

BART (Bidirectional and Auto-Regressive Transformers) is a language generation model developed by Facebook AI Research (FAIR) in 2019. BART is a pre-training sequence-to-sequence model that is capable of producing high-quality text in various tasks, such as summarization, machine translation, and text generation.

The BART model was developed by a team of researchers from FAIR, including Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. The development time for BART is not explicitly known, but it was first introduced in a research paper in 2019.

The BART model was pre-trained on a large corpus of text, including Wikipedia and web pages, and then fine-tuned on specific tasks using various datasets. The pre-training corpus contains approximately 800 million words, which is smaller than some other pre-trained models like GPT-3, but still large enough to capture the statistical patterns in natural language.

The BART architecture is based on the Transformer architecture, which is a type of neural network that is particularly well-suited for natural language processing tasks. BART is unique because it combines both auto-regressive and bidirectional components, which allows it to model both left and right contexts of the input sequence. This is achieved by using a denoising autoencoder objective during pre-training, which requires the model to generate the original input sequence from a corrupted version of the sequence.

One of the key features of BART is its ability to perform well on various language generation tasks. It has achieved state-of-the-art results on many benchmark datasets, such as the CNN/Daily Mail summarization dataset and the WMT14 English-to-German machine translation task. BART is also very flexible and can be fine-tuned on many different tasks with only a small amount of task-specific training data.

In terms of evaluation, the BART model is typically evaluated on standard benchmark datasets and compared against other state-of-the-art models. The quality of the generated text is measured using various metrics, such as BLEU score, ROUGE score, and human evaluation.

Overall, the BART model is a powerful language generation model that is capable of generating high-quality text for a variety of tasks. Its unique combination of bidirectional and auto-regressive components, along with its flexibility and ability to perform well on various tasks, make it a popular choice for researchers and practitioners in the field of natural language processing.