Group Final project: Preparation

8-Hengke Jiang, 11-Rock, 13-Yuting Jiang, 30-Ethan, 35-Jade Lin

linyh63@mail2.sysu.edu.cn

Table of Contents

Natural Language Processing(NLP)

History

Before the invention of computer, Alan Turing published an article titled "Computing Machinery and Intelligence" which proposed what is now called the Turing test as a criterion of intelligence.(1940)

With computer, machine translation(MT) and Statistical machine translation(SMT).(1950-1900, 1990-2010)

Multi-layer perceptron(MLP, 2003), recurrent neural network(RNN, 2010)...

Table of Contents

$$\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V$$

Transformers use self-attention mechanisms to process input sequences in parallel, making them highly efficient and effective.(6.2017[4])

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t\mathbf{I})$$

Base on Denoising Diffusion Probabilistic Models(2020, DDPM[2]) diffusion(2021, [1]) and continuous diffusions(2022, Diffusion-LM[3]) are introduced to improve non-autoregressive(NAR) text generation.

Bibliography

[1] Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg.
Structured denoising diffusion models in discrete state-spaces, 2023.

[2] Jonathan Ho, Ajay Jain, and Pieter Abbeel.
Denoising diffusion probabilistic models, 2020.

[3] Xiang Lisa Li, John Thickstun, Ishaan Gulrajani, Percy Liang, and Tatsunori B. Hashimoto.
Diffusion-lm improves controllable text generation, 2022.

[4] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin.
Attention is all you need, 2023.