* Positional Embedding
* Positional Encoding
* Self-Attention
* Greedy Decoding
* Random Sampling (top-k and top-p)
* Domain Specific Training -> BloombrgGPT paper
* ROUGE and BLEU score (Eval Matrices)