The term Transfer learning refers to the situation where a model is pre-trained on a data-rich dataset for one task and can be fine-tuned for specific downstream tasks. This is a very powerful technique that promotes collaboration and growth. Since training large models for each and every specific task is computationally and financially expensive, and it is often done using advanced GPUs and TPUs.The most important part of transfer learning is when a single standalone model pre-trained on a mixture of data-rich tasks can be used for all text-processing tasks.T5 is the first model to achieve this across various tasks (language translation, text summarization, text classification, etc.).The T5 Transformer Model was introduced in 2020 by the Google AI team and stands for Text-To-Text Transfer Transformer. T5 introduced the “Text-to-Text” framework, in which every NLP task (Translation, Classification, etc.) has the same underlying structure in which text is fed as input to the model and text is produced as output. This means we can use the same model, the same hyperparameters, and the same loss function across all the tasks.
The T5 transformer model works by using the same standard encoder-decoder structure as standard transformer models. It consists of 12-pair blocks of encoder-decoder.