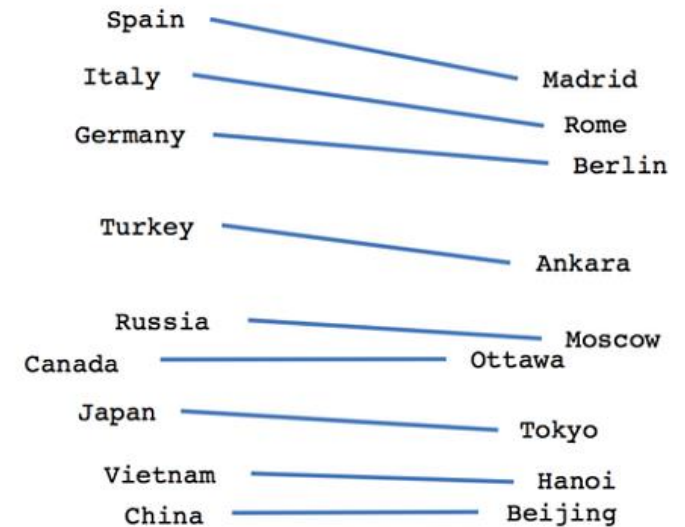
- Lo importante: propiedades semánticas se transforman en propiedades “espaciales”



Male-Female



Verb tense

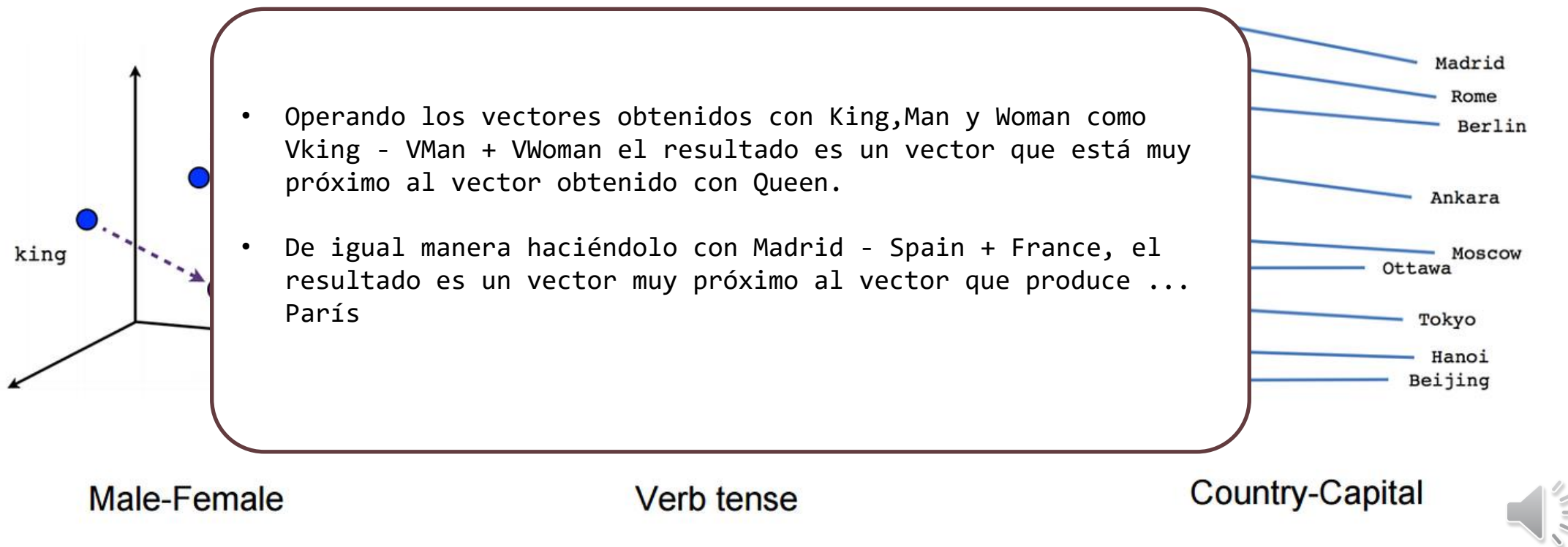


Country-Capital



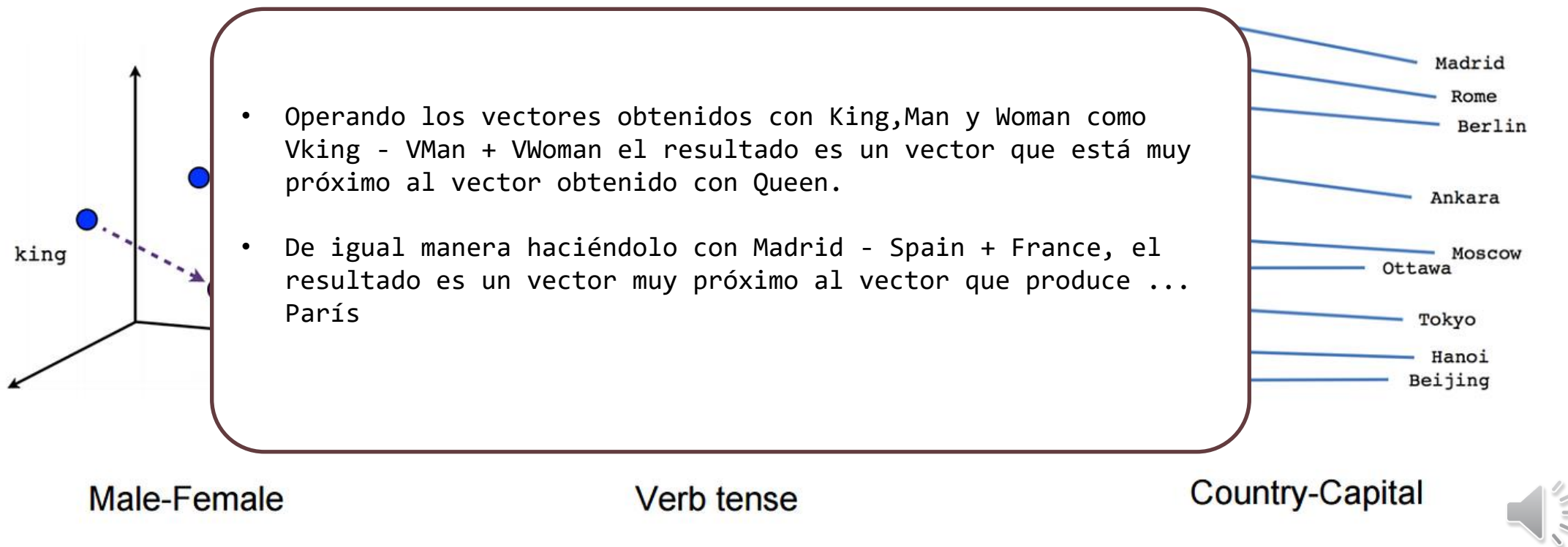
Word Embeddings

- Lo importante: propiedades semánticas se transforman en propiedades “espaciales”



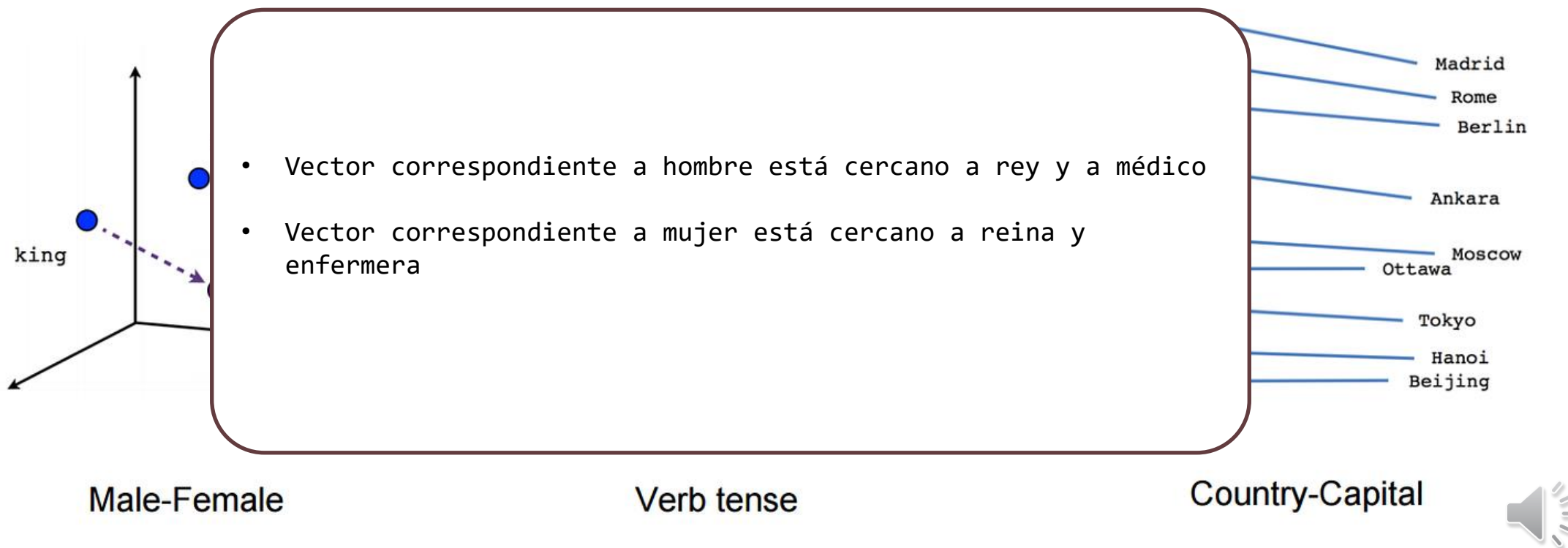
Word Embeddings

- Lo importante: propiedades semánticas se transforman en propiedades “espaciales”



Word Embeddings

- Lo importante: propiedades semánticas se transforman en propiedades “espaciales”



Sentence Embeddings



Sentence Embeddings

- El objetivo es análogo: Representar oraciones enteras como vectores de forma que se mantengan las propiedades semánticas y contextuales.



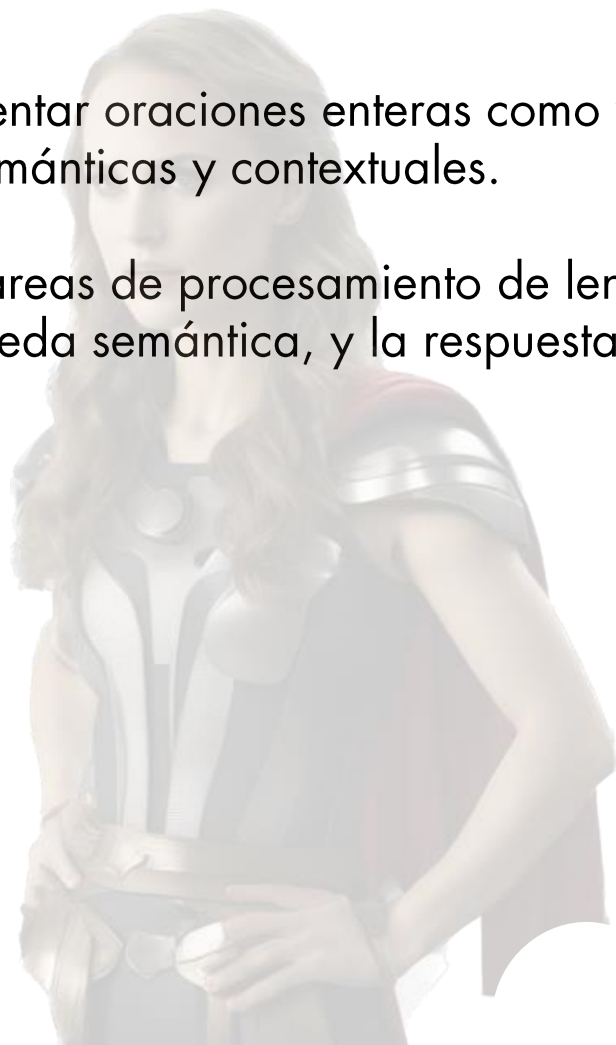
Sentence Embeddings

- El objetivo es análogo: Representar oraciones enteras como vectores de forma que se mantengan las propiedades semánticas y contextuales.
- Aplicable a una variedad de tareas de procesamiento de lenguaje natural (NLP), como la clasificación de texto, la búsqueda semántica, y la respuesta a preguntas, entre otras.



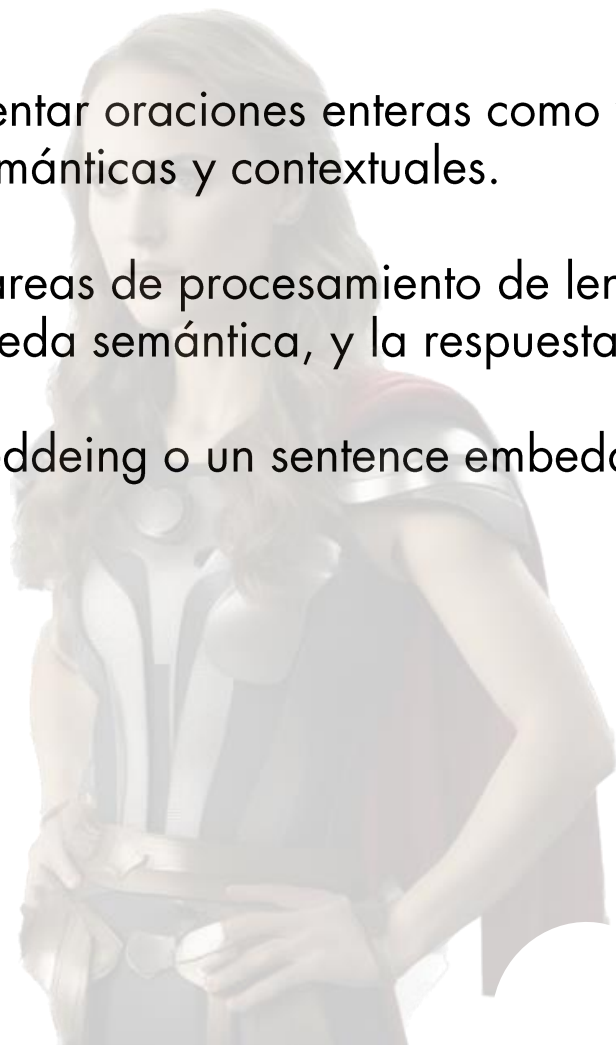
Sentence Embeddings

- El objetivo es análogo: Representar oraciones enteras como vectores de forma que se mantengan las propiedades semánticas y contextuales.
- Aplicable a una variedad de tareas de procesamiento de lenguaje natural (NLP), como la clasificación de texto, la búsqueda semántica, y la respuesta a preguntas, entre otras.



Sentence Embeddings

- El objetivo es análogo: Representar oraciones enteras como vectores de forma que se mantengan las propiedades semánticas y contextuales.
- Aplicable a una variedad de tareas de procesamiento de lenguaje natural (NLP), como la clasificación de texto, la búsqueda semántica, y la respuesta a preguntas, entre otras.
- ¿Cómo obtener un Word embedding o un sentence embedding o un sentence encoder?



Sentence Embeddings

- El objetivo es análogo: Representar oraciones enteras como vectores de forma que se mantengan las propiedades semánticas y contextuales.
- Aplicable a una variedad de tareas de procesamiento de lenguaje natural (NLP), como la clasificación de texto, la búsqueda semántica, y la respuesta a preguntas, entre otras.
- ¿Cómo obtener un Word embeddeing o un sentence embedding o un sentence encoder?
 1. Usar un WordEmbedding preentrenado (de Tensorflow Hub,Hugging Face), obtener los embeddings de las sentencias a tratar, obtener el centroide de dichos embeddings



Sentence Embeddings

- El objetivo es análogo: Representar oraciones enteras como vectores de forma que se mantengan las propiedades semánticas y contextuales.
- Aplicable a una variedad de tareas de procesamiento de lenguaje natural (NLP), como la clasificación de texto, la búsqueda semántica, y la respuesta a preguntas, entre otras.
- ¿Cómo obtener un Word embedding o un sentence embedding o un sentence encoder?
 1. Usar un WordEmbedding preentrenado (de Tensorflow Hub, Hugging Face), obtener los embeddings de las sentencias a tratar, obtener el centroide de dichos embeddings
 2. Usar directamente un modelo preentrenado para hacer sentence embeddings



Sentence Embeddings

- El objetivo es análogo: Representar oraciones enteras como vectores de forma que se mantengan las propiedades semánticas y contextuales.
- Aplicable a una variedad de tareas de procesamiento de lenguaje natural (NLP), como la clasificación de texto, la búsqueda semántica, y la respuesta a preguntas, entre otras.
- ¿Cómo obtener un Word embedding o un sentence embedding o un sentence encoder?
 1. Usar un WordEmbedding preentrenado (de Tensorflow Hub, Hugging Face), obtener los embeddings de las sentencias a tratar, obtener el centroide de dichos embeddings
 2. Usar directamente un modelo preentrenado para hacer sentence embeddings



