



# **Нейросетевые языковые модели BERT**

Евгений Борисов

# Нейросетевые языковые модели



# Нейросетевые языковые модели

## BERT (Bidirectional Encoder Representation Transformers)

Devlin J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding – 2018.

Языковая модель, основанная на Transformer Encoder

BERT Base – 12 блоков

BERT Large - 24 блока

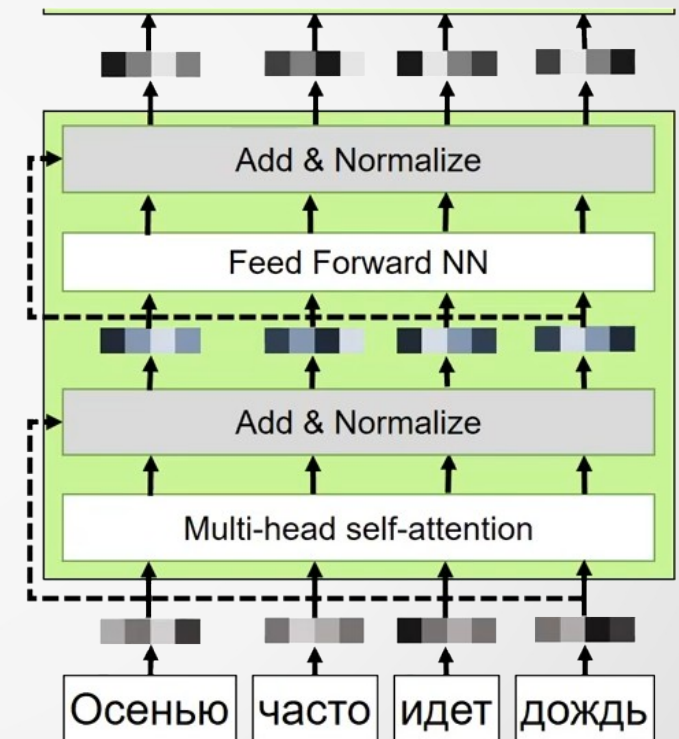
применяем обучаемый Positional Encoding

данные для обучение разбиты на пары предложений

обучаем модель определять связаны ли два предложения

маскируем слова и обучаем модель их предсказывать

$12(24) \times$

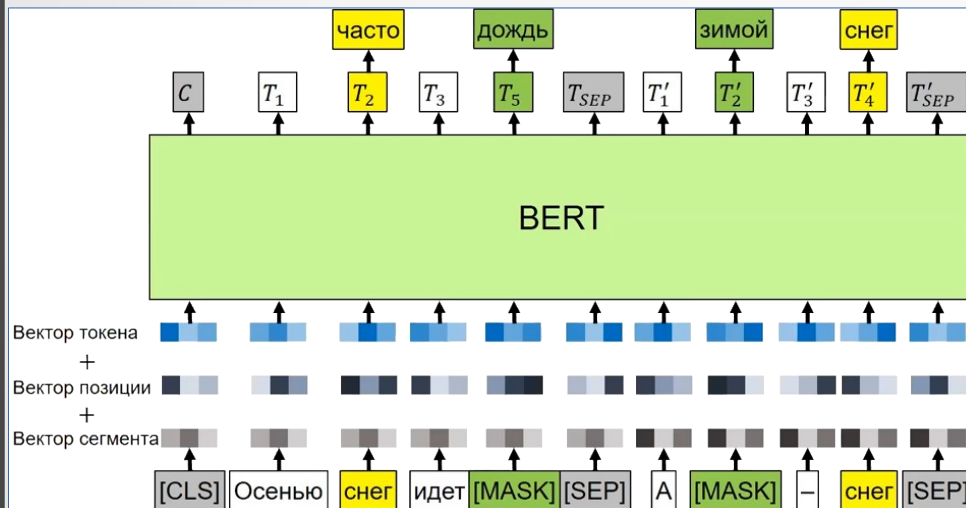


# Нейросетевые языковые модели

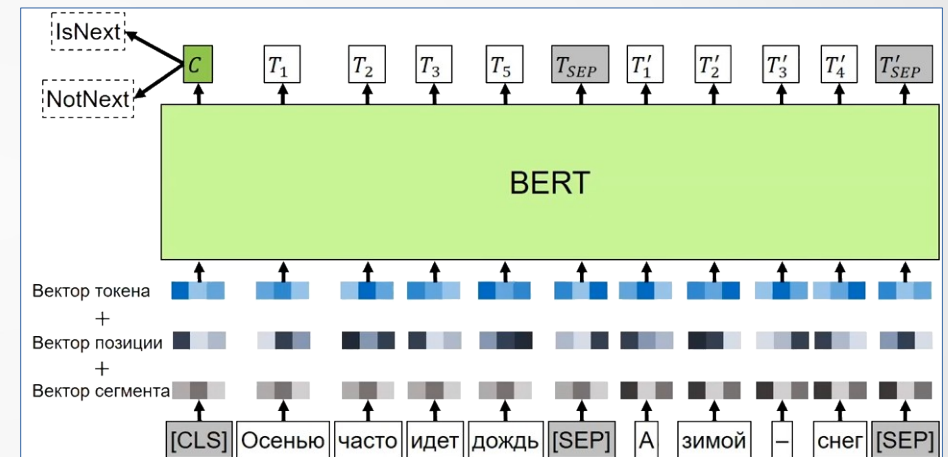
## BERT (Bidirectional Encoder Representation Transformers)

Devlin J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding – 2018.

- обучаемый Positional Encoding + дополнительный вектор номера сегмента



- маскируем слова и обучаем модель их предсказывать
- 15% токенов могут маскироваться, из них:
  - 80% - [MASK];
  - 10% - подменяем на случайные;
  - 10% - оставляем;



- обучаем модель определять связаны ли сегменты (опционально)

# Литература

## Литература

Борисов Е.С. Методы машинного обучения. 2024  
[https://github.com/mechanoid5/ml\\_lectorium\\_2024\\_I](https://github.com/mechanoid5/ml_lectorium_2024_I)

Борисов Е.С. Методы обработки текстов на естественном языке. 2024  
[https://github.com/mechanoid5/ml\\_nlp\\_2024\\_I](https://github.com/mechanoid5/ml_nlp_2024_I)

Майоров В.Д. Основы обработки текстов. 10. Языковые модели. ИСП РАН, 2021  
[https://www.youtube.com/watch?v=\\_8MGdpt4I9M](https://www.youtube.com/watch?v=_8MGdpt4I9M)

Тихомиров М.М. Основы обработки текстов. 14. Большие языковые модели. ИСП РАН, 2023  
[https://www.youtube.com/watch?v=EC6\\_rMs1vsY](https://www.youtube.com/watch?v=EC6_rMs1vsY)

Нейчев Радослав Self-Attention. Transformer overview. Лекторий ФПМИ, 2020  
<https://www.youtube.com/watch?v=UETKUIIYE6g>

Jay Alammar Transformer в картинках. (Перевод - Е.Смирнова, С. Шкарин)  
<https://habr.com/ru/articles/486358/>

Jay Alammar GPT-2 в картинках. (Перевод - Е.Смирнова, С. Шкарин)  
<https://habr.com/ru/articles/490842/>