

All Readings: Introduction to Large Language Models

Kindly note that the 30 minutes indicated on the platform considers the time that it may take you to browse through the reading resources provided. The total time required depends on the readings you decide to explore further.

Assembled readings on large language models:

- Module 2 Instruction Prompt: Link [here](#)
- Introduction to Large Language Models
<https://developers.google.com/machine-learning/resources/intro-llms>
- Language Models are Few-Shot Learners:
<https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf>
- LangChain on Vertex AI
<https://cloud.google.com/vertex-ai/generative-ai/docs/reasoning-engine/overview>
- Learn about LLMs, Gemini models, and Vertex AI
<https://cloud.google.com/vertex-ai/generative-ai/docs/learn-resources>
- Building AI-powered apps on Google Cloud databases using pgvector, LLMs and LangChain
<https://cloud.google.com/blog/products/databases/using-pgvector-llms-and-langchain-with-google-cloud-databases>
- Training Large Language Models on Google Cloud
<https://github.com/GoogleCloudPlatform/llm-pipeline-examples>
- Prompt Engineering for Generative AI
<https://developers.google.com/machine-learning/resources/prompt-eng>
- PaLM-E: An embodied multimodal language model:
<https://ai.googleblog.com/2023/03/palm-e-embodied-multimodal-language.html>
- Parameter-efficient fine-tuning of large-scale pre-trained language models
<https://www.nature.com/articles/s42256-023-00626-4>
- Parameter-Efficient Fine-Tuning of Large Language Models with LoRA and QLoRA
<https://www.analyticsvidhya.com/blog/2023/08/lora-and-qlora/>

- Solving a machine-learning mystery:
<https://news.mit.edu/2023/large-language-models-in-context-learning-0207>

Assembled readings on generative AI:

- Background: What is a Generative Model?
<https://developers.google.com/machine-learning/gan/generative>
- Gen AI for Developers
<https://cloud.google.com/ai/generative-ai?hl=en#developer-resources>
- Ask a Techspert: What is generative AI?
<https://blog.google/inside-google/googlers/ask-a-techspert/what-is-generative-ai/>
- What is generative AI?
<https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>
- Building the most open and innovative AI ecosystem:
<https://cloud.google.com/blog/products/ai-machine-learning/building-an-open-generative-ai-partner-ecosystem>
- Generative AI is here. Who Should Control It?
<https://www.nytimes.com/2022/10/21/podcasts/hard-fork-generative-artificial-intelligence.html>
- Stanford U & Google's Generative Agents Produce Believable Proxies of Human Behaviors:
<https://syncedreview.com/2023/04/12/stanford-u-googles-generative-agents-produce-believable-proxies-of-human-behaviours/>
- Generative AI: Perspectives from Stanford HAI:
https://hai.stanford.edu/sites/default/files/2023-03/Generative_AI_HAI_Perspectives.pdf
- Generative AI at Work:
https://www.nber.org/system/files/working_papers/w31161/w31161.pdf
- The future of generative AI is niche, not generalized:
<https://www.technologyreview.com/2023/04/27/1072102/the-future-of-generative-ai-is-niche-not-generalized/>
- The implications of Generative AI for businesses:
<https://www2.deloitte.com/us/en/pages/consulting/articles/generative-artificial-intelligence.html>
- Proactive Risk Management in Generative AI:
<https://www2.deloitte.com/us/en/pages/consulting/articles/responsible-use-of-generative-ai.html>
- How Generative AI Is Changing Creative Work:
<https://hbr.org/2022/11/how-generative-ai-is-changing-creative-work>

Additional Resources:

- Attention is All You Need: <https://research.google/pubs/pub46201/>
- Transformer: A Novel Neural Network Architecture for Language Understanding: <https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html>
- Transformer on Wikipedia: [https://en.wikipedia.org/wiki/Transformer_\(machine_learning_model\)#:%~:text=Transformers%20were%20introduced%20in%202017,allowing%20training%20on%20larger%20datasets](https://en.wikipedia.org/wiki/Transformer_(machine_learning_model)#:%~:text=Transformers%20were%20introduced%20in%202017,allowing%20training%20on%20larger%20datasets)
- What is Temperature in NLP? <https://lukesalamone.github.io/posts/what-is-temperature/>
- Model Garden: <https://cloud.google.com/model-garden>
- Auto-generated Summaries in Google Docs: <https://ai.googleblog.com/2022/03/auto-generated-summaries-in-google-docs.html>