
УДК 004.8 Жусип М.Н., Жаксыбаев Д.О.

Жусип М.Н.

магистр технических наук, преподаватель

Западно-Казахстанский аграрно-технический университет

им. Жангир хана

(г. Уральск, Казахстан)

Жаксыбаев Д.О.

PhD, и.о. доцента

Западно-Казахстанский аграрно-технический университет

им. Жангир хана

(г. Уральск, Казахстан)

СРАВНЕНИЕ ЧАТ-БОТОВ С ИСПОЛЬЗОВАНИЕМ ТРАНСФОРМЕРОВ И НЕЙРОСЕТЕЙ: ИССЛЕДОВАНИЕ ПРИМЕНЕНИЯ АРХИТЕКТУР GPT И BERT

***Аннотация:** в данной статье рассматривается сравнение двух популярных архитектур, используемых для создания чат-ботов, — GPT (Generative Pre-trained Transformer) и BERT (Bidirectional Encoder Representations from Transformers). Оба подхода основаны на трансформерах, но применяются по-разному: GPT фокусируется на генерации текста, а BERT — на глубоком понимании контекста и классификации.*

***Ключевые слова:** чат-боты, трансформеры, нейронные сети, обработка естественного языка, искусственный интеллект.*

Современные технологии искусственного интеллекта (ИИ) активно развиваются в области обработки естественного языка (NLP), что делает чат-боты важным инструментом взаимодействия с пользователями в различных сферах: от обслуживания клиентов до образования и медицины. Основными архитектурами, используемыми для создания продвинутых чат-ботов, являются

трансформеры, такие как GPT (Generative Pre-trained Transformer) и BERT (Bidirectional Encoder Representations from Transformers). Эти модели нейросетей существенно отличаются по подходам к обучению и применению.

Архитектура и принципы работы. GPT и BERT — это две популярные архитектуры трансформеров, которые по-разному подходят к обработке текста и решению задач обработки естественного языка (NLP). Оба алгоритма стали основополагающими для развития чат-ботов, но их внутренние механизмы и области применения существенно различаются.

GPT — это генеративная модель, разработанная компанией OpenAI. Основная особенность GPT заключается в том, что она использует однонаправленную архитектуру: модель анализирует текст слева направо, то есть предсказывает следующее слово на основе предыдущих токенов. Модель обучена на огромных корпусах текстов, что позволяет ей генерировать логически связанные и разнообразные ответы.

Ключевыми особенностями GPT являются:

- Генеративная природа: GPT генерирует текст на основе входного запроса, что позволяет создавать ответы в реальном времени. Чат-боты на основе GPT могут поддерживать сложные, открытые диалоги, где возможны множество вариантов развития беседы.
- Самообучение: Модель использует механизм self-attention, который позволяет ей учитывать контекст предыдущих слов для более точного предсказания следующего токена.
- Широкое применение: GPT подходит для задач, где требуется генерировать уникальный текст, таких как создание статей или ответы в чатах.

Однако GPT имеет и свои недостатки. Поскольку модель фокусируется на генерации, иногда её ответы могут быть слишком общими или не полностью корректными в специфических задачах, требующих точного анализа запроса.

BERT, разработанный Google, представляет собой двунаправленную архитектуру, которая анализирует текст как слева направо, так и справа налево. Это позволяет модели учитывать полный контекст слова в предложении, а не

только его предыдущие токены. В отличие от GPT, BERT не генерирует текст, а предназначен для задач, связанных с классификацией и пониманием текста, таких как определение намерений пользователя или выбор правильного ответа на запрос.

Ключевые особенности BERT:

- **Двунаправленность:** BERT анализирует контекст с обеих сторон, что улучшает понимание значений слов в зависимости от их окружения. Это особенно полезно в задачах, где важно точное понимание смысла.
- **Предназначение для классификации:** BERT чаще всего используется для ранжирования ответов и определения наилучшего из них, а не для генерации текста.
- **Маскированное языковое моделирование:** BERT обучается, "маскируя" некоторые слова в предложении и пытаясь их восстановить, что помогает ему глубже анализировать взаимосвязь слов в контексте.

Недостатком BERT является то, что эта модель не способна генерировать текст, а только выбирать из предложенных вариантов. Это ограничивает её использование в чат-ботах, которые требуют большего разнообразия ответов.

Сравнение производительности. В контексте обработки диалогов обе архитектуры проявляют себя по-разному. GPT превосходит в задачах, связанных с генерацией текста и динамическими диалогами. Это делает его идеальным для чат-ботов, которым нужно поддерживать длительные и разнообразные беседы. BERT, с другой стороны, лучше подходит для точного понимания запросов и ранжирования ответов, что делает его незаменимым в задачах, связанных с обслуживанием клиентов или где важна точность ответа.

Для достижения наилучших результатов часто используются гибридные решения. В таких системах BERT используется для анализа и понимания запроса пользователя, а GPT — для генерации ответа на основе этого анализа. Это позволяет создавать более сбалансированные чат-боты. В будущем комбинированное использование этих архитектур может значительно улучшить качество чат-ботов и расширить их применение в различных сферах.

СПИСОК ЛИТЕРАТУРЫ:

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, 30, 5998-6008;
2. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners. *OpenAI Blog*. Retrieved from <https://openai.com/research/language-models-are-unsupervised-multitask-learners>;
3. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692*

Zhussip M.N., Zhaxybayev D.O.

Zhussip M.N.

West Kazakhstan Agrarian-Technical University named after Zhangir Khan
(Uralsk, Kazakhstan)

Zhaxybayev D.O.

West Kazakhstan Agrarian-Technical University named after Zhangir Khan
(Uralsk, Kazakhstan)

COMPARISON OF CHATBOTS USING TRANSFORMERS AND NEURAL NETWORKS: APPLICATION STUDY OF GPT AND BERT

Abstract: paper discusses a comparison of two popular architectures used to create chatbots, GPT (Generative Pre-trained Transformer) and BERT (Bidirectional Encoder Representations from Transformers). Both approaches are based on Transformers but are applied in different ways: GPT focuses on text generation, while BERT focuses on deep context understanding and classification.

Keywords: chatbots, GPT, BERT, transformers, neural networks, natural language processing, artificial intelligence.