

# ALBERT for Question Answering on SQuAD 2.0

Тохчуков Данил

Факультет вычислительной математики и кибернетики  
МГУ имени М. В. Ломоносова

2023

Вопросно-ответные системы используются для того, чтобы помочь людям эффективно находить релевантную информацию. В статье решается задача Question Answering на датасете SQuAD 2.0. Среди существующих вопросно-ответных систем рассматриваются SOTA системы на основе предобученного трансформера ALBERT, и вопросно-ответная система Bidirectional Attention Flow (BIDAF). Также предлагается собственная вопросно-ответная система на базе ALBERT и Bidirectional Attention Flow (BIDAF). Система сохраняет логику BIDAF и использует идеи из SOTA систем на базе ALBERT.

**Article:** Endangered Species Act

**Paragraph:** “... Other legislation followed, including the Migratory Bird Conservation Act of 1929, a 1937 treaty prohibiting the hunting of right and gray whales, and the Bald Eagle Protection Act of 1940. These later laws had a low cost to society—the species were relatively rare—and little opposition was raised.”

**Question 1:** “Which laws faced significant opposition?”

**Plausible Answer:** later laws

**Question 2:** “What was the name of the 1937 treaty?”

**Plausible Answer:** Bald Eagle Protection Act

Датасет SQuAD 2.0 состоит из 130,319 объектов обучающей выборки и 11,873 объектов тестовой выборки. В обучающей выборке 33% объектов не имеют ответа на вопрос, в валидационной же 50% объектов без ответа.

# Постановка задачи

Задача состоит в том, чтобы найти правильный ответ на заданный вопрос по тексту, причём ответ на вопрос непрерывен и целиком находится в тексте — это значит что он не разделяется другими токенами из текста. Требуется найти начало и конец ответа – задача классификации.

Если в тексте не будет искомого ответа, модель должна вывести что ответа нет.

# Метрики

Для измерения качества моделей применяются две метрики: Exact Match (EM) score и F1 score.

**Exact Match** это двоичная мера (истина/ложь) того, точно ли выходные данные системы соответствуют основному истинному ответу (exact match accuracy). Это довольно строгий показатель.

**F1** менее строгая метрика. Она является средним гармоническим precision и recall модели.

$$F1 = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

**A Lite BERT (ALBERT)** – языковая модель с механизмом Self-Attention. На основе ALBERT строилась SOTA вопросно-ответная система ALBERT (ensemble):

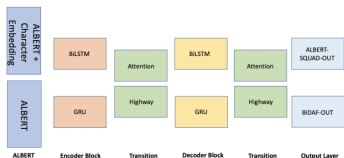
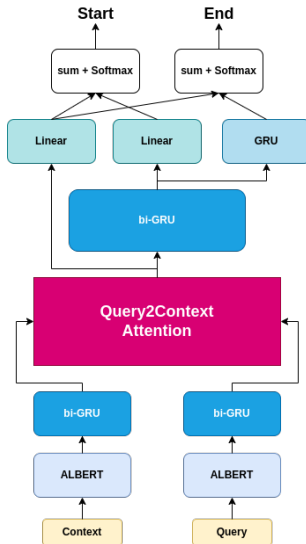


Figure 1. Schema of Model Architecture

**BIDAF** – полноценная вопросно-ответная система, успешно применявшаяся к задаче Question Answering на датасете SQuAD 1.0.

# Предложенная система



# Заключение

В работе была решена задача Question Answering на датасете SQuAD 2.0. Также в работе были представлены основные существующие системы для решения Question Answering задачи. Также была предложена архитектура "облегченной" вопросно-ответной системы на базе языковой модели ALBERT.