

Модели-декодировщики

[Ask a question](#)

Декодировщики используют только компонент декодирования трансформера. На каждом этапе для текущего слова слой внимания может получить доступ только к словам, которые были расположены до него в предложении. Такие модели часто называются *авторегрессионными моделями*.

Процесс предобучения декодировщиков обычно заключается в предсказании следующего слова в предложении.

Такие модели лучше всего подходят для задач, связанных с генерацией текста.

Представителями этого семейства моделей являются:

- CTRL
- GPT
- GPT-2
- Transformer XL