CS6220 (2023) Reading List 1

Each title is linked to the URL where you can download the paper (click on the title). You are asked to choose 2 papers of your favorite subject. We recommend that you choose two papers on similar topics for better learning experience on a specific subject.

* [1-5] are about MapReduce (popular reads)
* [6-9] are about Prompt Engineering for Pretrained Large Language Models (LLMs).

1. Jeffrey Dean and Sanjay Ghemawat. [MapReduce: Simplified Data Processing on Large Clusters.](http://static.googleusercontent.com/media/research.google.com/en/us/archive/mapreduce-osdi04.pdf) In Proceedings of OSDI 2004, San Francisco, CA, 2004
2. Fay Chang et. Al. [Bigtable: A Distributed Storage System for Structured Data](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/archive/bigtable-osdi06.pdf) In Proceedings of OSDI 2006, Seattle, WA, 2006. <http://hadoop.apache.org/>
3. Mike Burrows. [The Chubby lock service for loose-coupled distributed systems.](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/archive/chubby-osdi06.pdf)In 7th USENIX Symposium on Operating Systems Design and Implementation (OSDI), 2006.
4. Giuseppe DeCandia et. Al. [Dynamo: Amazon’s Highly Available Key-value Store.](http://www.allthingsdistributed.com/files/amazon-dynamo-sosp2007.pdf) In SOSP ‘07
5. Cheng-Tao Chu et. Al. [Map-Reduce for Machine Learning on Multicore.](http://www.cs.stanford.edu/people/ang/papers/nips06-mapreducemulticore.pdf)NIPS’06.
6. [Language Models are Few-Shot Learners](https://arxiv.org/abs/2005.14165)
7. [Chain-of-Thought Prompting Elicits Reasoning in Large Language Models](https://arxiv.org/abs/2201.11903), Jason Wei, et.al., 2022
8. [Tree of Thoughts: Deliberate Problem Solving with Large Language Models](https://arxiv.org/abs/2305.10601)
9. [Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks.](https://arxiv.org/pdf/2005.11401)