

Модели-кодировщики

[Ask a question](#)

Кодировщики используют только компонент кодировщика трансформера. На каждом этапе слой внимания может использовать все слова исходного предложения. Эти модели часто характеризуют как имеющие двунаправленное внимание (англ. bi-directional attention), и часто называют моделями *автокодировщиками*.

Предварительное обучение этих моделей обычно заключается в том, чтобы как-то исказить предложение (например, путем маскировки в нем случайных слов) и поставить перед моделью задачу найти или восстановить исходное предложение.

Кодировщики лучше всего подходят для задач, требующих *понимания* всего предложения, таких как классификация предложений, распознавание именованных сущностей (и, в более общем смысле, классификация слов) и ответы на вопросы с извлечением информации из контекста (выделительные вопросно-ответные системы).

К представителям этого семейства моделей относятся:

- ALBERT
- BERT
- DistilBERT
- ELECTRA
- RoBERTa