GETTING TO KNOW LangChain

IN THIS WORKSHOP

- Provide a guided tour of LangChain:
  - Introduction to Large Language Models
  - Introduction to LangChain
  - Models, Prompts and Parsers
  - Chains

LARGE LANGUAGE MODELS (LLMs)

- Large language models are sophisticated artificial intelligence systems designed to understand and generate human-like text.
- They are built on deep learning architectures, enabling them to process and generate human-like text based on patterns in data.
- Characteristics:

MASSIVE DATA TRAINING

COMPLEX ARCHITECTURES

NATURAL LANGUAGE UNDERSTANDING

LARGE LANGUAGE MODELS (LLMs)

1 Billion computations per Second

1 Minute

LARGE LANGUAGE MODELS (LLMs)

1 Billion computations per Second

1 Hour

LARGE LANGUAGE MODELS (LLMs)

1 Billion computations per Second

1 Day

LARGE LANGUAGE MODELS (LLMs)

1 Billion computations per Second

1 Month

LARGE LANGUAGE MODELS (LLMs)

1 Billion computations per Second

1 Year

LARGE LANGUAGE MODELS (LLMs)

1 Billion computations per Second

100 Year

LARGE LANGUAGE MODELS (LLMs)

1 Billion computations per Second

10,000 Year

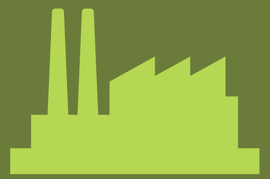
LARGE LANGUAGE MODELS (LLMs)

1 Billion computations per Second

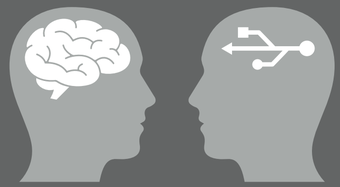
1,000,000 Year

USE CASES

CONTENT GENERATION



CHATBOTS



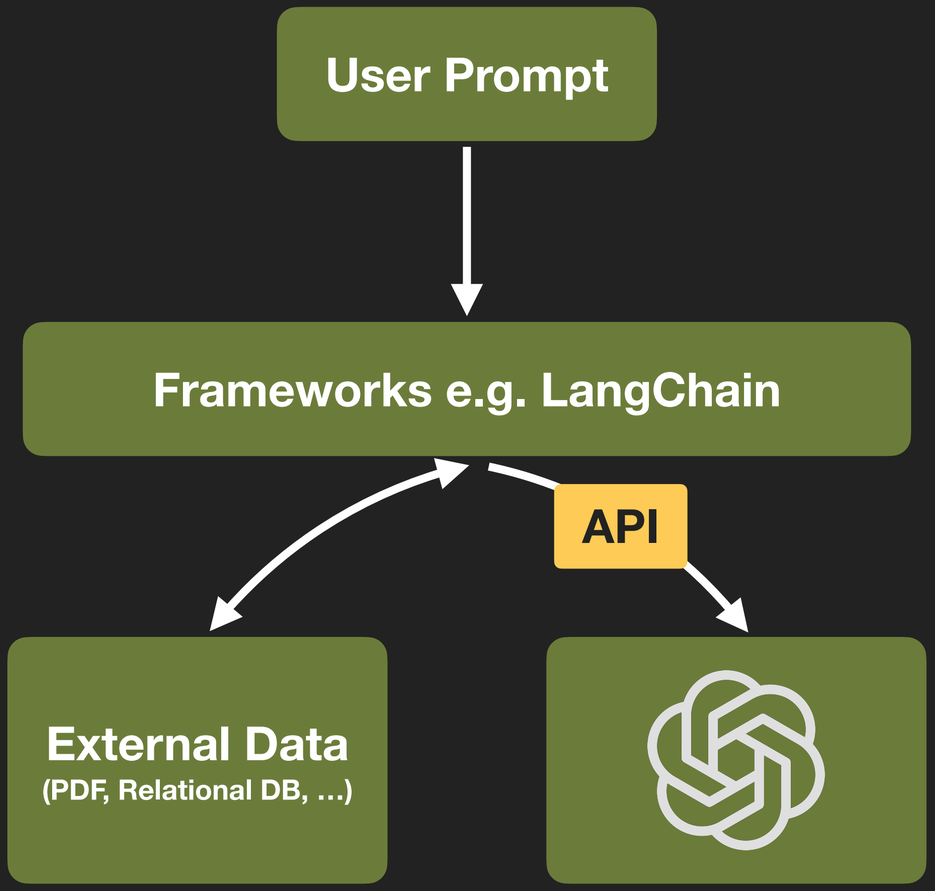
PROGRAMMING

SUMMARIZATION

RAG (RETRIEVAL AUGMENTED GENERATION)

- RAG allows us to retrieve snippets of information from external data sources and argument it to user's prompts to get tailored responses from Llama2

- LangChain: A framework that makes it easier to implement RAG



LangChain 🦜🔗

- is a framework for developing applications powered by language models. It enables applications that:
  - Are context-aware: connect a language model to sources of context (prompt instructions, few shot examples, content to ground its response in, etc.)
  - Reason: rely on a language model to reason (about how to answer based on provided context, what actions to take, etc.)

Templates

A collection of easily deployable reference architectures for a wide variety of tasks.

LangServe

A library for deploying LangChain chains as a REST API.

LangSmith

A developer platform that lets you debug, test, evaluate, and monitor chains built on any LLM framework and seamlessly integrates with LangChain.

Temperature

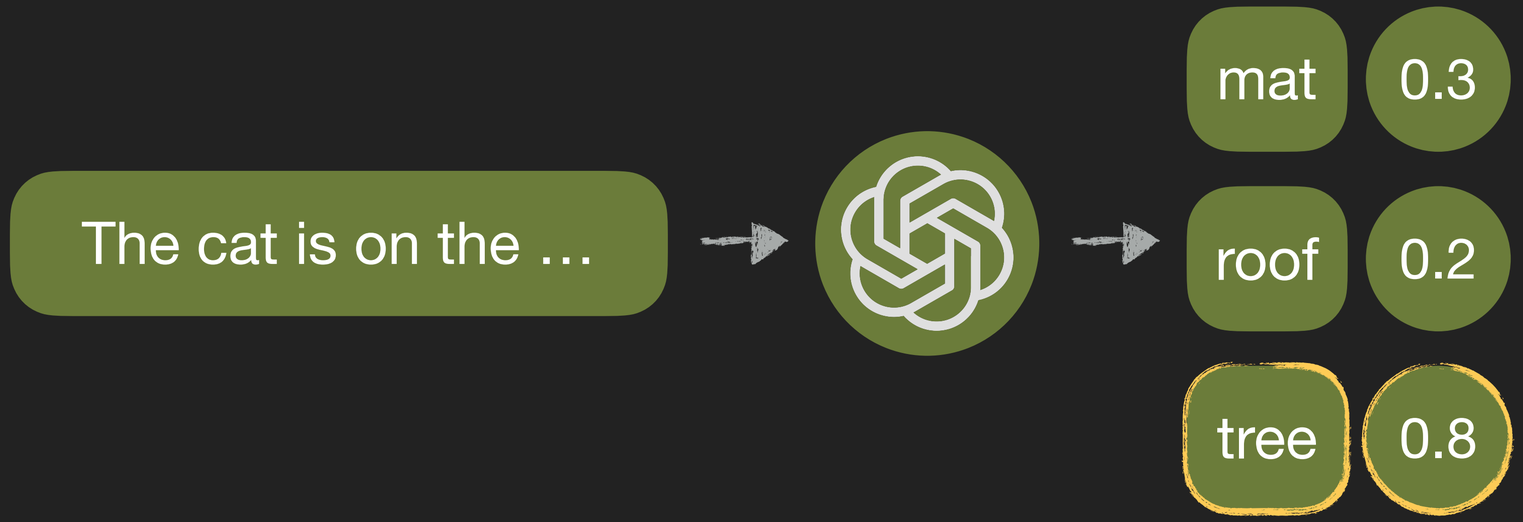
- LLM temperature is a parameter that influences the language model's output, determining whether the output is more random and creative or more predictable.

Higher temperature → more creative outputs

Lower temperature → more predictable outputs

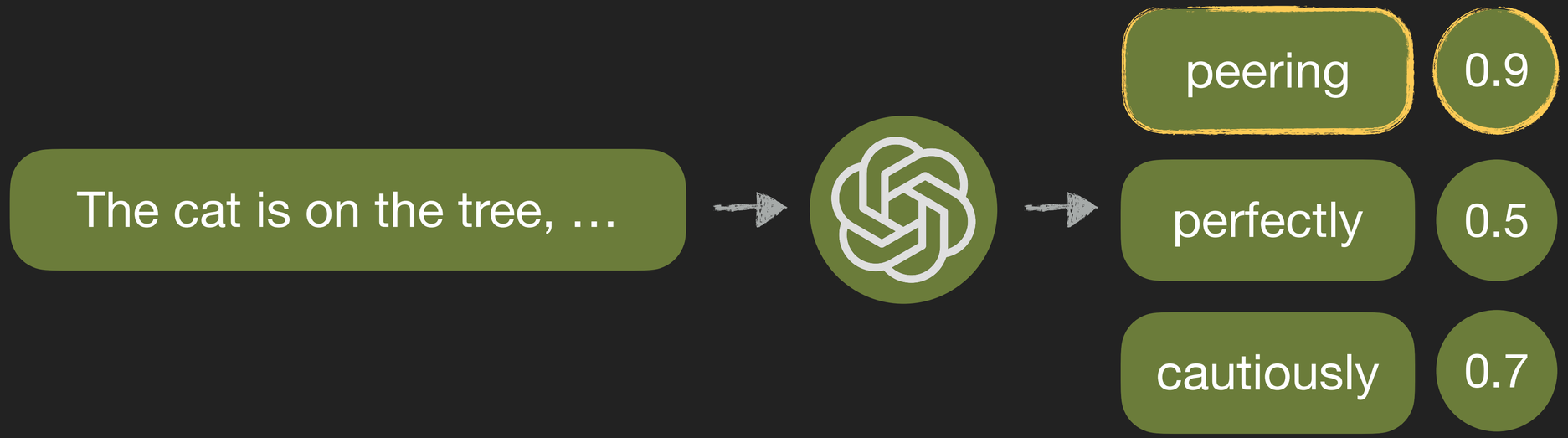
Temperature

- When generating text, the model considers a range of possible next words or tokens, each with a certain probability.



Temperature

- When generating text, the model considers a range of possible next words or tokens, each with a certain probability.



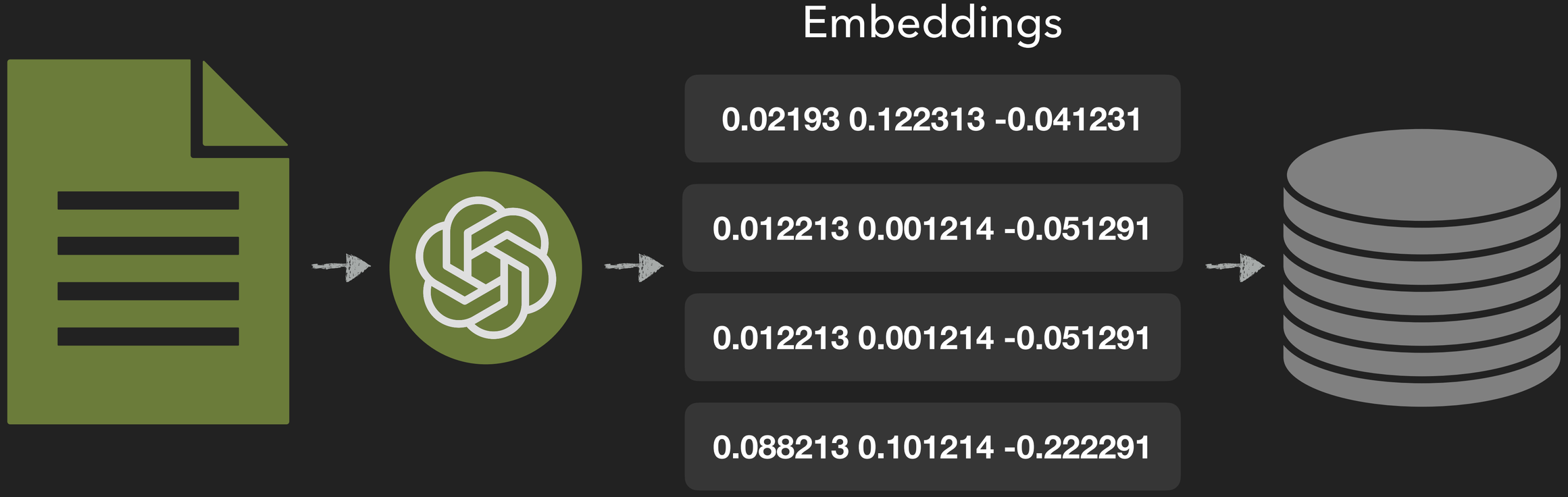
MODELS, PROMPTS AND PARSERS CODE :

Question Answering with LangChain

- LLMs are limited to input size
- You cannot feed 1000 thousands reviews to a LLM and ask questions about these reviews

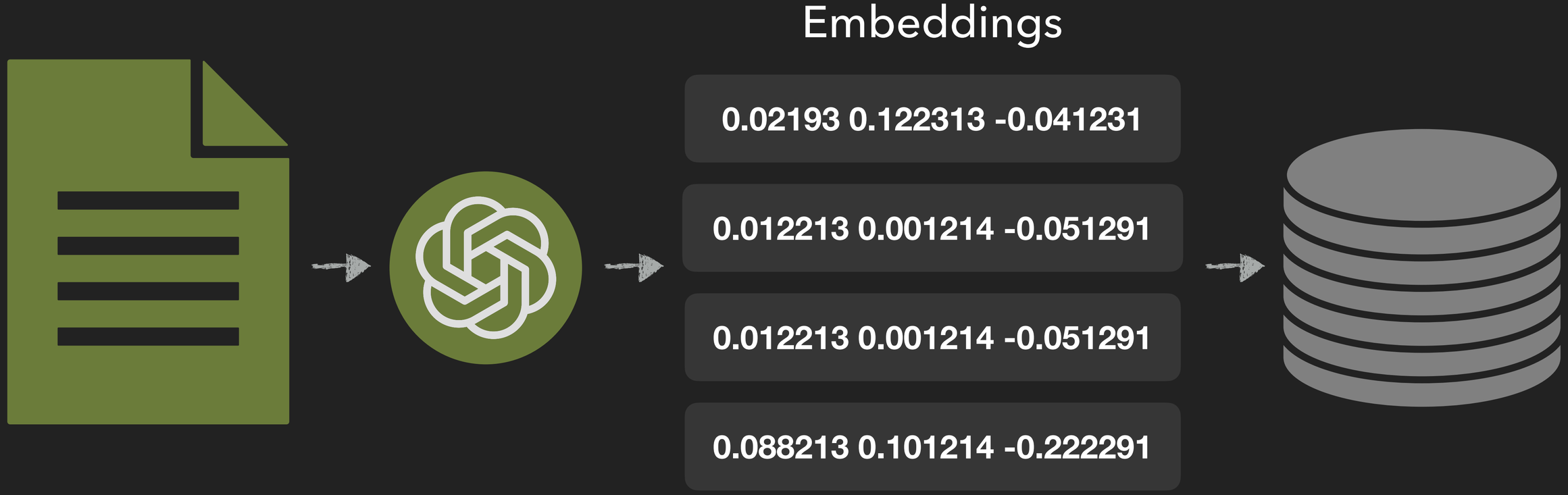
- Solution?

Embeddings

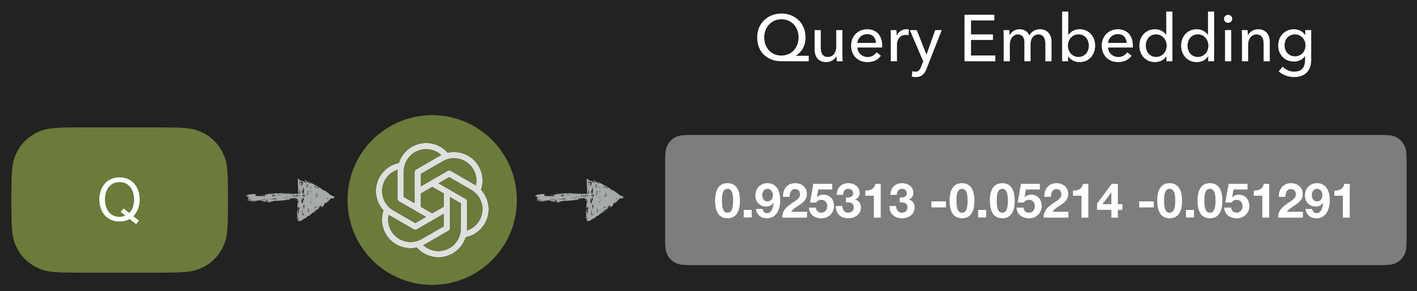

Embeddings
0.02193 0.122313 -0.041231
0.012213 0.001214 -0.051291
0.012213 0.001214 -0.051291
0.088213 0.101214 -0.222291

EMBEDDING CODE :

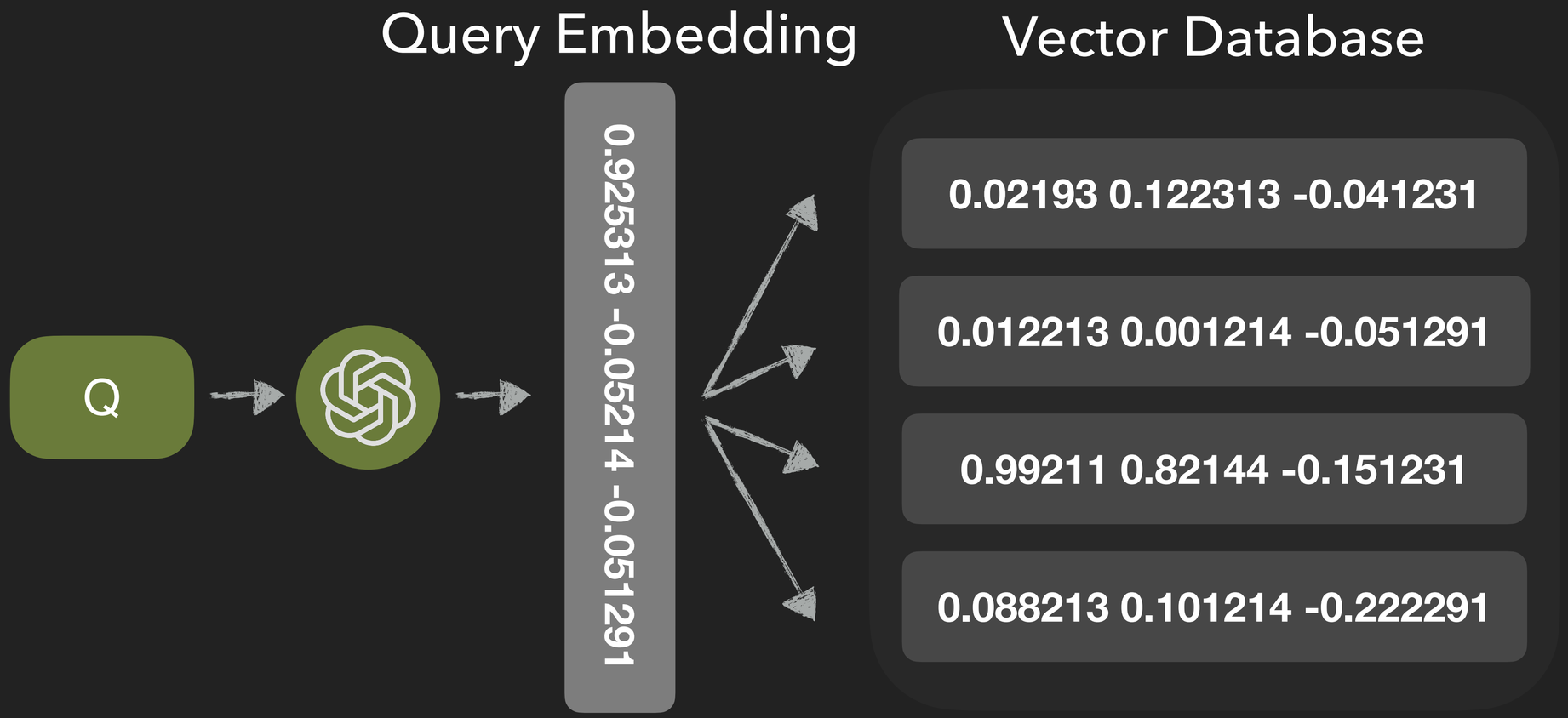
Step 1: Get The Database Embeddings



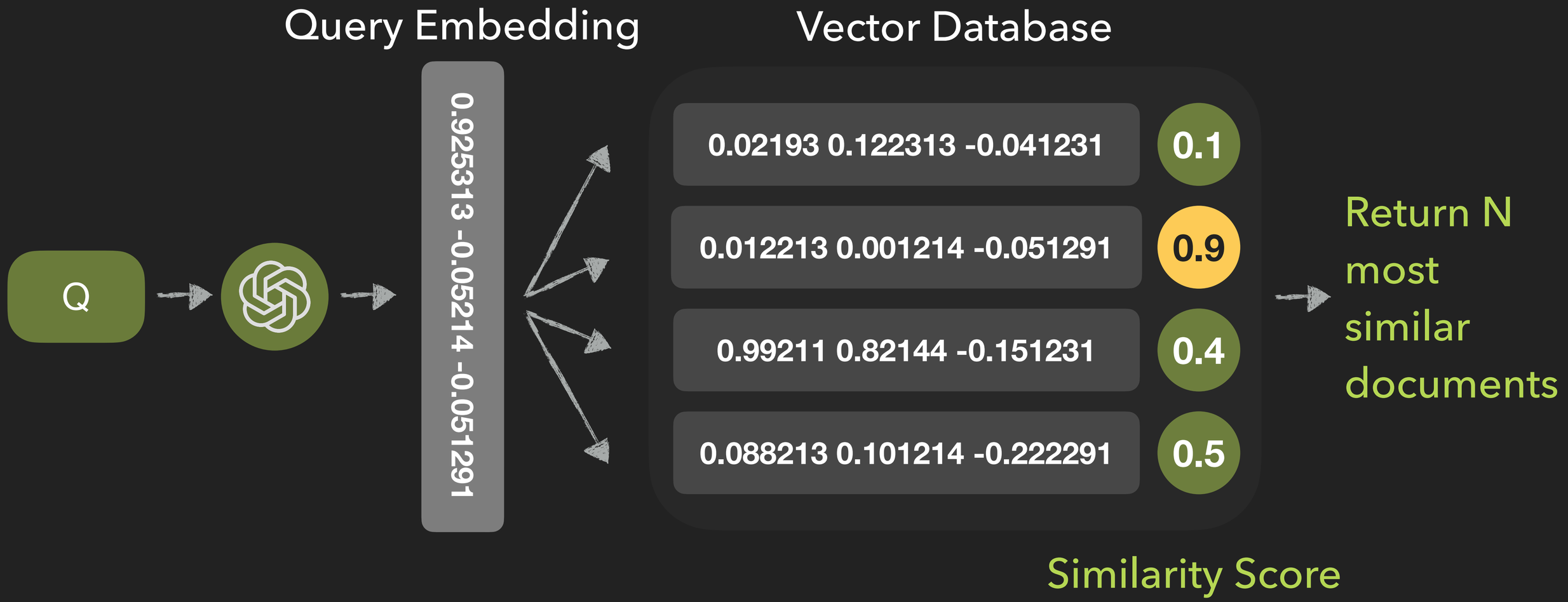
Step 2: Get The Query Embedding



Step 3: Calculate Similarity Scores



Step 3: Calculate Similarity Scores



QA CODE: