BART MODEL

BART stands for Bidirectional and Auto-Regressive Transformers. It is a denoising autoencoder for pretraining sequence-to-sequence models. It is trained by corrupting text with an arbitrary noising function, and learning a model to reconstruct the original text. It uses a standard Transformer-based neural machine translation architecture with a bidirectional encoder (like BERT) and a left-to-right decoder (like GPT). This means the encoder’s attention mask is fully visible, like BERT, and the decoder’s attention mask is causal, like GPT2.

BART has been used for various tasks like summarization, abstraction questions answering and others.