Dataset VietNews

1. Tải & tiền xử lý bằng VnCoreNLP hoặc PyVi

2. thiết lập một pipeline trích rút gồm TF-IDF→TextRank→PhoBERT-Cosine.

3 đo ROUGE-1/2/L

4. viết báo cáo + demo Notebook.

Transformer (Encoder)

1. Multi-headed Attention

1.1 Self-Attention

- An input associates with others input

- Feed an input into 3 fully-connected layers to crates query, key and value vectors

Metric

ROUGE-Score