

TEXT AND LANGUAGE GENERATION

Overview of Language Models

Definition: Models designed to understand and generate human language for diverse applications like translation, summarization, and conversation.

Key Models

GPT (Generative Pre-trained Transformer)

Definition: An autoregressive model that generates text by predicting the next word in a sequence.

BERT (Bidirectional Encoder Representations from Transformers)

Definition: A bidirectional model that learns context from both directions, enhancing text understanding.

Training Language Models

Data: Large, diverse text datasets to cover varied language contexts.

Loss Functions

Cross-Entropy Loss: Measures the accuracy of next-word predictions.
Masked Language Modeling (MLM) Loss: Used to predict masked words, helping the model learn bidirectional context.

Optimization: Algorithms like gradient descent, commonly using the Adam optimizer, to improve model accuracy.

Evaluation Metrics for Text Generation

Perplexity: Measures model prediction accuracy; lower perplexity indicates better performance.

BLEU (Bilingual Evaluation

Understudy): Compares generated and reference text to measure similarity.
ROUGE (Recall-Oriented Understudy for Gisting Evaluation): Assesses overlap in words between generated and reference texts.

Human Evaluation: Essential for assessing text coherence, fluency, and creativity.

Applications of GPT and BERT

GPT Applications

Creative Writing: Automates storytelling and content creation.
Dialogue Generation: Powers chatbots and conversational AI.

BERT Applications

Information Retrieval: Enhances relevance in search engines.
Sentiment Analysis: Assesses emotions in text for feedback analysis.

Summarization: Extracts key ideas from documents and articles.

Hands-on: Implementing a Text Generation Model

Libraries: Commonly uses Hugging Face's transformers library, PyTorch, or TensorFlow.

Sample Code:

Basic code for generating text with GPT-2, ideal for hands-on practice.