Pegasus model:

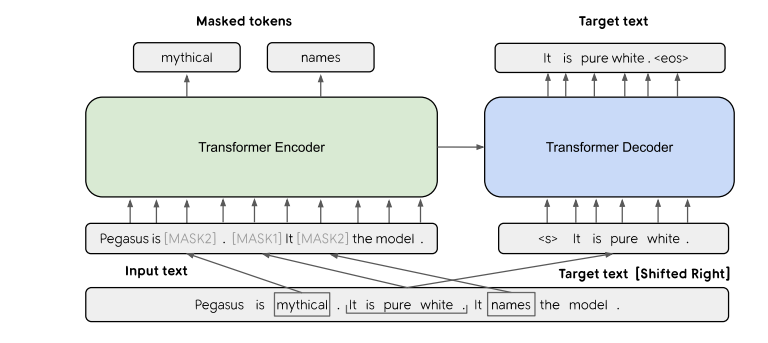
This is a standard transformer based model, with encoder and decoder architecture.

This model is pre-trained on objective called “Gap sentence prediction”. This is specifically aimed for Text summarization. Hypothesis is that using a pre-training objective that more closely resembles the downstream task leads to better and faster fine-tuning performance

Gap sentence prediction: PEGASUS randomly selects a subset of sentences and removes them from the input. These sentences are masked, i.e replaced my masked token.

The remaining sentences in the document are provided as context for the model. The goal is for the model to learn to reconstruct the missing sentences based on the context of the remaining sentences.

The model tries to predict the removed sentences from the context. The objective is to minimize the loss between the predicted sentences and the actual ones



* This model is pretrained on huge data using “Gap sentence prediction” for abstract summarization.
* While fine tuning, there is no “Gap sentence prediction” done.

1. Articles are passed to the encoder .
2. Encoder generates context embeddings
3. Starting sequence and the encoder context embedding is passed to the decoder.

ROUGE score:

ROUGE score measures the similarity between the machine-generated summary and the reference summaries using overlapping n-grams.

ROUGE = ∑ (Recall of n-grams)

Where:

* Recall of n-grams is the number of n-grams that appear in both the machine-generated summary and the reference summaries divided by the total number of n-grams in the reference summaries.

'rouge1': 0.2594917893942396, 'rouge2': 0.10725189124774678