ClosedAI

“Математика является учением об отношениях между понятиями, лишенными какого бы то ни было содержания.” **(Давид Гильберт)**

Тензорное разложение как инструмент сжатия моделей трансформеров

Герман Корж: разложения Таккера, TT, бонус

Степан Масчан: каноническое разложение

Владислав Морозов: дизайн эксперимента, оформление

Тензорное разложение

В полилинейной алгебре тензорная декомпозиция - это любая схема для выражения "тензора данных" в виде последовательности элементарных операций, действующих на другие, часто более простые тензоры. Многие тензорные разложения обобщают матричные разложения.

Наиболее распространенные разложения

- **Каноническое разложение** - запись тензора в виде суммы скелетонов.
- **Разложение Такера** - SVD-разложения вдоль каждой размерности и объединение результатов.
- **Иерархическое разложение Таккера** (HT)
- **TT-разложение** - тензорный поезд
- **Квантизованные тензорные аппроксимации** (формат QTT)

ClosedAI

Каноническое разложение

$$a_{ijk} = \sum_{\alpha=1}^{r} u_{i\alpha} v_{j\alpha} w_{k\alpha}$$

Свойства:

- для d-мерного тензора требуется хранить nrd элементов
- Единственно при условии несильных ограничений
- Отсутствует устойчивый алгоритм для вычисления наилучшей аппроксимации ранга r

ClosedAI

Разложение Таккера

$$a_{ijk} = \sum_{\alpha_1,\alpha_2,\alpha_3=1}^{r_1,r_2,r_3} g_{\alpha_1\alpha_2\alpha_3} u_{i\alpha_1} v_{j\alpha_2} w_{k\alpha_3}.$$

Свойства:

- Устойчивый алгоритм (SVD)
- Экспоненциальный рост

ClosedAI

Разложение TT

$$a_{i_1 i_2 \ldots i_d} = \sum_{\alpha_1, \ldots, \alpha_{d-1}} g_{i_1 \alpha_1} g_{\alpha_1 i_2 \alpha_2} \cdots g_{\alpha_{d-2} i_{d-1} \alpha_{d-1}} g_{\alpha_{d-1} i_d}$$

Свойства:

- Устойчивый TT-SVD алгоритм
- Требуется памяти dnr^2

ClosedAI

Дизайн эксперимента

- MarianMT
- BLUE
- Tatoeba
-

Каноническое, Таккера, TT

ClosedAI

MarianMT

MarianMT (Marian Machine Translation) — это набор моделей нейронного машинного перевода, разработанных в рамках проекта Marian, который активно поддерживается исследовательской группой на Facebook AI Research (FAIR).

Основные характеристики:

- Многоязычная поддержка
- Современные методы перевода
- Быстрая работа
- Гибкость и настраиваемость

ClosedAI

BLEU (Bilingual Evaluation Understudy)

Одна из наиболее широко используемых автоматических оценок для оценки качества машинного перевода. Предложена в 2002 году. В настоящий момент стала стандартом в области оценки машинного перевода.

Характеристики

- оценивает качество перевода, сравнивая автоматический перевод с одним или несколькими эталонными (реальными) переводами, произведенными человеком.
- Основной принцип работы BLEU заключается в использовании n-грамм — последовательностей из n слов.
- использует механизм, который ограничивает количество совпадений, учитываемое для каждой n-граммы. Например, если слово "дом" встречается трижды в реальном переводе, а в машине — пять, для подсчета будет учитываться только три совпадения.

ClosedAI

Tatoeba(от японского татоэба (яп. 例えば) «например»)

крупная многоязычная платформа, созданная для изучения языков и упрощения доступа к примерам использования языковых структур. Она основана на открытых данных и является сообществом, вовлеченным в создание и расширение базы данных с примерами предложений на различных языках.

https://tatoeba.org/

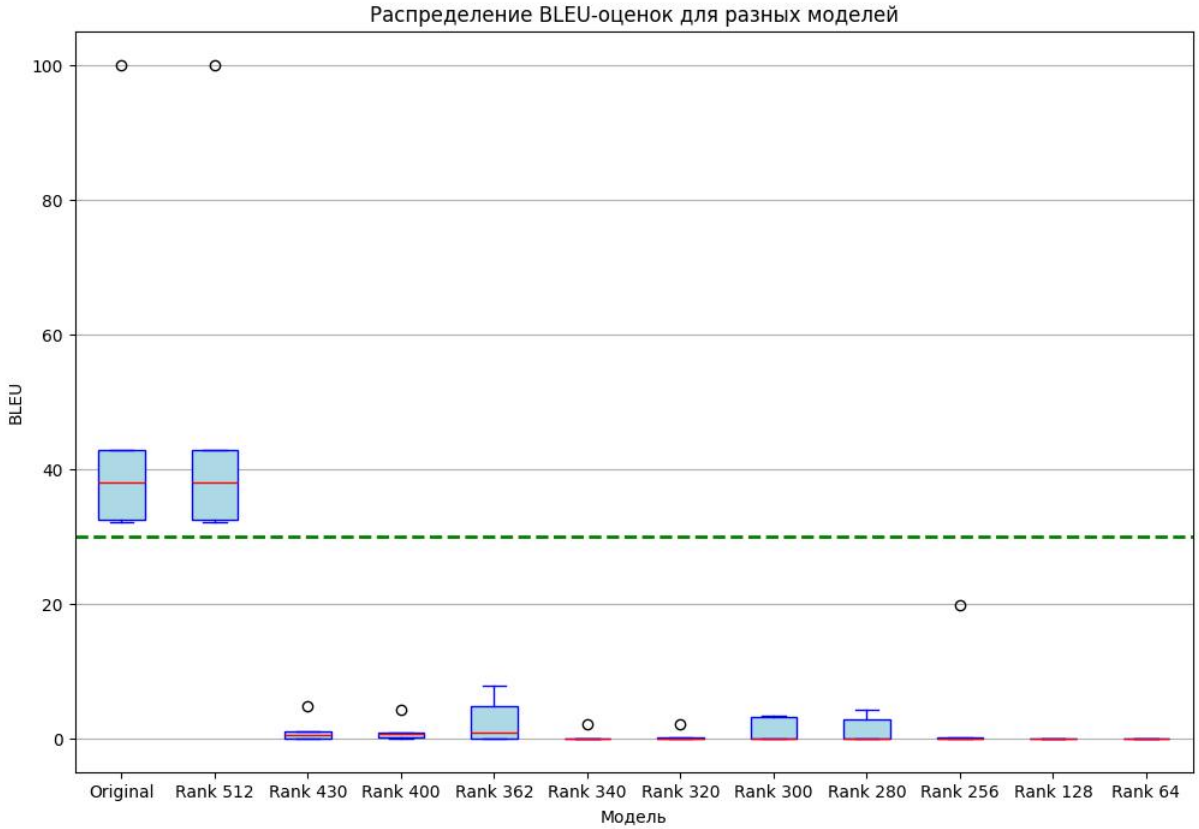
ClosedAI

“Когда я ходил купаться в шторм и меня уносило далеко от берега, то становилось безумно жалко младшую сестренку, которую я обижал в детстве.” (**Джером К. Джером**)

Результаты и выводы

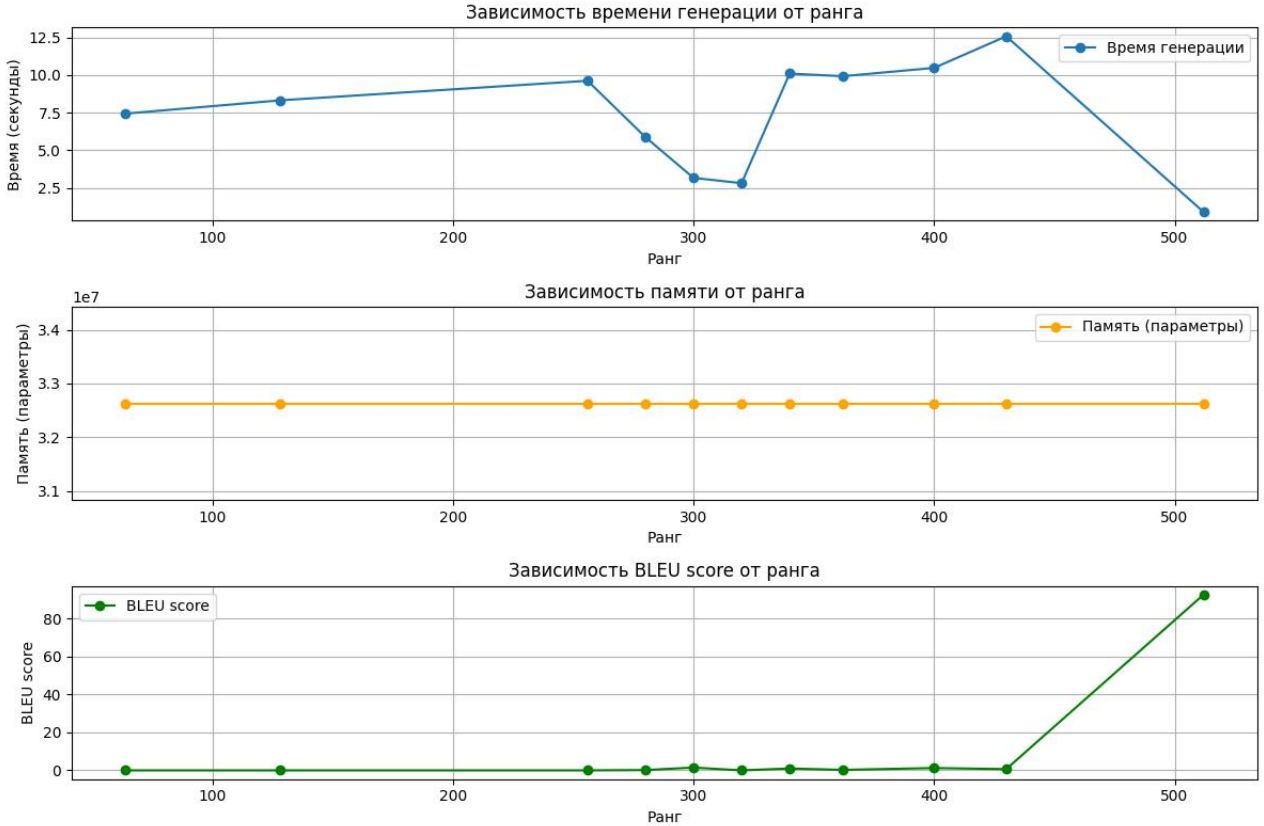
ClosedAI

Каноническое разложение



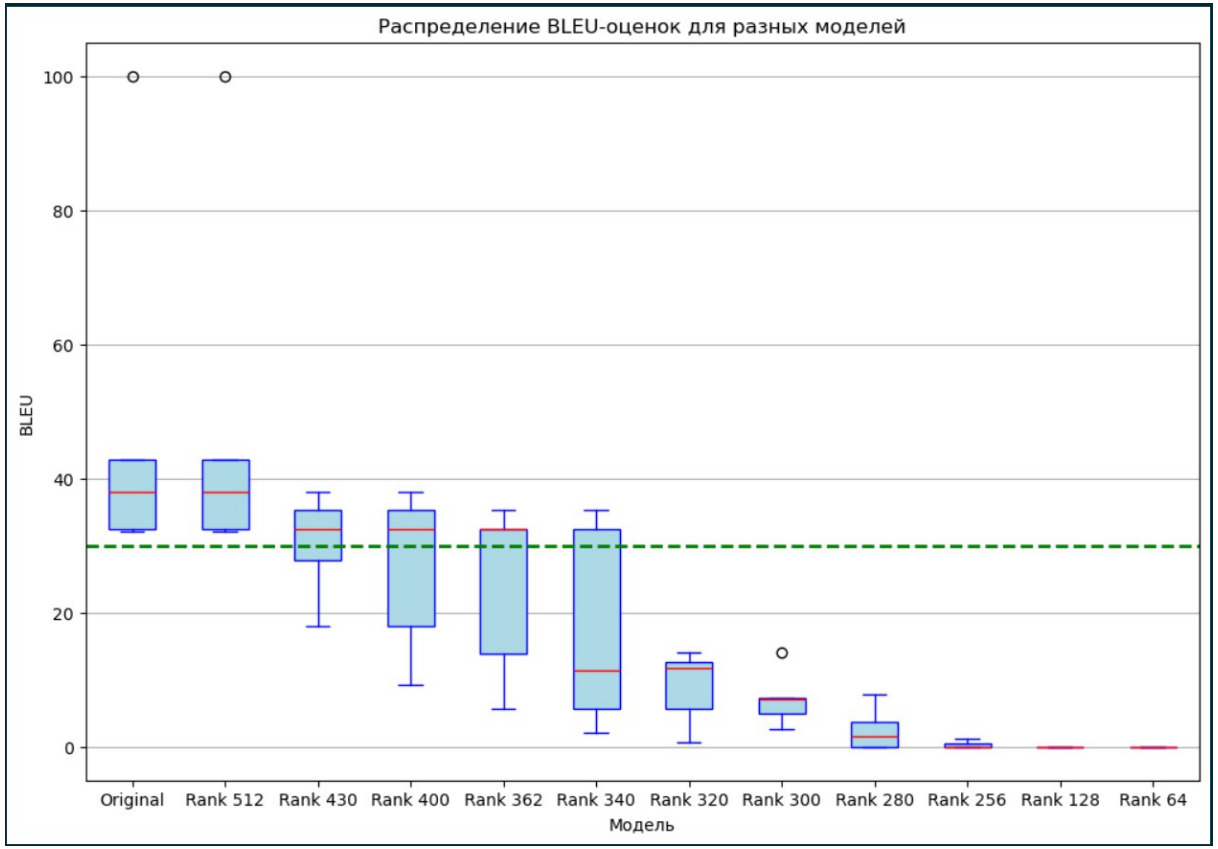
ClosedAI

Каноническое разложение



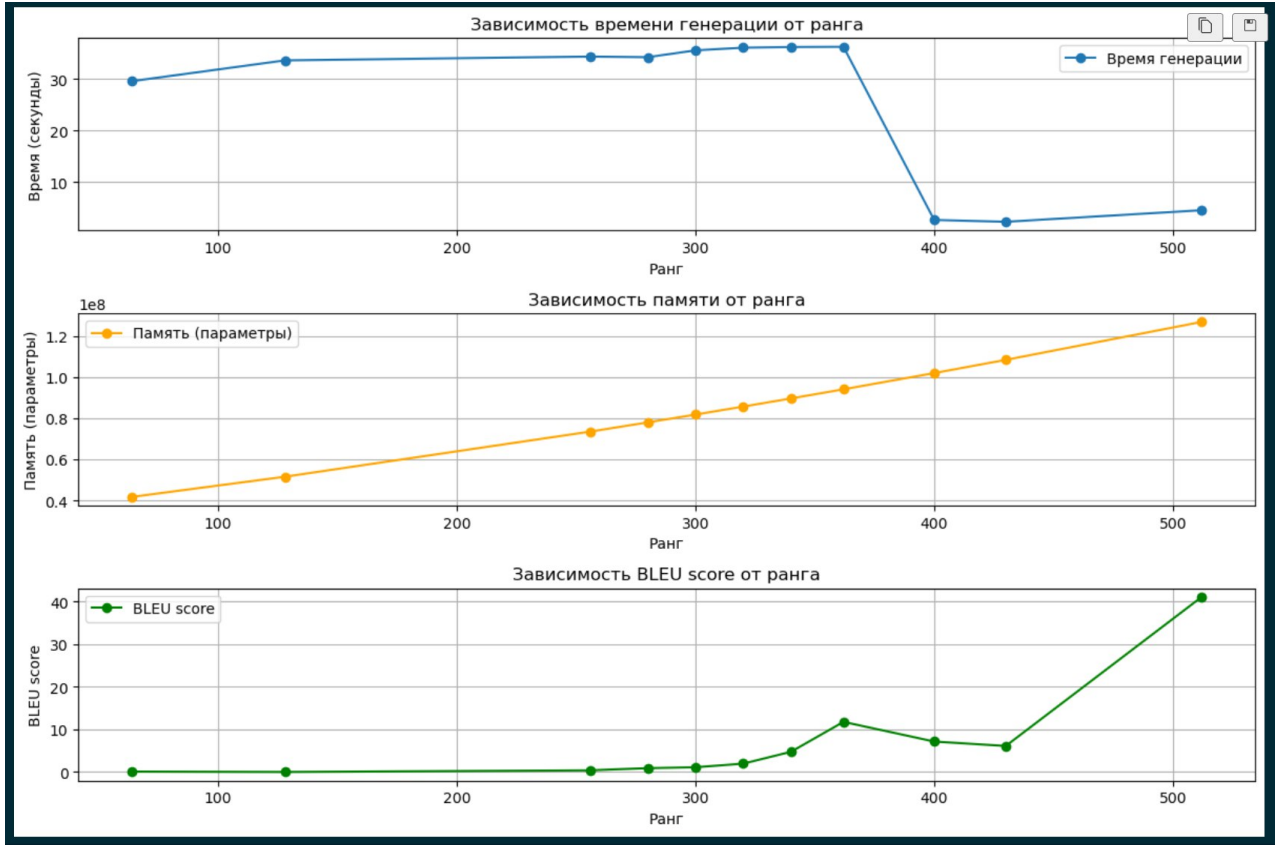
ClosedAI

Разложение Таккера



ClosedAI

Разложение Таккера



ClosedAI

Разложение TT

ClosedAI

Разложение TT

ClosedAI

“Зимой контуры чужой жизни более отчетливы.
Для путешественника это — бонус”
(Иосиф Бродский)

Бонус

ClosedAI

Использование TT для картинок

Original Image

Reconstructed Image

Original Image

Reconstructed Image

ClosedAI

https://github.com/stevenmaschan/nla-project.git