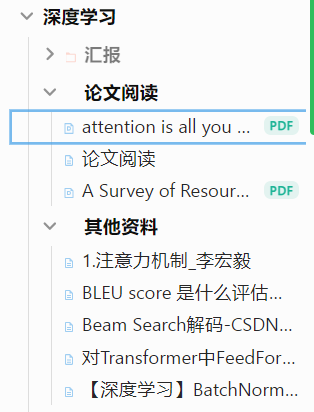
* 看了前半部分A Survey of Resource-efficient LLM and Multimodal Foundation Models，到注意力机制 https://arXiv:2401.08092v2
* 去看attention is all you need 论文 https://arxiv.org/abs/1706.03762
* 然后去看了李宏毅老师的 注意力机制
* 主要资料



* 论文笔记见附件