

UNDERSTANDING THE CONCEPT OF “SELF ATTENTION MECHANISM” IN GENERATIVE AI & LLM

Generative AI Deep Dives,
Key concepts for Transformers - Part 3

**GENERATIVE AI
FOR ALL**

 **DINESHLAL**



DINESH LAL
(DIRECTOR, DATA SCIENCE)



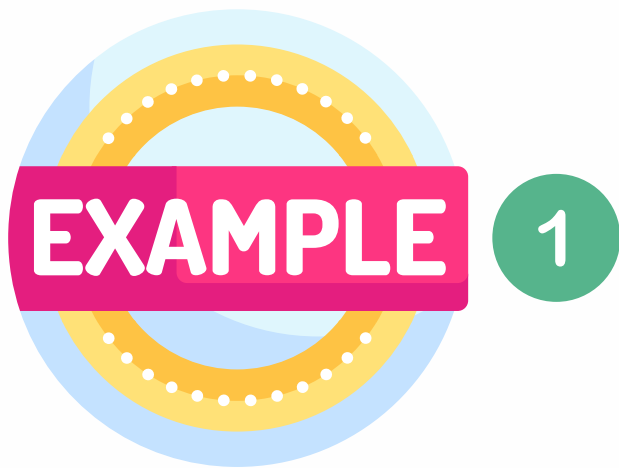
WHAT IS COVERED IN THIS DOCUMENT?

- This document explains the topic **Self Attention Mechanism**
- Document takes 5 different example of English sentences, and explains how different words around a word in a sentence makes all the difference in meaning
- In the last slide, proper definition is provided for self attention mechanism



WHAT IS SELF ATTENTION?

- Imagine you're reading a sentence. To understand a word's meaning, you consider other words around it. Lets see some examples:



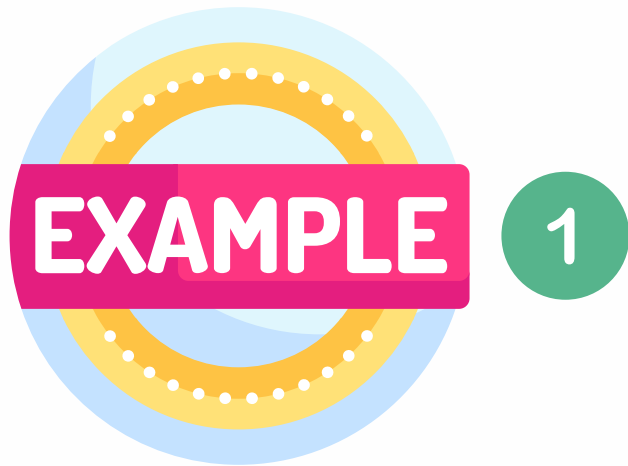
Sentence 1: "Let's visit the bank tomorrow."

Sentence 2: "I need to deposit money at the bank."

- The word "bank" has different meanings depending on the context.
- In sentence 1, it likely refers to a financial institution (e.g., for deposits). In sentence 2, it might indicate the edge of a river.



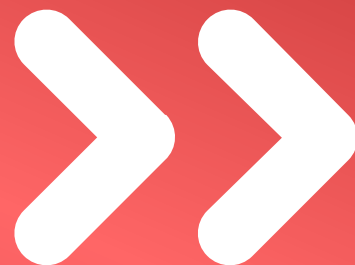
WHAT IS SELF ATTENTION?



Sentence 1: "Let's visit the bank tomorrow."
Sentence 2: "I need to deposit money at the bank."

Continued...

- Self-attention allows the model to analyze surrounding words.
- For sentence 1, "visit" and "tomorrow" suggest a location, while in sentence 2, "deposit" implies a financial context. This helps the model assign the correct meaning to "bank."



WHAT IS SELF ATTENTION?



EXAMPLE

2

Sentence : "This movie was fantastic, **NOT!**" (said with a sarcastic tone)

- Understanding sarcasm requires recognizing the contradiction between the literal meaning of the words and the intended sentiment.
- Self-attention allows the model to consider "NOT" in relation to "fantastic."
- By analyzing surrounding words and their emotional connotations, the model can pick up on the sarcastic intent.

WHAT IS SELF ATTENTION?



EXAMPLE

3

Sentence 1: "**The food wasn't bad.**" (Implies a somewhat positive sentiment)

Sentence 2: "**The food was not good.**" (Clearly expresses a negative sentiment)

- While both sentences use negation ("not" and "wasn't"), the sentiment differs.
- Self-attention helps the model understand the impact of negation based on surrounding words.



WHAT IS SELF ATTENTION?



EXAMPLE

3

Continued...

Sentence 1: "The food wasn't bad." (Implies a somewhat positive sentiment)

Sentence 2: "The food was not good." (Clearly expresses a negative sentiment)

- In sentence 1, "bad" has a negative connotation, but "wasn't" softens it.
- In sentence 2, both "not" and "good" are negative, leading to a more negative overall sentiment.

WHAT IS SELF ATTENTION?

EXAMPLE

4

Sentence: "Alice went to the store. She bought some apples."

- Here, "she" refers to Alice. Self-attention allows the model to track word relationships throughout the sentence.
- It recognizes that "Alice" is the closest noun preceding "she," establishing the connection between the pronoun and the antecedent.
- Without self-attention, the model might struggle to determine who "she" refers to if there were multiple female names mentioned earlier in the text.

WHAT IS SELF ATTENTION?



EXAMPLE

5

Sentence 1: "Let's break the ice and introduce ourselves."

(Meaning: Overcome initial awkwardness)

Sentence 2: "The ice on the lake is finally thick enough for skating."

(Literal meaning: Frozen water)

- The phrase "break the ice" has different meanings depending on the context. Self-attention helps the model disambiguate the idiom based on surrounding words

WHAT IS SELF ATTENTION?



Continued...

Sentence 1: "Let's break the ice and introduce ourselves."

(Meaning: Overcome initial awkwardness)

Sentence 2: "The ice on the lake is finally thick enough for skating."

(Literal meaning: Frozen water)

- In sentence 1, words like "introduce" and "ourselves" suggest social interaction, hinting at the figurative meaning of breaking the initial awkwardness in a new situation.



WHAT IS SELF ATTENTION?

EXAMPLE

5

Continued...

Sentence 1: "Let's break the ice and introduce ourselves."

(Meaning: Overcome initial awkwardness)

Sentence 2: "The ice on the lake is finally thick enough for skating."

(Literal meaning: Frozen water)

- In sentence 2, "lake" and "skating" point towards a literal interpretation of breaking solid ice on a frozen body of water.

WHAT IS SELF ATTENTION?



Continued...

Sentence 1: "Let's break the ice and introduce ourselves."

(Meaning: Overcome initial awkwardness)

Sentence 2: "The ice on the lake is finally thick enough for skating." (Literal meaning:

Frozen water)

- By analyzing the relationships between "break the ice" and surrounding words, self-attention allows the model to assign the appropriate meaning in each sentence.



DEFINITION

Self Attention Mechanism:

- Self-attention allows the model to **go beyond the literal meaning of individual words** and capture the nuances of language based on context.
- By considering surrounding words and their relationships, the model gains a more comprehensive understanding of the sentence's meaning.
- It allows the model to focus on relevant parts of an input sequence when processing each element.



Thank You

**SPECIAL THANKS TO CHATGPT, OPEN AI, COPILOT, GEMINI
FOR THE SUPPORT ON CONTENT**

