

Embedding Models

Instructor

Dipanjan Sarkar

Head of Community & Principal AI Scientist at Analytics Vidhya

Google Developer Expert - ML & Cloud Champion Innovator

Published Author

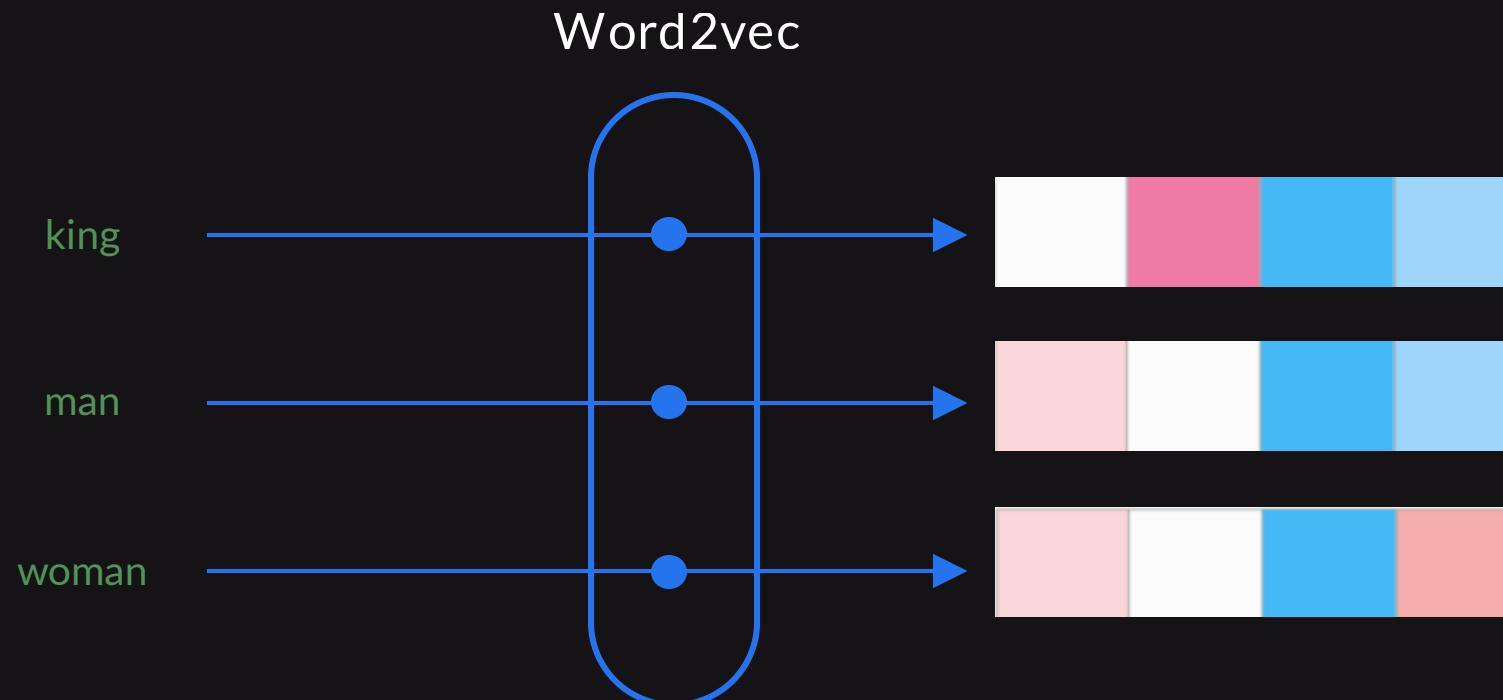


Outline

- What are embeddings?
- What are embedding models?
- Pre-trained embedding models
- Embedding similarity
- Embedding Models in LangChain
- Common usage patterns

Embeddings are learnt from several documents

A neural network language model, such as Word2Vec or Transformers, learns vector representations for each word, known as embeddings.



Embeddings encode the meaning of words

Let's say you take the [Big Five personality traits test](#) and get some scores as shown here

Openness to experience.....	79 out of 100
Agreeableness.....	75 out of 100
Conscientiousness.....	42 out of 100
Negative emotionality.....	50 out of 100
Extraversion.....	58 out of 100

Embeddings encode the meaning of words

For any person you can use these five attributes or dimensions to represent their personality

Openness to experience.....79 out of 100
Agreeableness..... 75 out of 100
Conscientiousness..... 42 out of 100
Negative emotionality..... 50 out of 100
Extraversion..... 58 out of 100

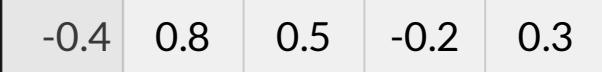
	Trait #1	Trait #2	Trait #3	Trait #4	Trait #5
Jay	-0.4	0.8	0.5	-0.2	0.3
Person #1	-0.3	0.2	0.3	-0.4	0.9
Person #2	-0.5	-0.4	-0.2	0.7	-0.1

Embeddings encode the meaning of words

The beauty of embeddings is that you can use them to make machines understand the meaning of words as numeric vectors

- 1 We can represent people (and things) as vectors of numbers (which is great for machines!).
- 2 We can easily calculate how similar vectors are to each other.

1. We can represent people (and things) as vectors of numbers (which is great for machines!).

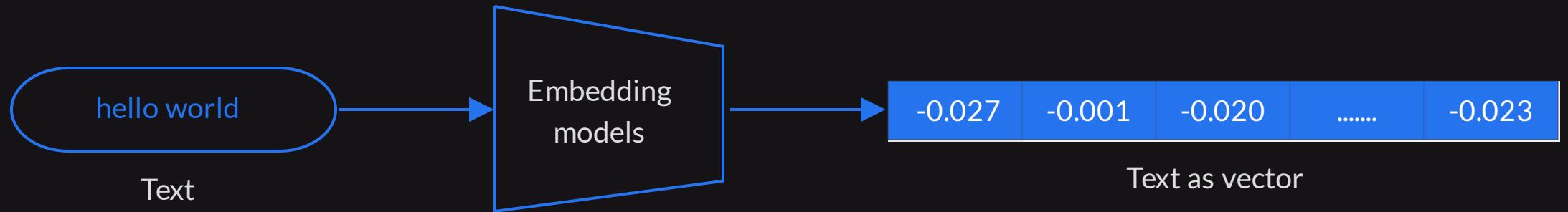
Jay 

2. We can easily calculate how similar vectors are to each other..

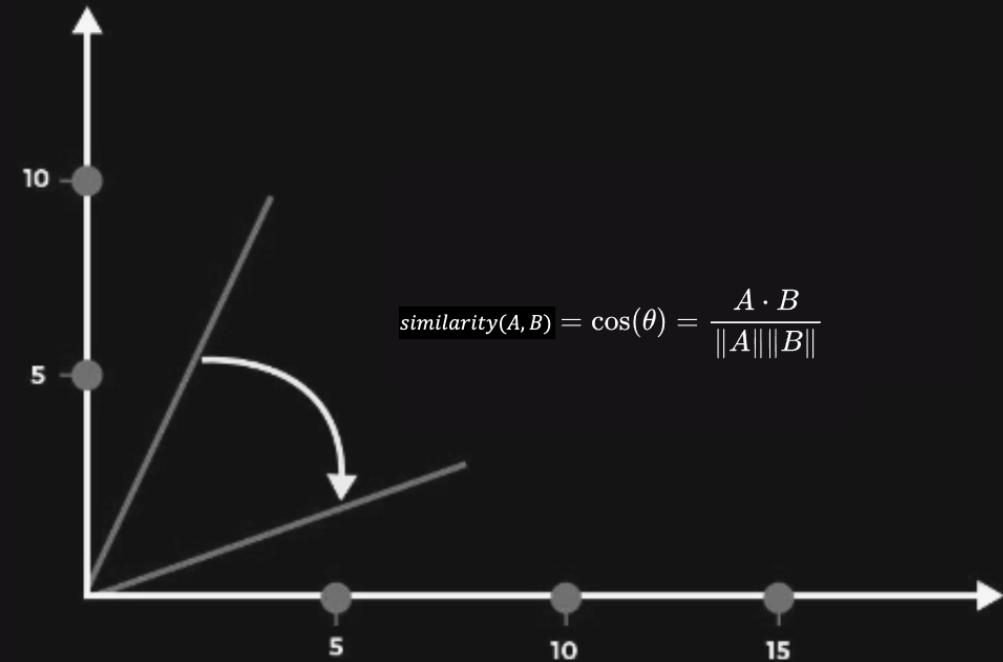
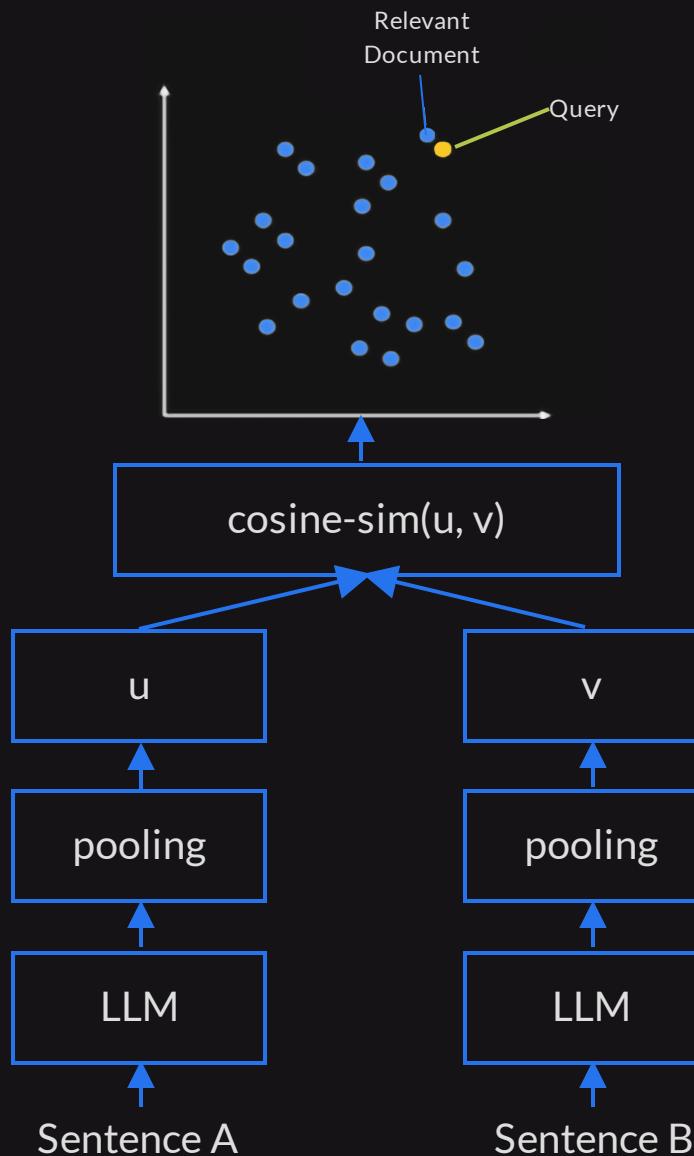
The people most similar to **Jay** are:

	Cosine_similarity
Person #1	0.86
Person #2	0.5
Person #3	-0.20

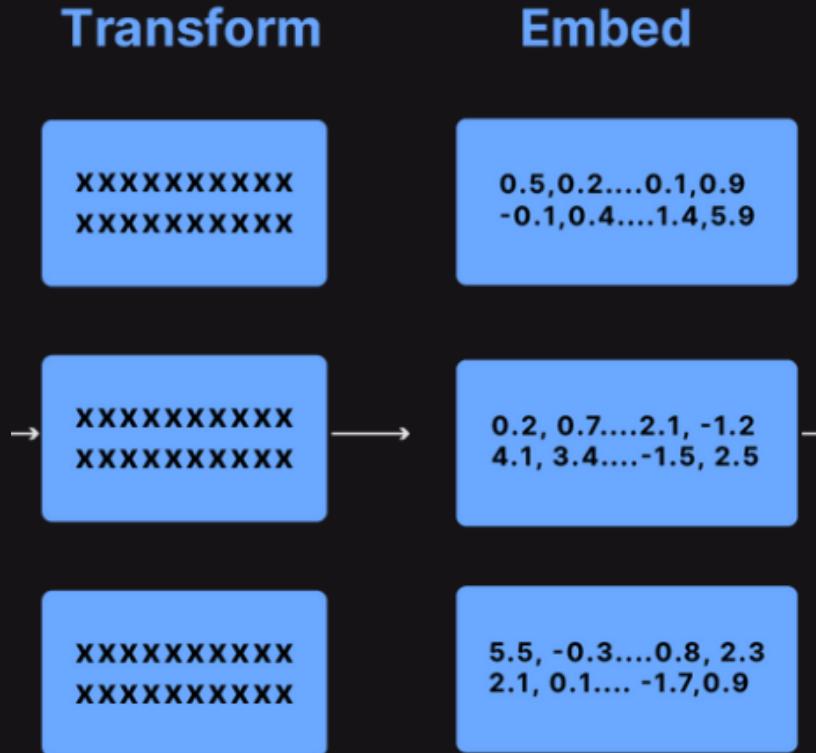
Pre-trained embedding models can be used directly



Embedding Similarity - Useful to find similar documents



Embedding Models



- LangChain supports API wrappers for all popular pre-trained embedding models
- Embedding models are typically used to convert documents or document chunks into embedding vectors
- Very useful in search engines and RAG systems

LangChain connects to all popular embedding models



LLM Embedding Model from [Open AI](#) or
[Google Gemini](#) or [HuggingFace](#)

Common usage patterns of embedding models in LangChain

```
from langchain_openai import OpenAIEMBEDDINGS  
  
embeddings_model = OpenAIEMBEDDINGS()  
embeddings = embeddings_model.embed_documents(  
    [  
        "Hi there!",  
        "Oh, hello!",  
        "What's your name?",  
        "My friends call me World",  
        "Hello World!"  
    ]  
)  
  
len(embeddings), len(embeddings[0])  
# Output  
# (5, 1536)
```

embed_documents - Embeds a list of documents - like a knowledgebase

```
from langchain_openai import OpenAIEMBEDDINGS  
  
embeddings_model = OpenAIEMBEDDINGS()  
embedded_query = embeddings_model.embed_query("Tell me what is AI?")  
  
len(embedded_query), embedded_query[:5]  
# Output  
# (1536, [0.00535, -0.000499, 0.03888, -0.00300, -0.00900])
```

embed_query - Embeds a single piece of text for the purpose of comparing to other embedded pieces of texts

Thank You
