# Article information:

OpenAI GPT-3 Text Embeddings - Really a new state-of-the-art in dense text embeddings? | by Nils Reimers | Medium
[https://medium.com/@nils\_reimers/openai-gpt-3-text-embeddings-really-a-new-state-of-the-art-in-dense-text-embeddings-6571fe3ec9d9](https://medium.com/%40nils_reimers/openai-gpt-3-text-embeddings-really-a-new-state-of-the-art-in-dense-text-embeddings-6571fe3ec9d9)

# Article summary:

1. OpenAI released an embeddings endpoint for GPT-3 that allows users to derive dense text embeddings for a given input text.

2. Nils Reimers benchmarked the text similarity on 14 datasets and text search embeddings on 6 datasets from various domains.

3. The OpenAI text similarity models performed worse than the state of the art, and were slower and more expensive than open alternatives.

# Article rating:

Appears moderately imbalanced: The article provides some useful information, but is missing several important points or pieces of evidence that would be required to present the discussed topics in a balanced and reliable way. You are encouraged to seek a more balanced perspective on the presented issues by exploring the provided research topics and looking at different information sources.

# Article analysis:

The article by Nils Reimers provides an overview of OpenAI’s new GPT-3 Text Embeddings, which are claimed to be state-of-the-art in dense text embeddings. The author benchmarks the performance of these new GPT-3 based embeddings against open alternatives, and finds that they perform worse than the state of the art models from 2018 such as Universal Sentence Encoder, as well as smaller models with just 22M parameters that can run in your Browser. Furthermore, they are slower and more expensive than open alternatives, generating extremely high-dimensional embeddings which require much more memory and slow down downstream applications.

The article is generally reliable in its reporting of facts and findings; however, it does not provide any counterarguments or explore any potential benefits of using OpenAI’s GPT-3 Text Embeddings over open alternatives. It also does not discuss any possible risks associated with using these new embeddings or note any potential biases in their results. Additionally, while the author does mention that there are four types of models available from OpenAI (Ada, Babbage, Curie and Davinci), he does not provide any information about how these differ from each other or what advantages/disadvantages each one has over the others. Finally, while the author does mention that encoding 10 million documents with the smallest OpenAI model will cost about $80,000 compared to $1 for an equally strong open model running on cloud, he fails to mention how much operating costs would be for using OpenAI’s models versus open models for an application with 1 million monthly queries - this could be a significant factor when deciding which type of model to use for a particular task.

# Topics for further research:

* OpenAI GPT-3 Text Embeddings advantages
* OpenAI GPT-3 Text Embeddings disadvantages
. OpenAI GPT-3 Text Embeddings biases
* OpenAI GPT-3 Text Embeddings comparison
* OpenAI GPT-3 Text Embeddings operating costs
* OpenAI GPT-3 Text Embeddings model differences

# Report location:

<https://www.fullpicture.app/item/21a2b9686bd25d54d3906cde6bcc07f7>