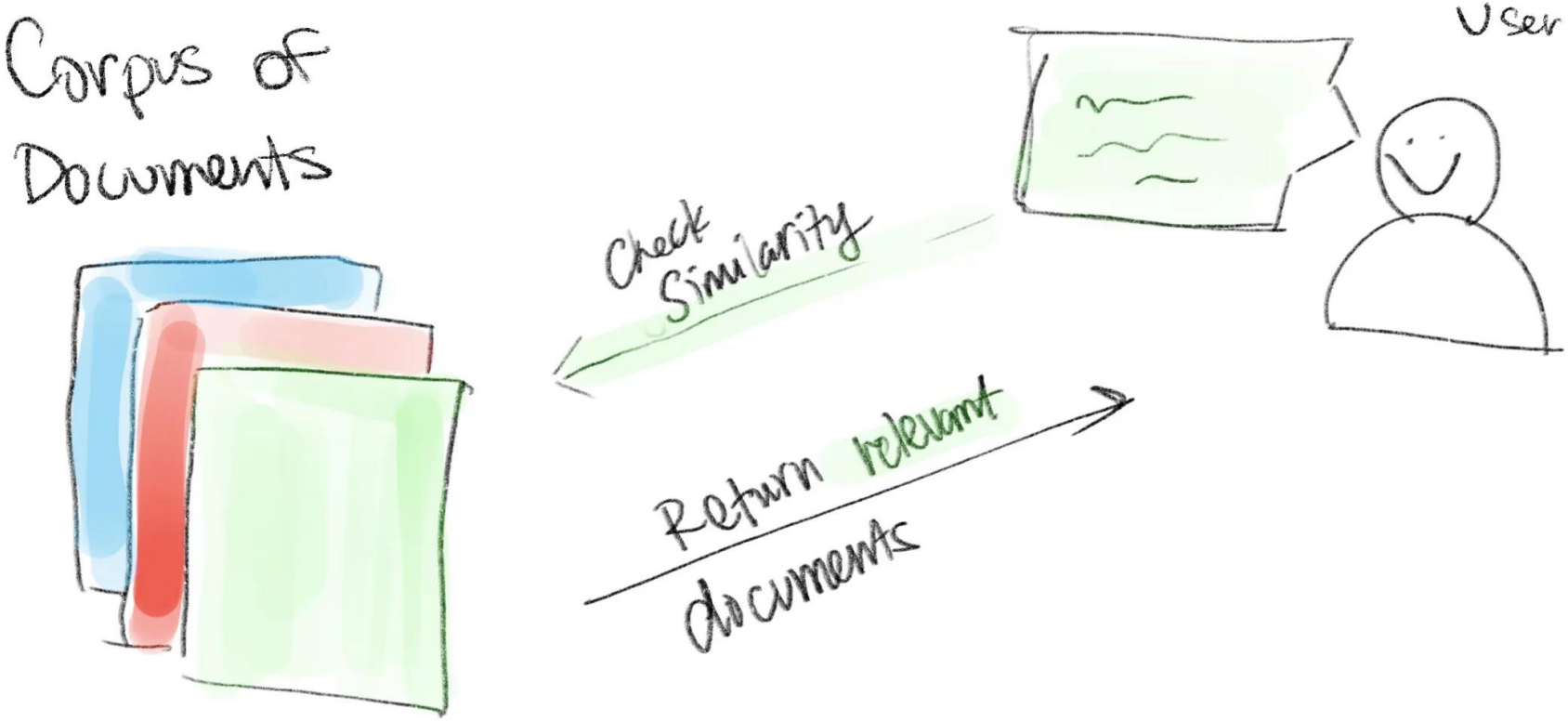
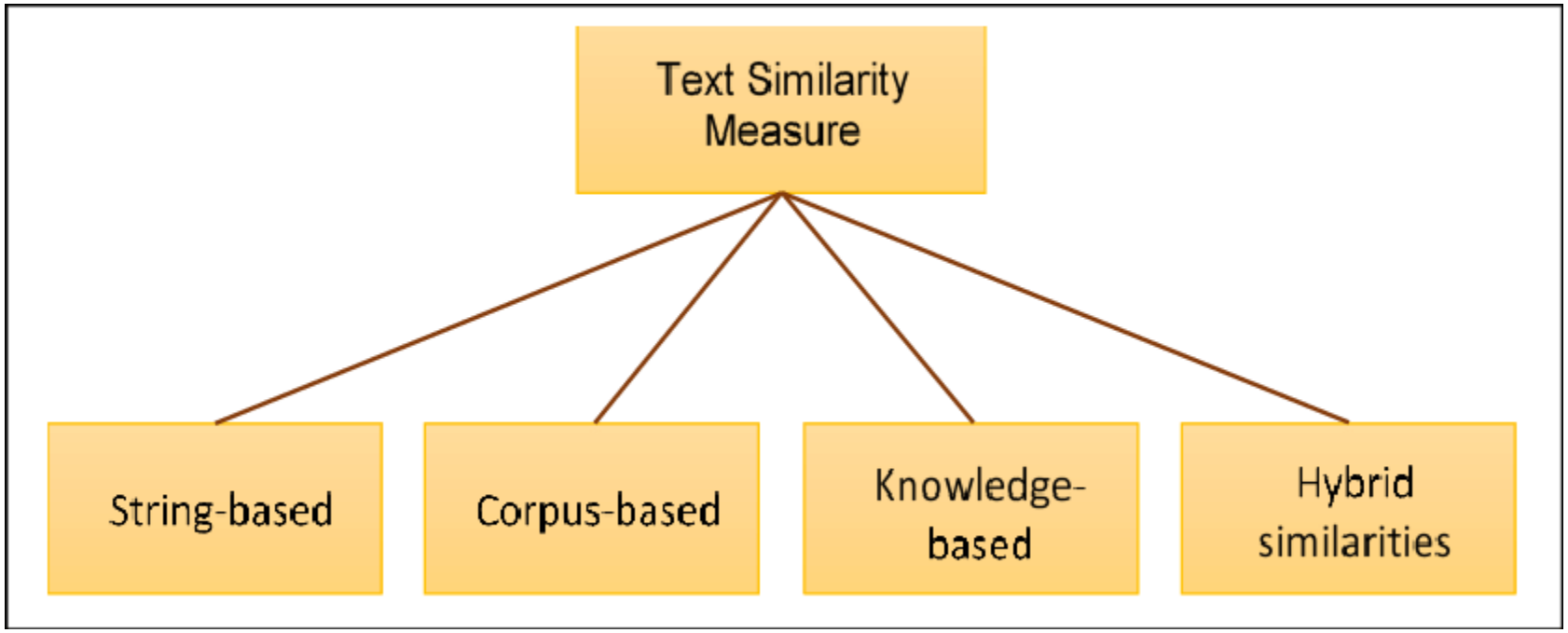
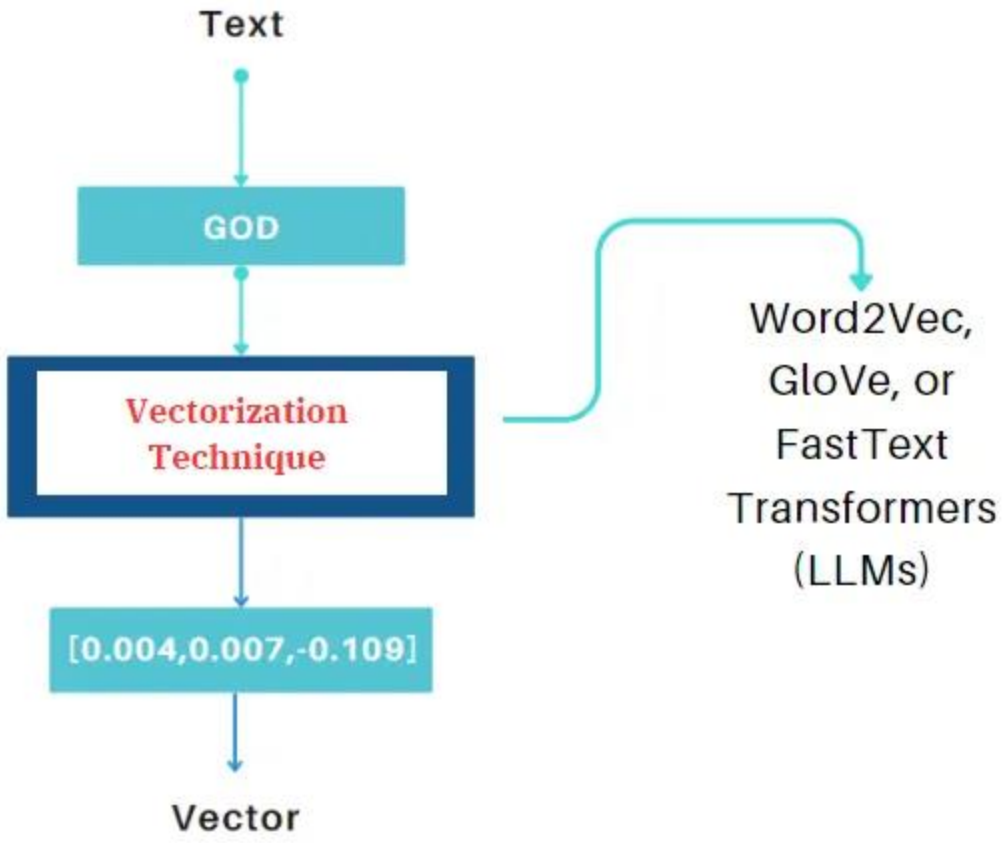


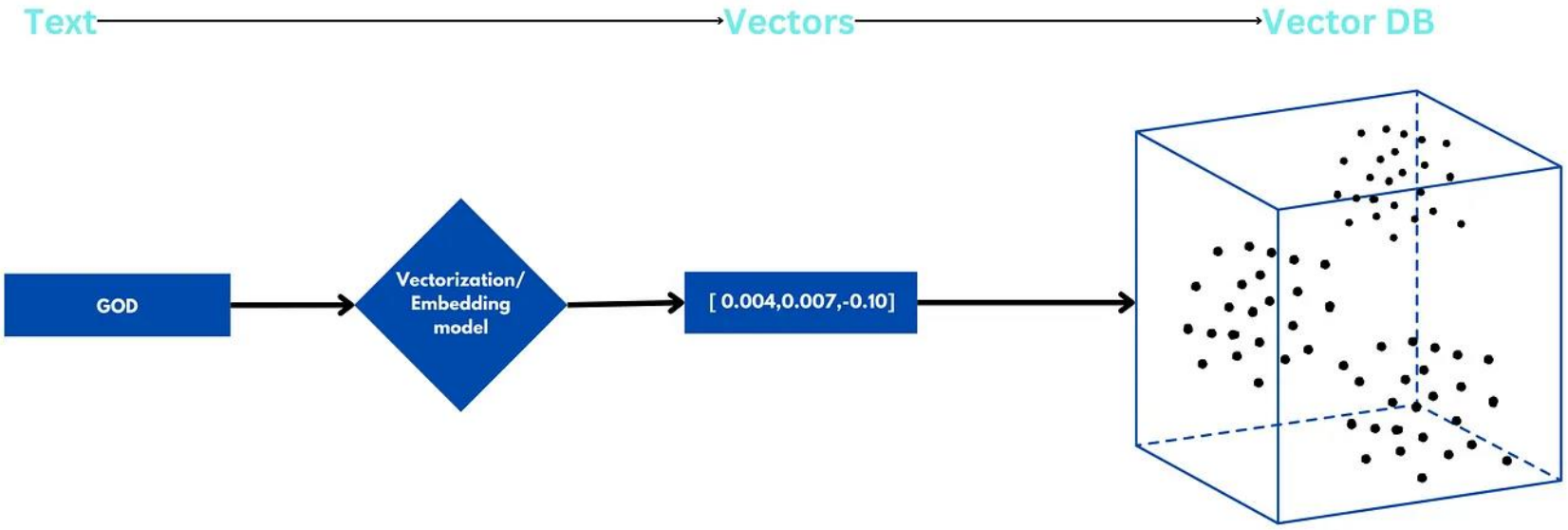
Bill Chambers
@BLLCHMBRS
LearnByBuilding.ai


Text Similarity Measure
String-based
Corpus-based
Knowledge-based
Hybrid similarities

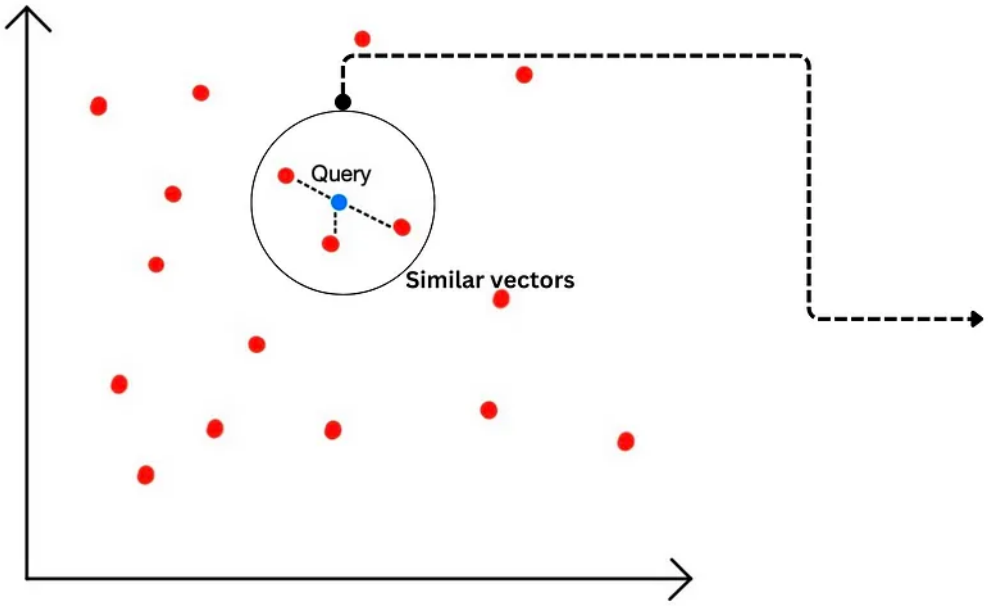
Step 1


Text
GOD
Vectorization Technique
[0.004,0.007,-0.109]
Vector
Word2Vec,
GloVe, or
FastText
Transformers
(LLMs)

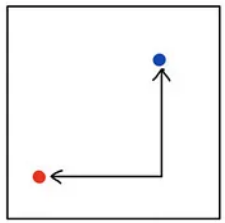
Step 2

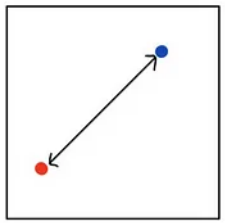

Text
Vectors
Vector DB
GOD
Vectorization/
Embedding
model
[0.004,0.007,-0.10]

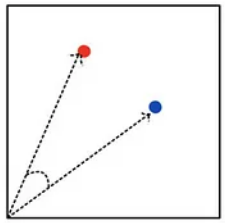
Step 3

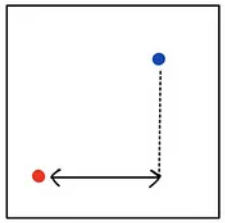


3 nearest neighbors for a query in a vector space


Manhattan


Euclidean

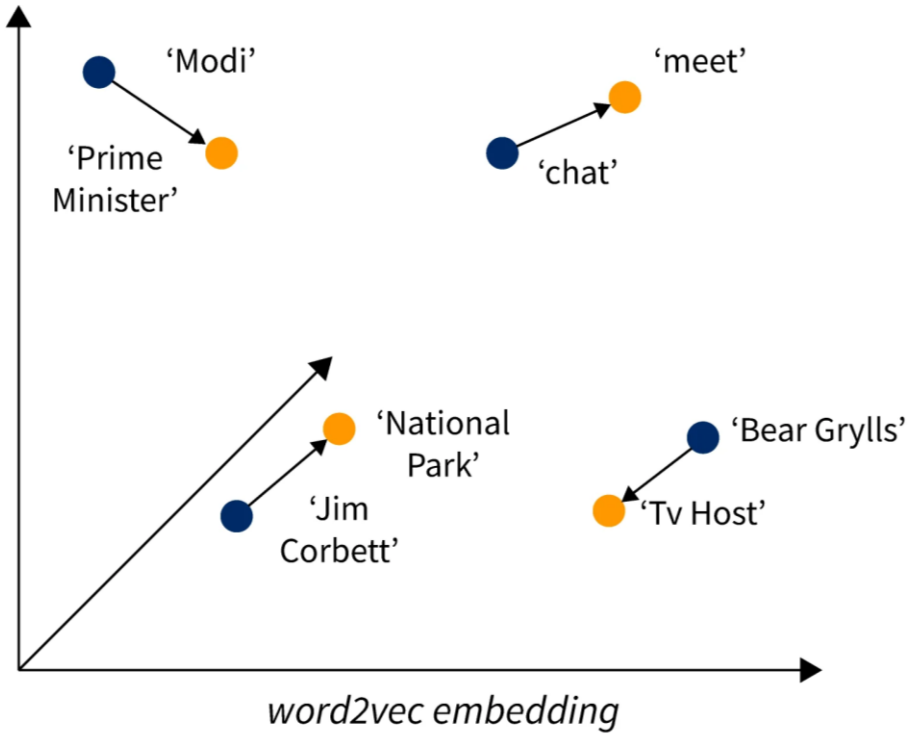

Cosine


Chebyshev

Similarity Measurement Techniques

Document 1

Modi *had a* **chat** *with* **Bear Grylls** *in* **Jim Corbett**



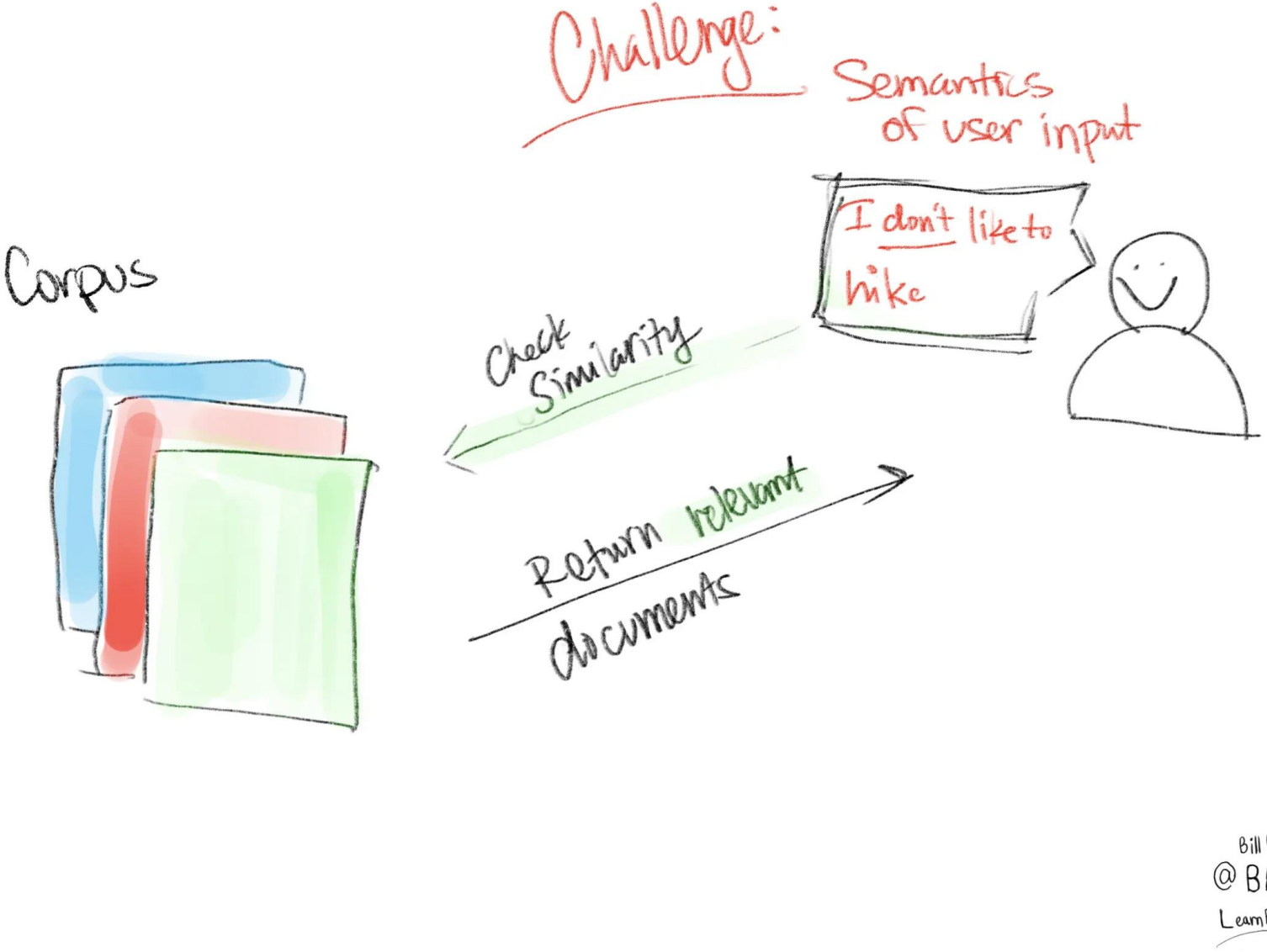
Document 2

The **prime minister meets** *the* **TV Host** *in a* **National Park**

You should have at least 8 GB of RAM available to run the 7B models, 16 GB to run the 13B models, and 32 GB to run the 33B models.

https://ollama.com/download

(metaverse) c:\Python3\my_project\project_llm>ollama run llama3.2



Bill Chambers
@BLLCHMBRS
LearnByBuilding.ai