|  |  |
| --- | --- |
| lu135925on3bu_tmp_3360867a00ce4d37 | **Министерство науки и высшего образования Российской Федерации Федеральное государственное бюджетное образовательное учреждение высшего образования** **«Московский государственный технический университет имени Н.Э. Баумана** **(национальный исследовательский университет)» (МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ Информатика и системы управления и искусственный интеллект

КАФЕДРА                  Системы обработки информации и управления

**Лабораторная работа №8** **по курсу «Методы машинного обучения в автоматизированных системах обработки информации и управления»**

Подготовили:

**Чжан Чжиси**

**ИУ5И-25М**

03.06.2024

Проверил:

**Гапанюк Ю. Е.**

*2024 г*.

**Задание:**

Для произвольного предложения или текста решите следующие задачи:

1. Токенизация.
2. Частеречная разметка.
3. Лемматизация.
4. Выделение (распознавание) именованных сущностей.
5. Разбор предложения.

Сначала я разложила текст на лексемы с помощью библиотеки natasha и разбила его на разделы. Затем я лемматизировала каждую лексему и извлекла именованные сущности, такие как имена людей и организаций. Наконец, я разобрала предложение на составляющие и вывела его текст и части речи.

Текст для анализа：Солнце светит ярко, птицы поют, а ветер шумит в листве

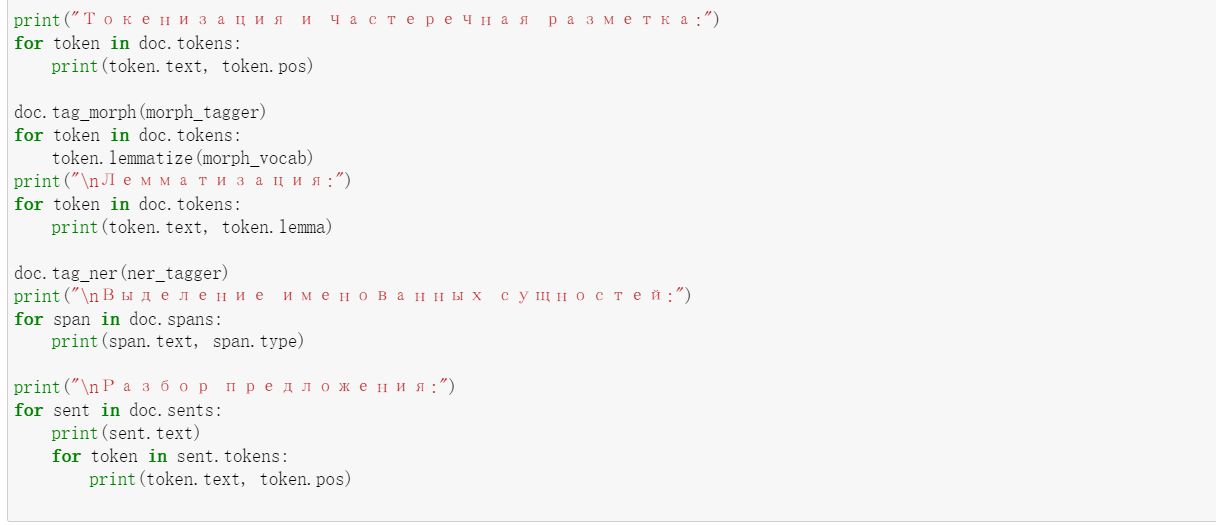


Рис.1-Все коды работают

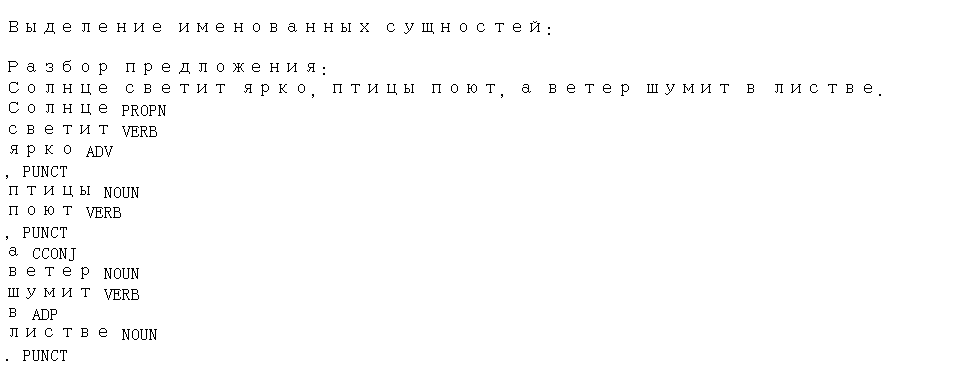
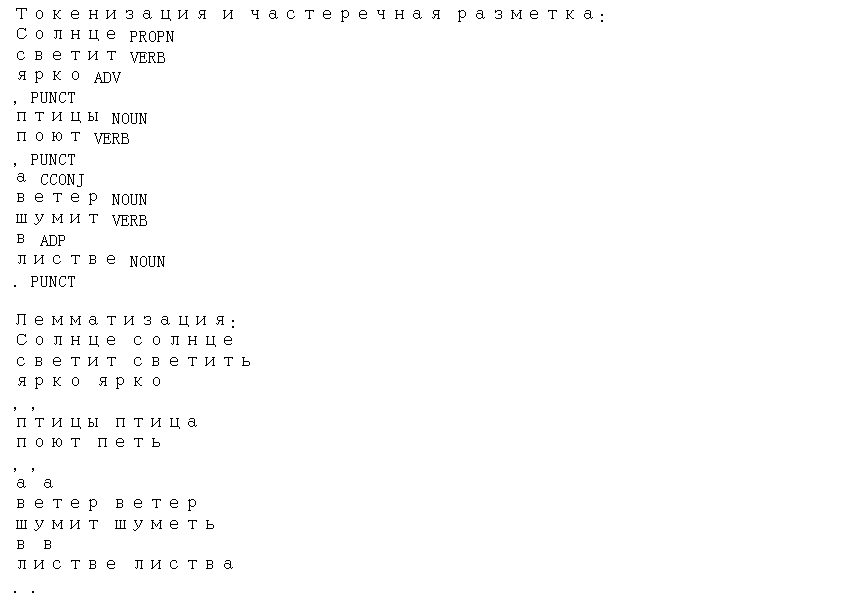


Рис.2-Результаты предварительной обработки текста