

Bert

1

EMLO GPT

- EMLO~~recurrent~~**RNN**
 - GPT~~recurrent~~**transformer**

Bert transformer

1

1. NSP
• NSP
 2. Bert
• Bert

A decorative horizontal bar consisting of a series of small, evenly spaced rectangles.

1

1. မြန်မာစာMLM
 - မြန်မာစာမျက်နှာမျက်နှာမျက်နှာ
 - ၁၅%မြန်မာစာ၏ ၈၀%မြန်မာစာမျက်နှာမျက်နှာ၏ ၁၀%မြန်မာစာ၏ ၁၀%မြန်မာစာ
 - မြန်မာစာbertမျက်နှာမျက်နှာမျက်နှာမျက်နှာ
 2. မြန်မာစာNSP
 - မြန်မာစာမျက်နှာမျက်နှာမျက်နှာမျက်နှာ
 - ၅၀%မြန်မာစာ၏ ၅၀%မြန်မာစာ

1

$\boxed{\text{[CLS]}}$ = Token Embedding + Segment Embedding + Position Embedding

- **Token Embedding** WordPiece 通过将 "unhappiness" → "un", "#happiness" 的方法实现。
 - **Segment Embedding** 将每个子词映射到一个 0 到 1 之间的值。
 - **Position Embedding** 将位置映射到一个 512 维的 Token 上。

1

WordPiece

□□□□□□□□□□

-
 -

□□□□□□□□□□ # # □□□□ # # i n g □□□□□□□□□□

မြတ်စွာ

1. မြတ်စွာနေရာတွင် အမြတ်စွာ ပေါ်လေ့ရှိသူများ
2. မြတ်စွာနေရာတွင် အမြတ်စွာ ပေါ်လေ့ရှိသူများ

$\text{score}(s) = \log P(s) - (\log P(s_1) + \log P(s_2))$

မြတ်စွာ

- $P(s)$ မြတ်စွာနေရာတွင် ပေါ်လေ့ရှိသူများ
- $P(s_1)$ နှင့် $P(s_2)$ မြတ်စွာနေရာတွင် ပေါ်လေ့ရှိသူများ