

Как найти себя в мире LLM на практике, а не в теории

Михаил Пожидаев

3 октября 2025 г.

Место под Солнцем

Как выйти на траекторию научных исследований?

Доступность обучающих данных

Без обучающих данных невозможно обучить LLM с нужными навыками. Для обучения LLM обучающие данные в некоторых областях могут быть недоступны.

Полнота навыков моделей

На HuggingFace представлено множество открытых LLM, но перечень их навыков ограничен, и нередко русский язык поддерживается недостаточно. Компании не заинтересованы в публикации материалов, способных привести к стратегической конкуренции.

Доступность оборудования

Можно не пояснять, все и так понимают, о чём речь.

Multi-head attention

Ключевое математическое выражение XXI века!

$$\alpha(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V$$

Выражение внутреннего внимания зарядило интеллектуальной магией языковые модели, но его возможности всё ещё до конца неисследованы.

- ▶ Можем ли применять Multi-head attention для обработки данных, отличных от текстовых?
- ▶ Можем ли применять Multi-head attention для предварительного кодирования текста и решения классических лингвистических задач?

Фундаментальная лингвистика

С применением Multi-head attention и Fine tuning

Морфологическое тегирование:

1. Добавляем линейную классифицирующую ячейку после МНА.
2. Используем датасет Sintagrus и перекрывающуюся токенизацию для открытого словаря.
3. Получаемая точность определения частей речи 98.2%.

Оценка расстояния до зависимых слов:

1. Используем модель RuGPT от Сбера.
2. Собираем предложения с числительными и самостоятельно сокращаем числительные.
3. Производим Fine tuning модели, передавая пары исходных и сокращённых предложений.

Новые устремления

low-Rank Adaptation

$$W = W_0 + \delta W$$

$$\delta W = AB$$

- ▶ W — обучаемая матрица весовых коэффициентов;
- ▶ W_0 — исходная «замороженная» матрица;
- ▶ δW — матрица для накопления данных точной настройки;
- ▶ A и B — матрицы пониженной размерности.

Миражи интеллекта

От теста Тьюринга до больших языковых моделей

Информация о книге: marigostra.ru/nlp-book



Спасибо за внимание!

Домашняя страница: <https://marigostra.ru>
Канал в Телеграм: <https://t.me/MarigostraRu>

