МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«РОССИЙСКИЙ БИОТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ (РОСБИОТЕХ)»

Институт промышленной инженерии, информационных технологий и мехатроники

Кафедра «Информатика и вычислительная техника пищевых производств»

Направление:

**ЛАБОРАТОРНАЯ РАБОТА № 22-24**

*на тему:*

«СОЗДАНИЕ ЧАТ-БОТА В ТЕЛЕГРАМЕ С ИСПОЛЬЗОВАНИЕМ RNN.

ЗНАКОМСТВО С ТРАНСФОРМЕРАМИ»

Вариант № 7

|  |  |  |
| --- | --- | --- |
| Выполнил: |  | Студент 4 курса, гр. 21-ИУ-3  Морозова М.П. |
| Проверил: |  | Ящун Т.В. |

**Цель работы:** разработать с использованием модели RNN на базе ячеек LSTM чат-бота, способного общаться с пользователями на естественном языке и выполнять различные задачи, используя знания из определённой предметной области; познакомиться с возможностями трансформеров для решения задач в области NLP.

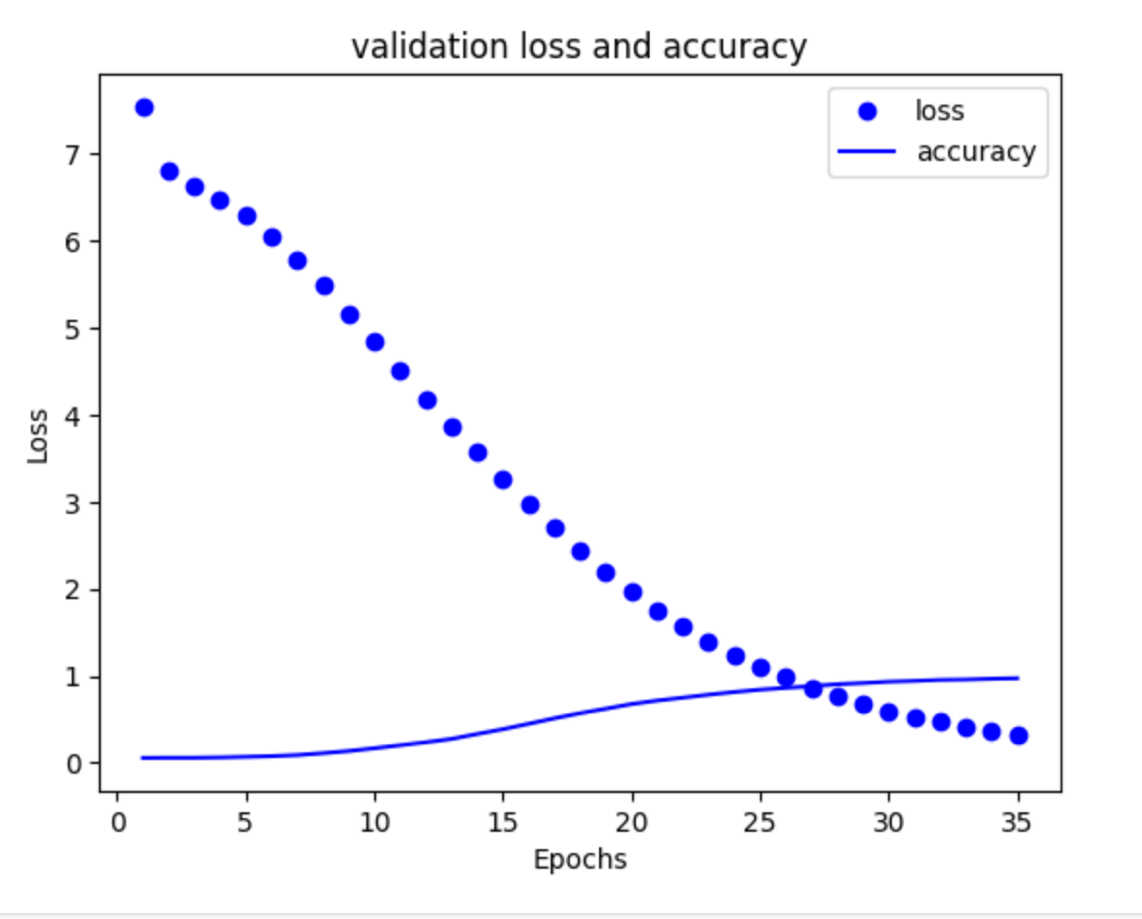
**ЗАДАНИЕ 1.** СОЗДАНИЕ ЧАТ-БОТА В ТЕЛЕГРАМЕ С ИСПОЛЬЗОВАНИЕМ РЕКУРРЕНТНЫХ LSTM

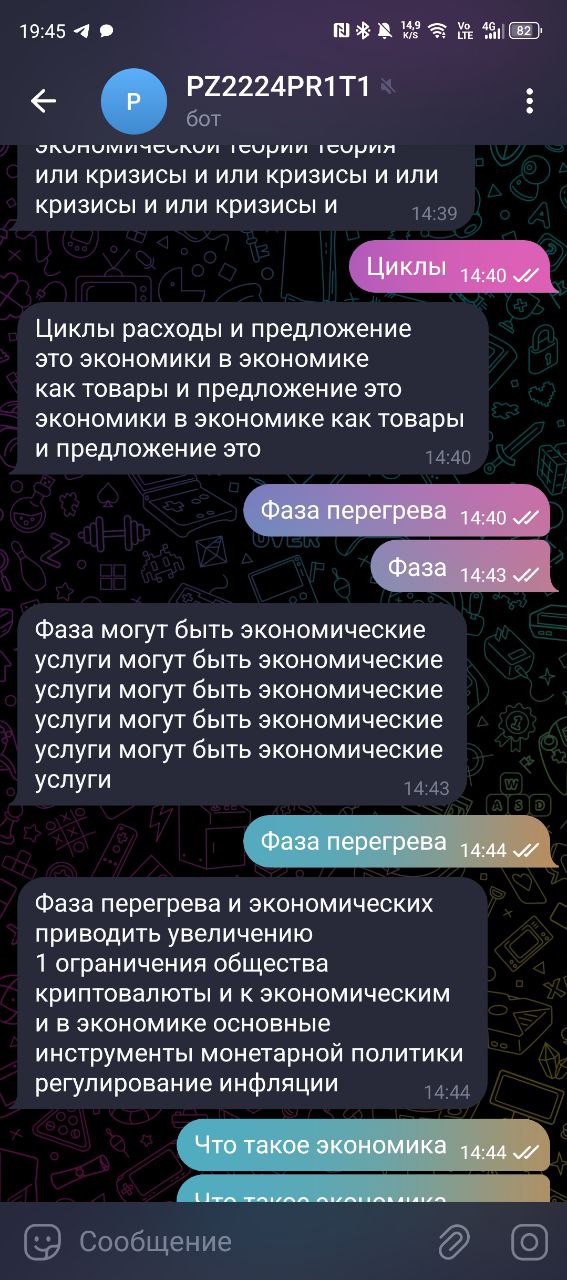
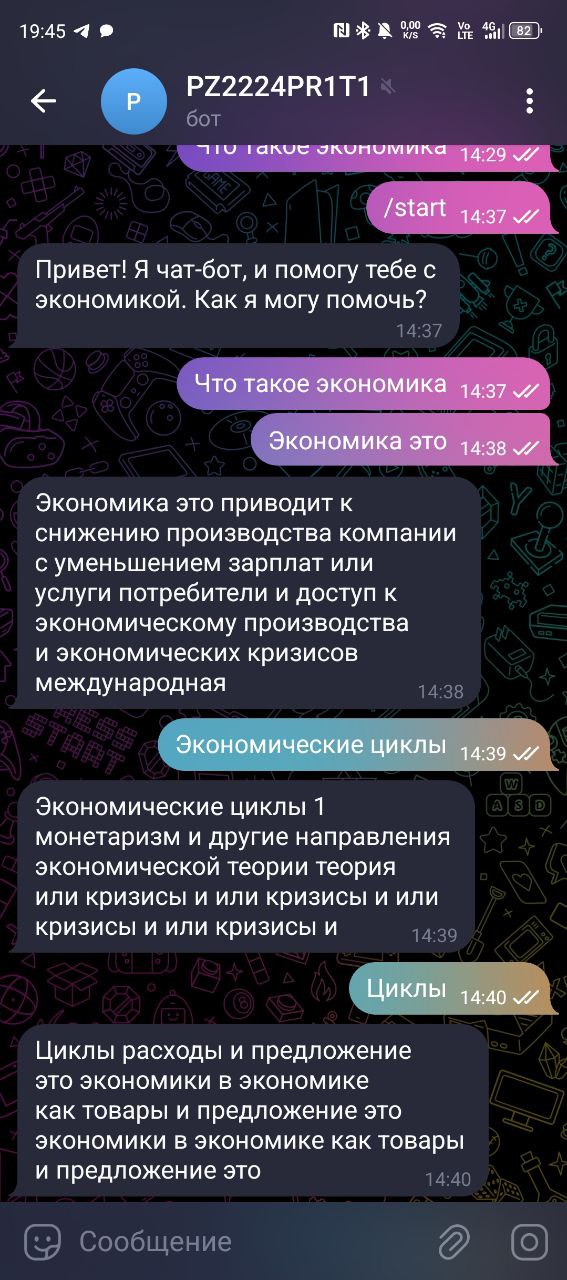
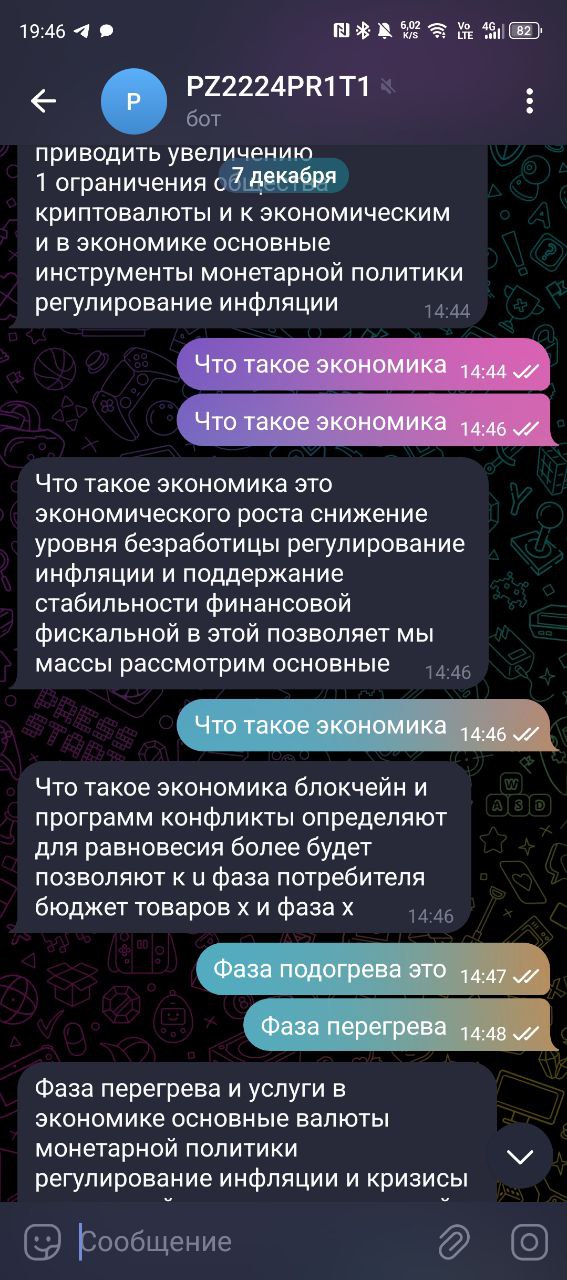
**Ссылка на GoogleCollab:** <https://colab.research.google.com/drive/1Zyl_CeQHkCWWJr0IpBro2TrWb8gUOxc8?usp=sharing>

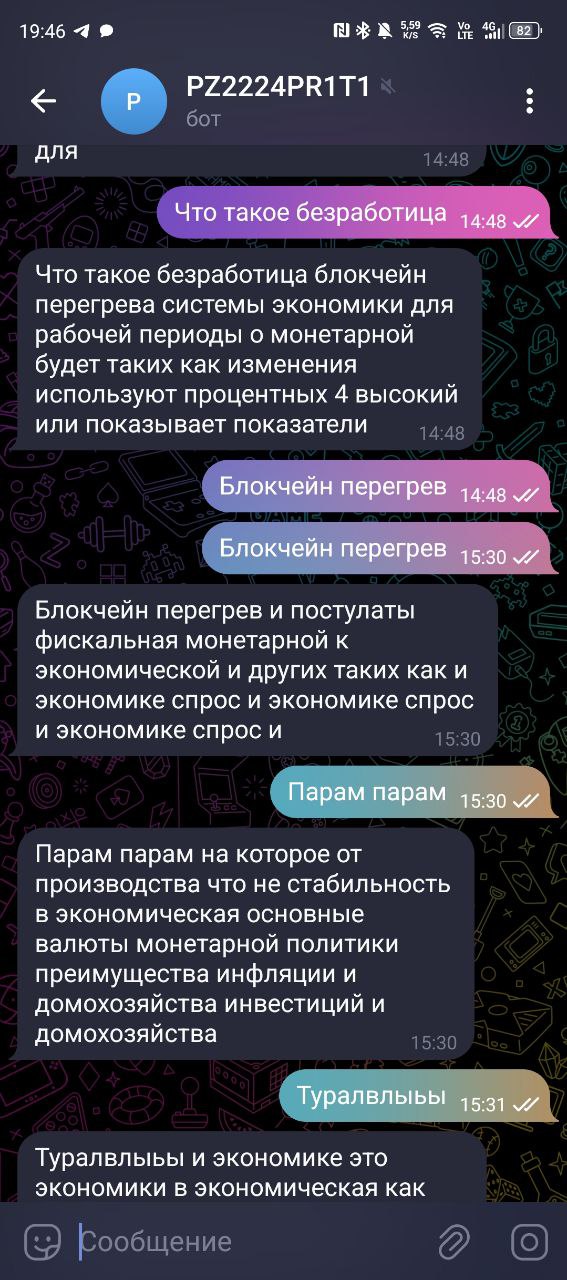
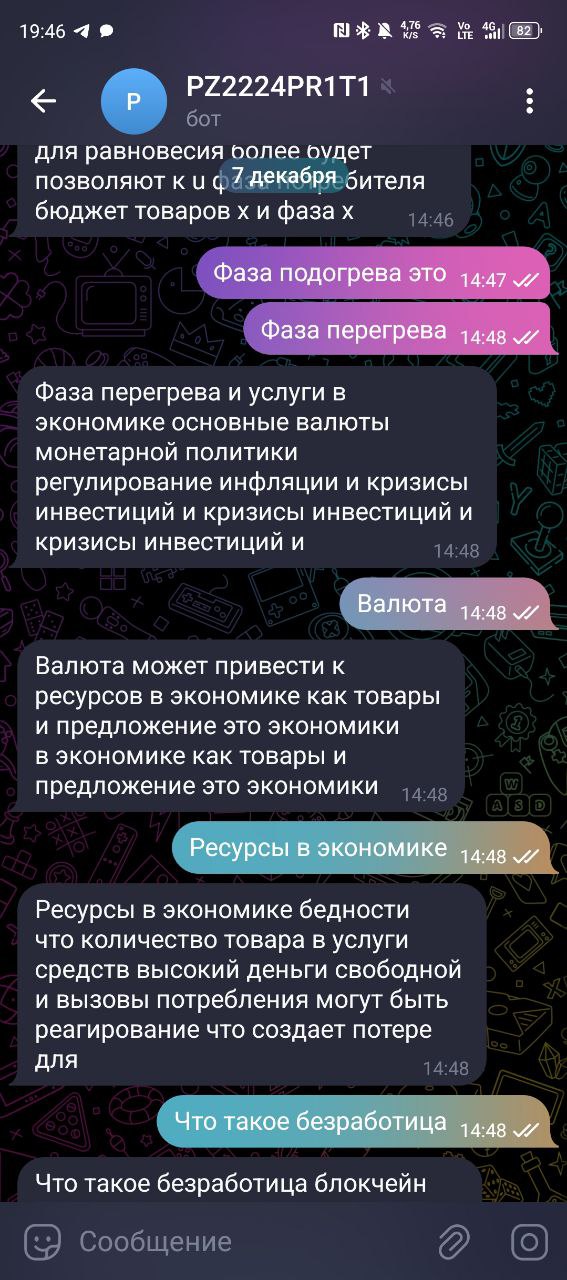
Сравнительная таблица показателей точности работы RNN

****

**Выводы:** ни число нейронов, ни количество слов для прогноза не оказали сильного влияния на показатели точности.



****

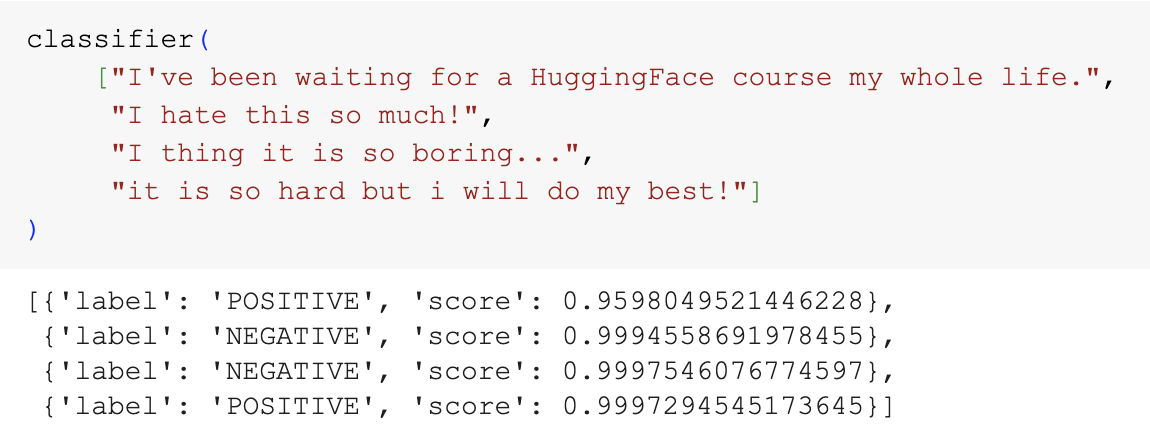
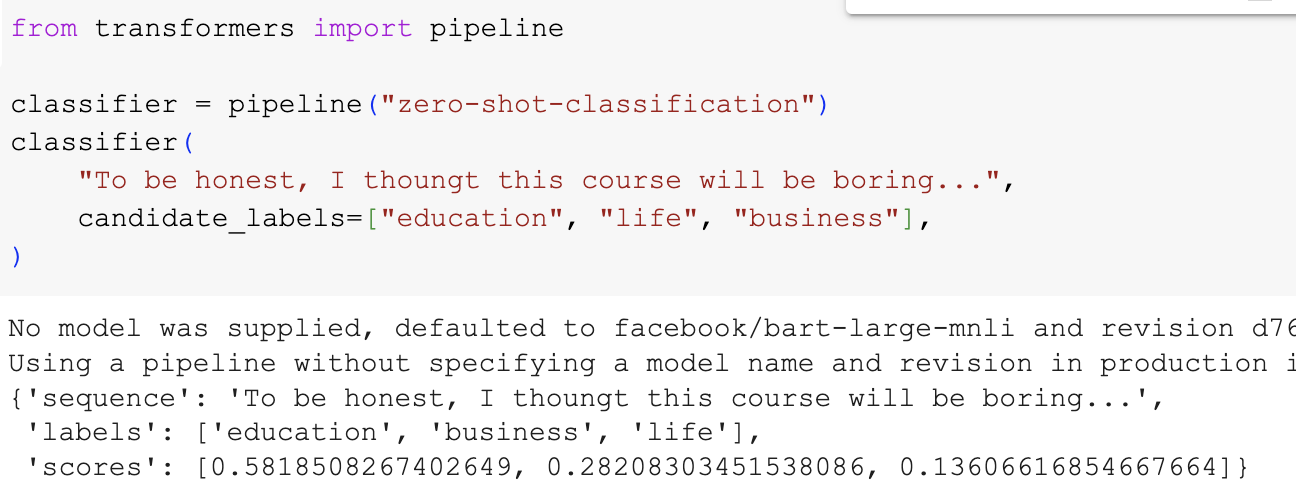
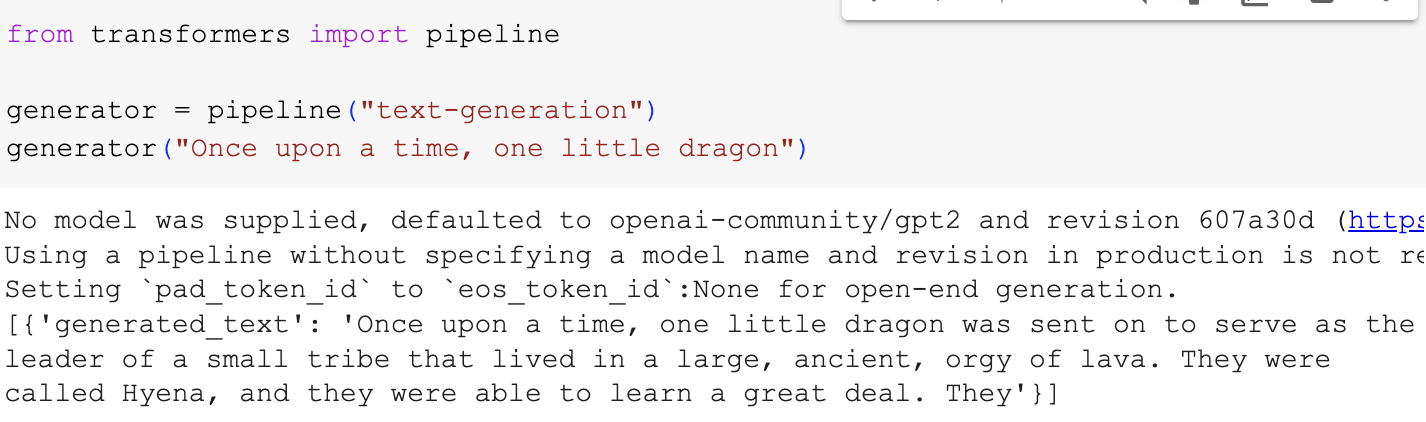
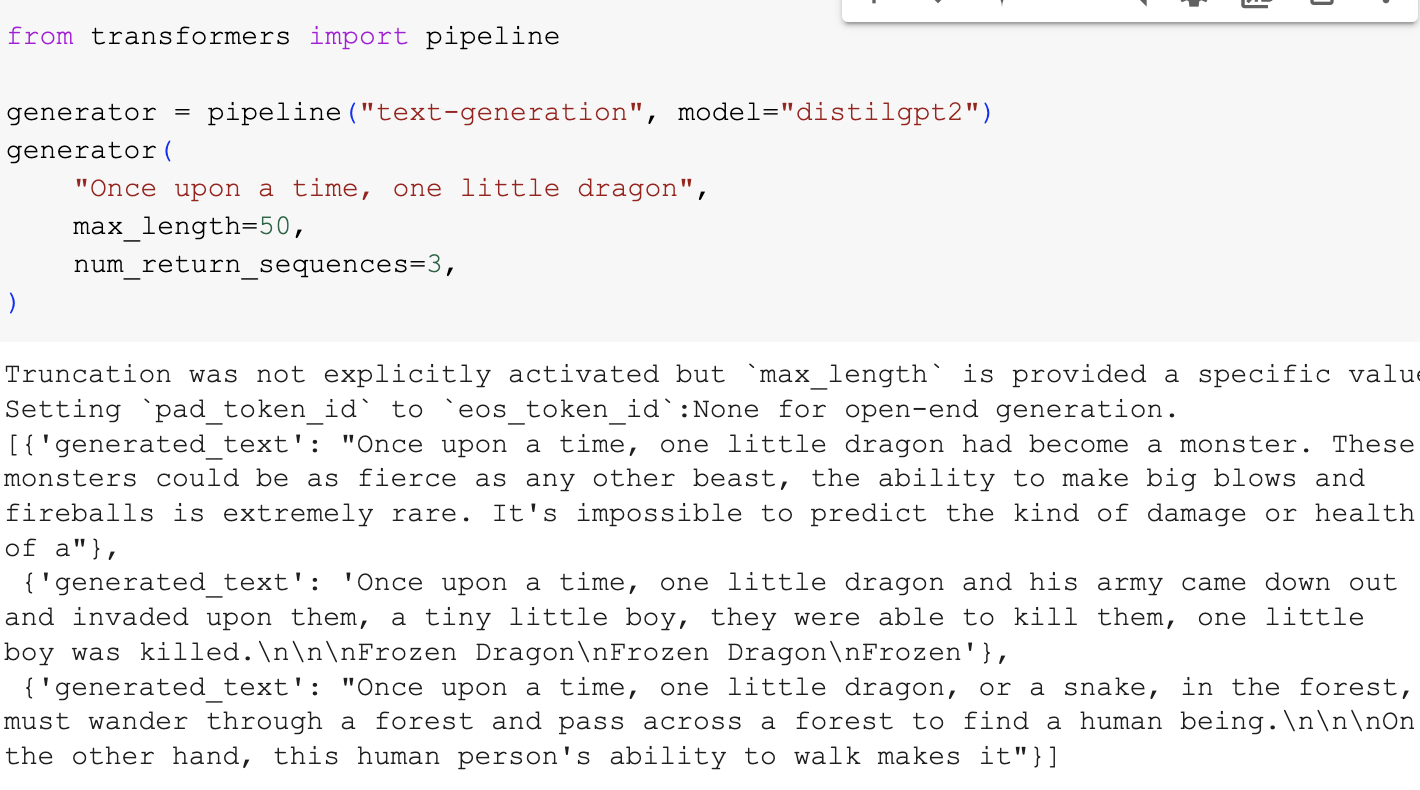
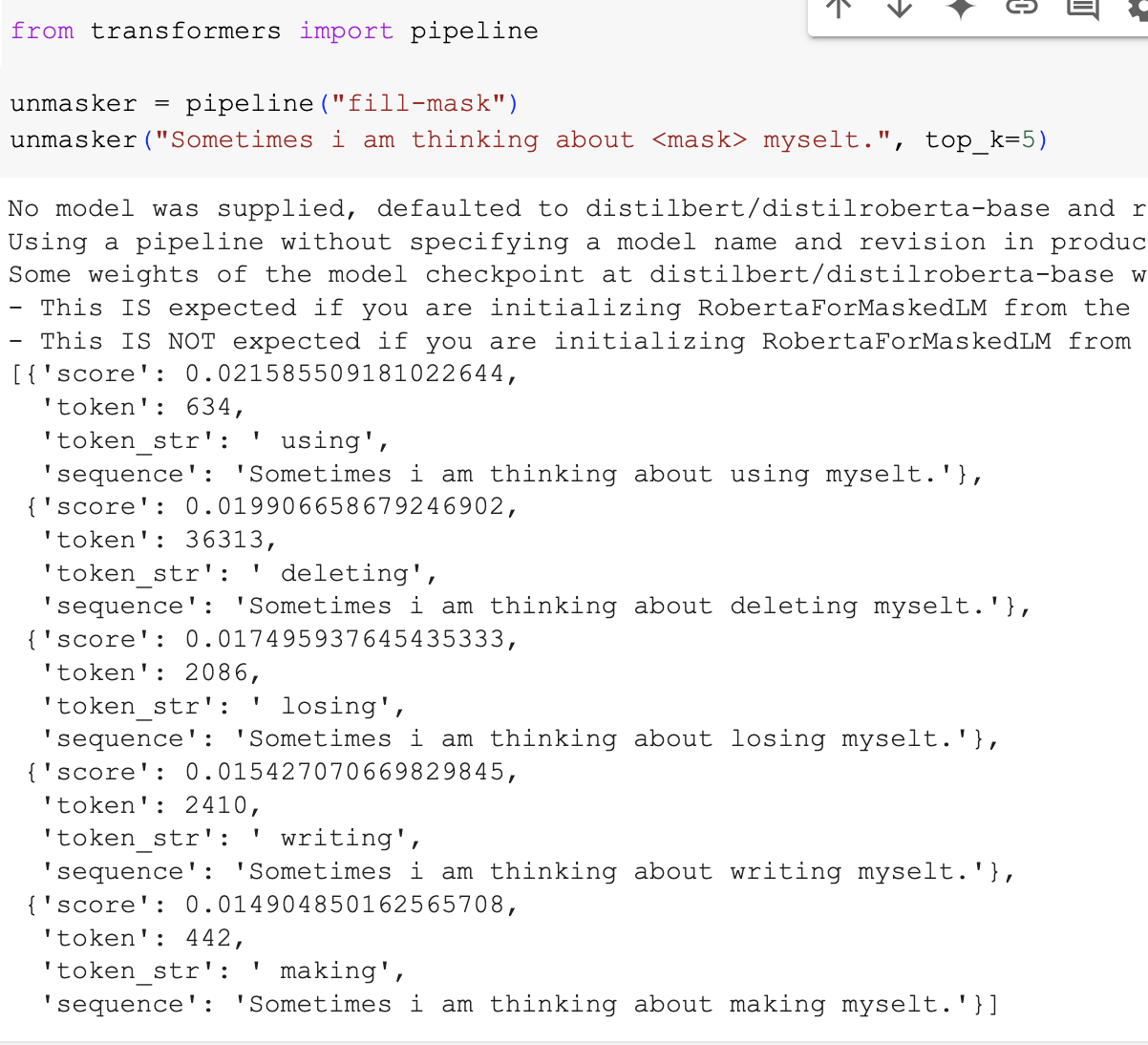
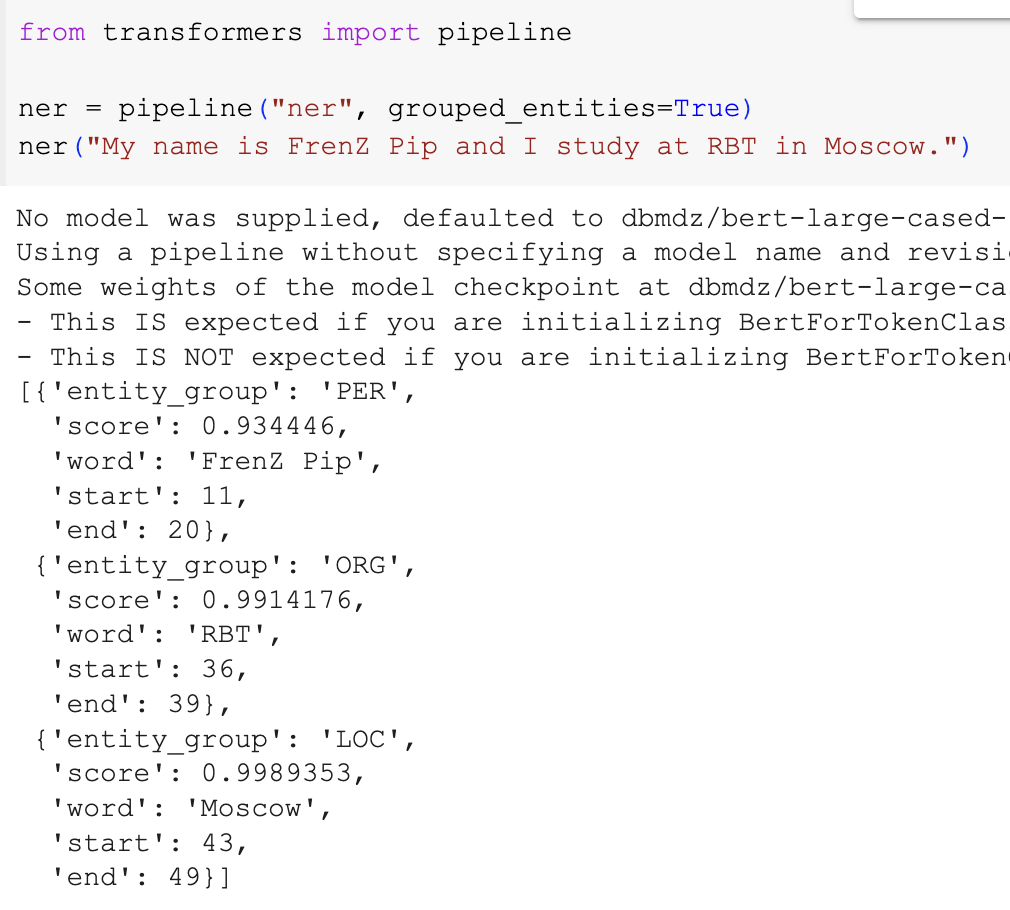
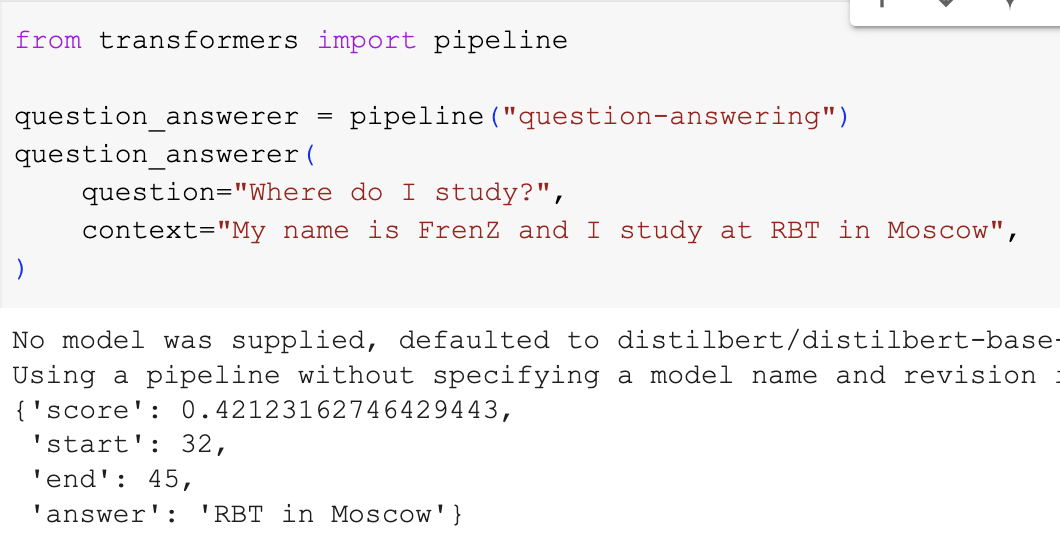
****

**Выводы:** обученная модель использует слова из предметной области, но составленные ею предложения не несут смысла.

**ЗАДАНИЕ 2.** ЗНАКОМСТВО С ТРАНСФОРМЕРАМИ

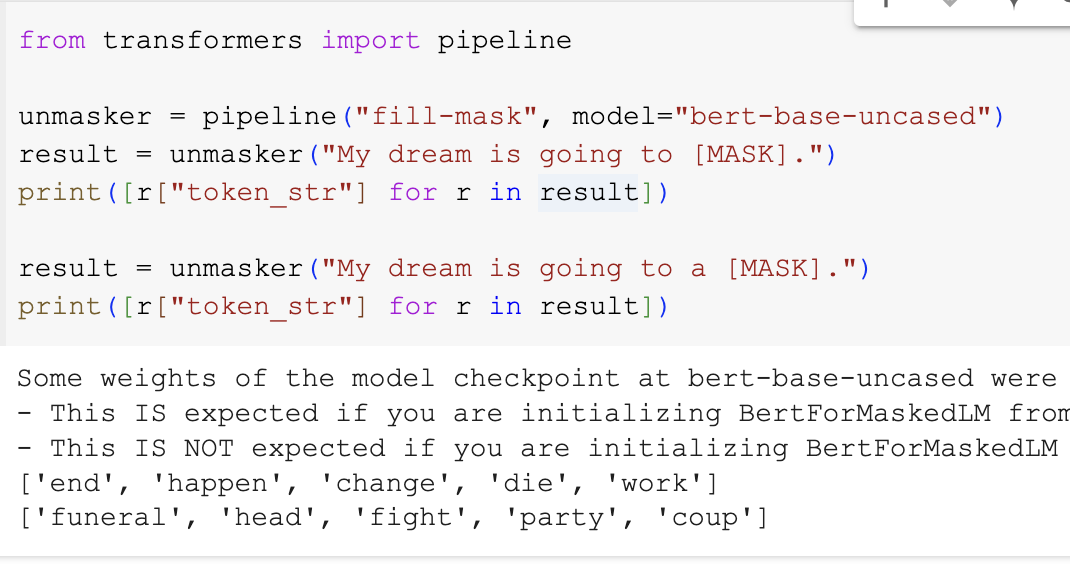
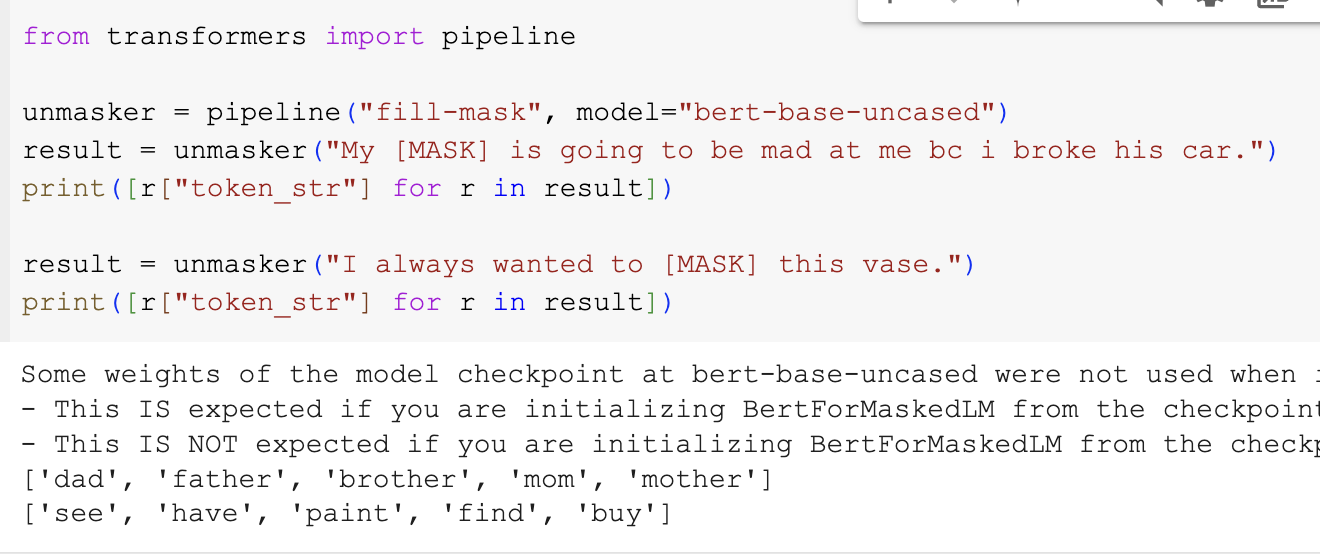
1. **Трансформеры: на что они способны?**

Ссылка на GoogleCollab: <https://colab.research.google.com/drive/1TTexxegOJ_mY5IkxTS0uV5p_w0G8P6Ra?usp=sharing>

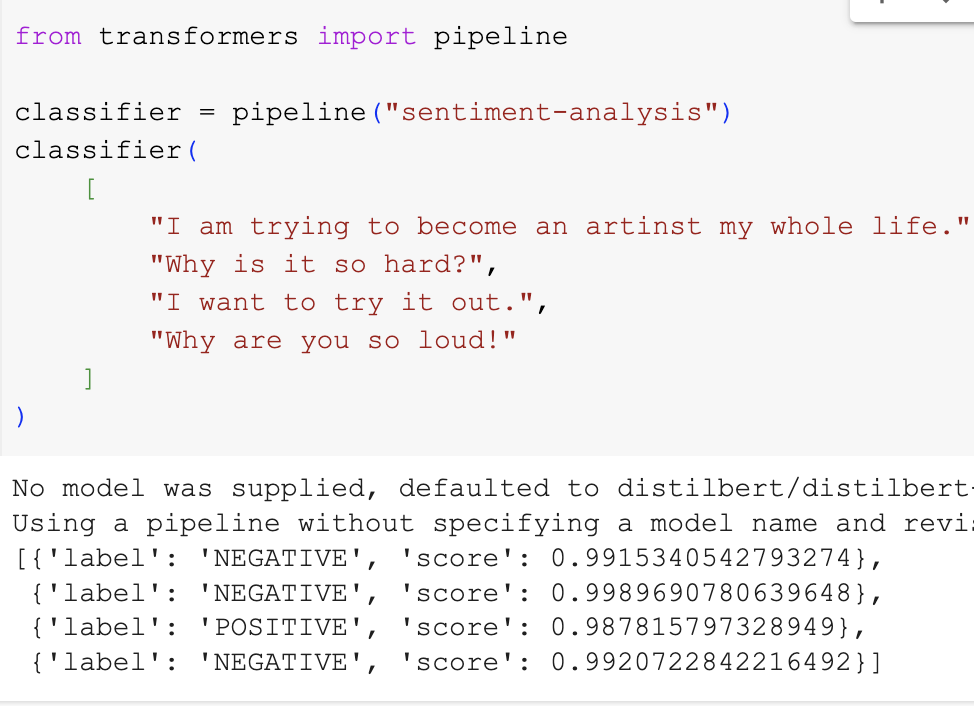
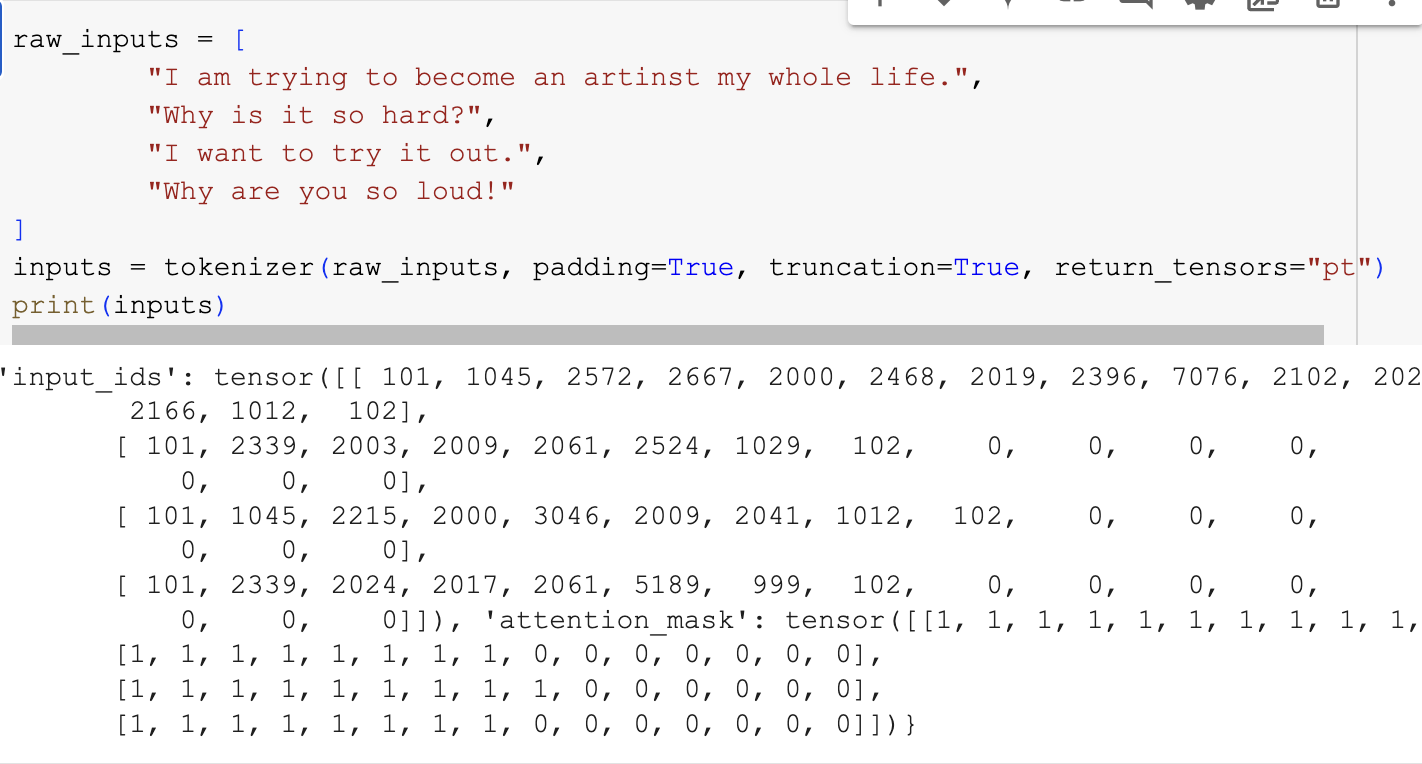


1. **Предвзятости и ограничения**

Ссылка на GoogleCollab: <https://colab.research.google.com/drive/1ePlYu_5P_2ySDgZl0xIhykUzeWUZy-Ia?usp=sharing>

****

1. **За конвейером**

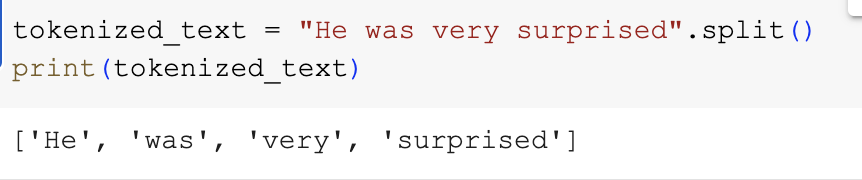
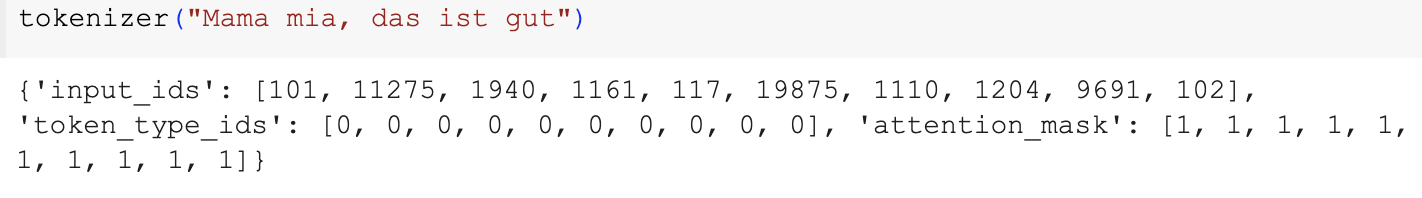
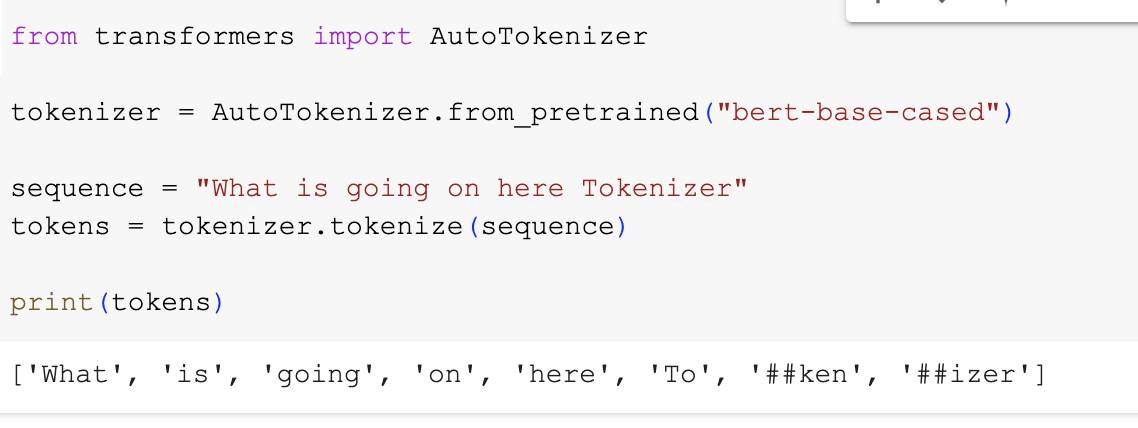
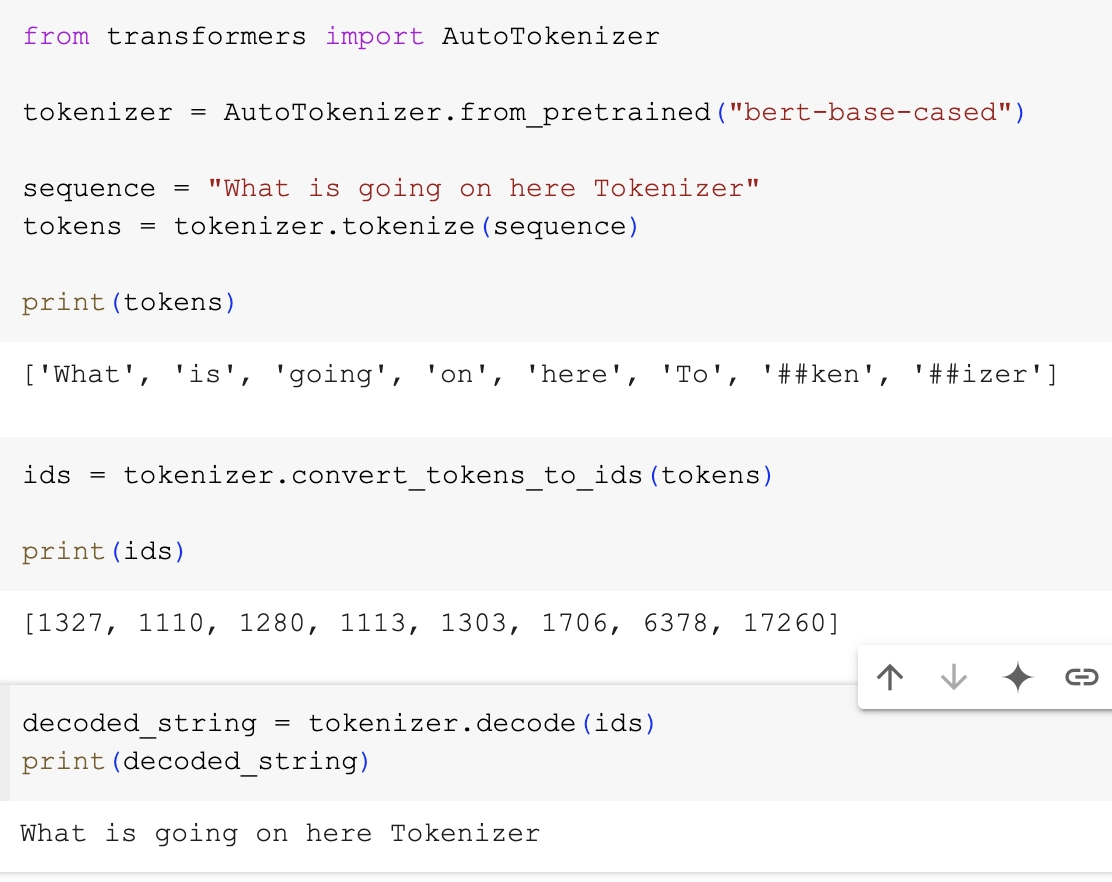
Ссылка на GoogleCollab: <https://colab.research.google.com/drive/15-_NYE3sq87pFsi32jeEYFBTcauGZAyn?usp=sharing> ****

1. **Модели**

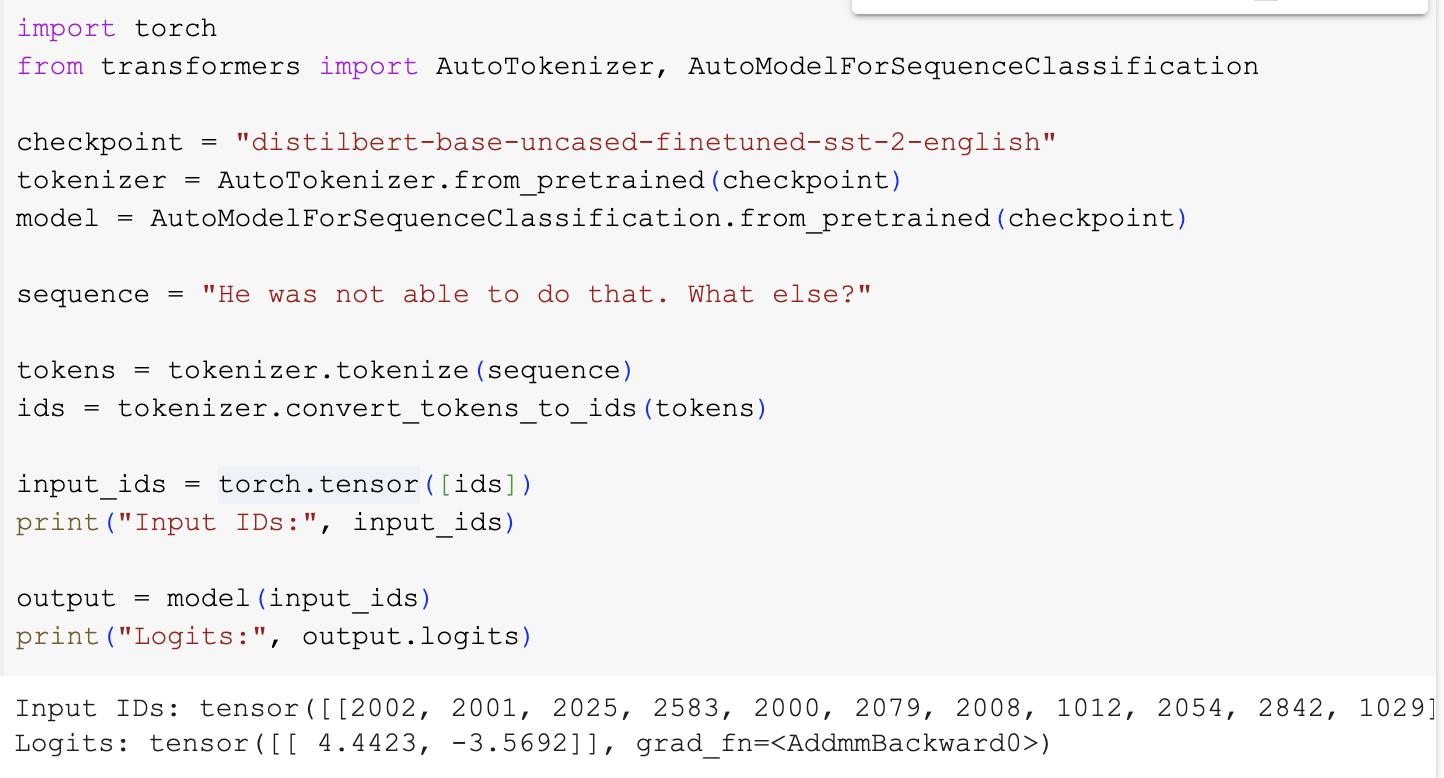
Ссылка на GoogleCollab: <https://colab.research.google.com/drive/13Tq3SoCiW4386K1_CqqBUrpNNKXssStL?usp=sharing>

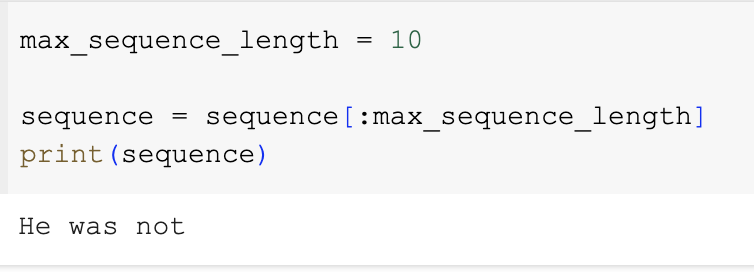
1. **Токенизаторы**

Ссылка на GoogleCollab: <https://colab.research.google.com/drive/1TBPVzF2CF1CmoJ8jaJdrUQuq9iYXLTaQ?usp=sharing>

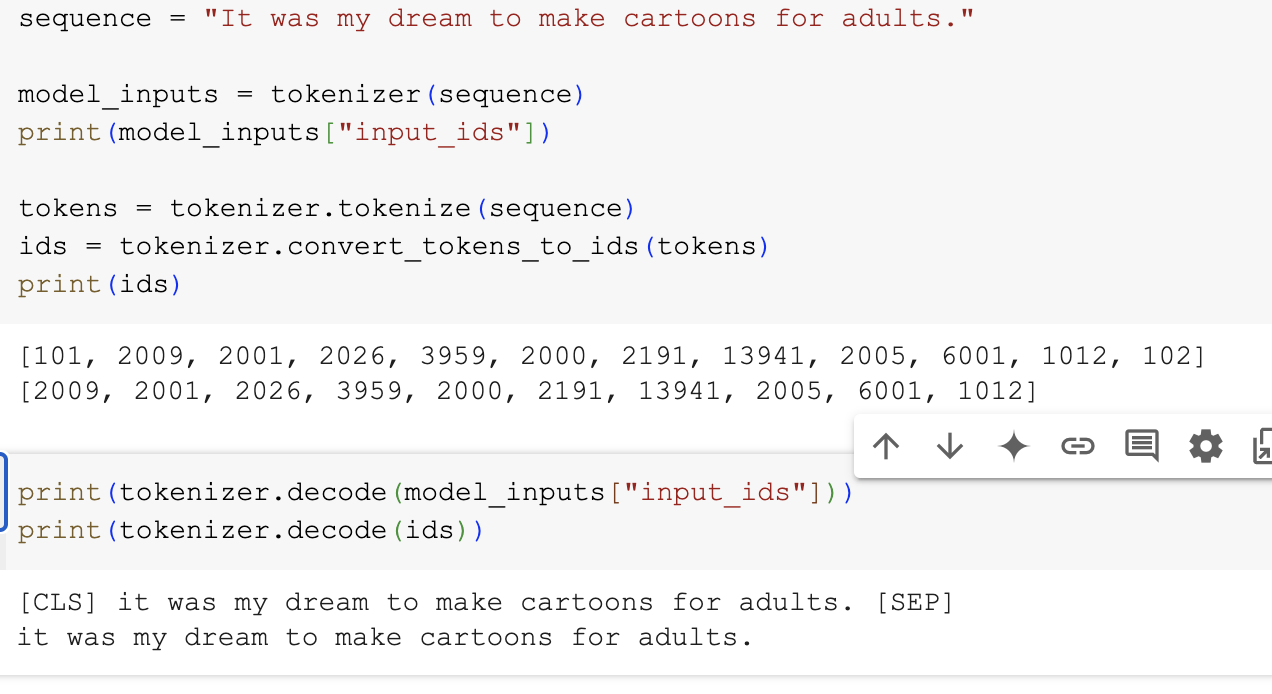
****

1. **Обработка нескольких последовательностей**

Ссылка на GoogleCollab: <https://colab.research.google.com/drive/1Q6XYufHVqQGe9fMHwPzrU-YsFIdbXQ1B?usp=sharing> ****

****

1. **Собираем все воедино**

Ссылка на GoogleCollab: <https://colab.research.google.com/drive/1YqwGW7inZRRQNIYLSAEzGwHXU4__DHYZ?usp=sharing> ****

1. **Предобработка данных**

Ссылка на GoogleCollab: <https://colab.research.google.com/drive/1cMVbDvvaGqaJH-0FwaKD0OK5AEBwohdV?usp=sharing>

1. **Fine-tuning модели с использованием Trainer API**

Ссылка на GoogleCollab: <https://colab.research.google.com/drive/1XBlyTzvOSMJyg71P265h64c0Cn-kx8tO?usp=sharing>

1. **Полное обучение**

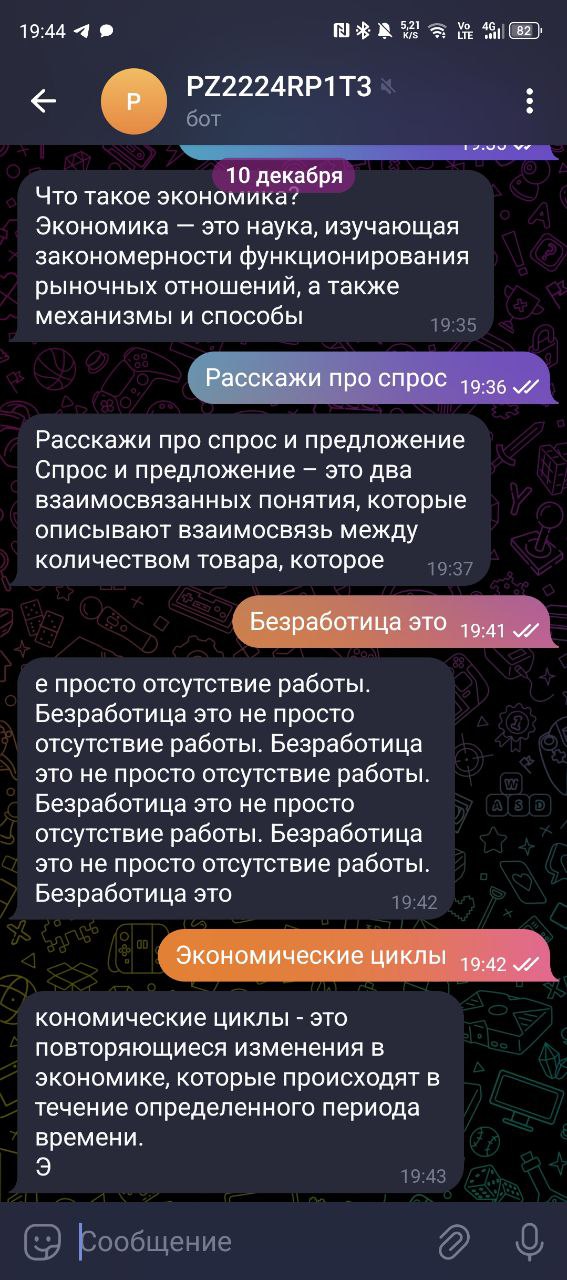
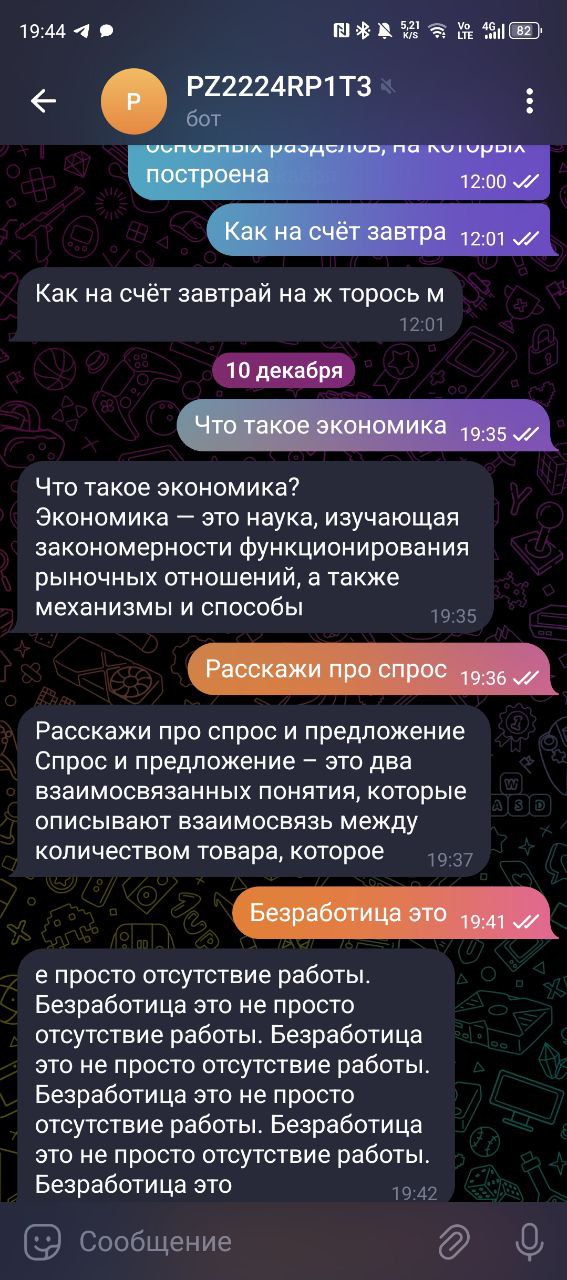
Ссылка на GoogleCollab: <https://colab.research.google.com/drive/1XJCB_LvfeSD33cvwqnQaPrRu7s0vdXAV?usp=sharing>

**ЗАДАНИЕ 3.** СОЗДАНИЕ ЧАТ-БОТА В ТЕЛЕГРАМЕ С ИСПОЛЬЗОВАНИЕМ ТРАНСФОРМЕРА

Ссылка на GoogleCollab:

<https://colab.research.google.com/drive/1ReWYpEDOBxAE_zvEUUERlQOetxBYv9YG?usp=sharing>

До обучения:



После обучения:

