**Abstractive Text Summarization**

**Program flow:**

1. **Data preparation:** Cleaning and transforming data into text, summary pair and the same is split into train, validate, and test.
2. **Pre-training**: Pre-trained PEGASUS model is imported into the Google Colab environment with TPU/GPU runtime instance.
3. **Import Fine tuning Data:** Fine-tuning data (WikiHow Knowledge Base/ Amazon Fine Food Reviews data) is imported into Google Colab environment.
4. **Convert Python Data frames into .TFRecord\*:** Transform the imported fine-tuning data into .TFRecord data to optimize the computing efficiency of the model during fine-tuning.
5. **Register fine-tuning dataset:** Register .TFRecord\* fine-tuning dataset in the

cd/content/PEGASUS/params/public\_params.py

1. **Fine-tuning PEGASUS:** The registered training dataset is used to fine-tune the model.
2. **Evaluating the model:** The fine-tuned model is then evaluated on the registered test data. The predictions of the model and the summarization statistics can be found in the below path**.**

/content/PEGASUS/ckpt/pegasus\_ckpt/WikiHow

/content/PEGASUS/ckpt/pegasus\_ckpt/amazon