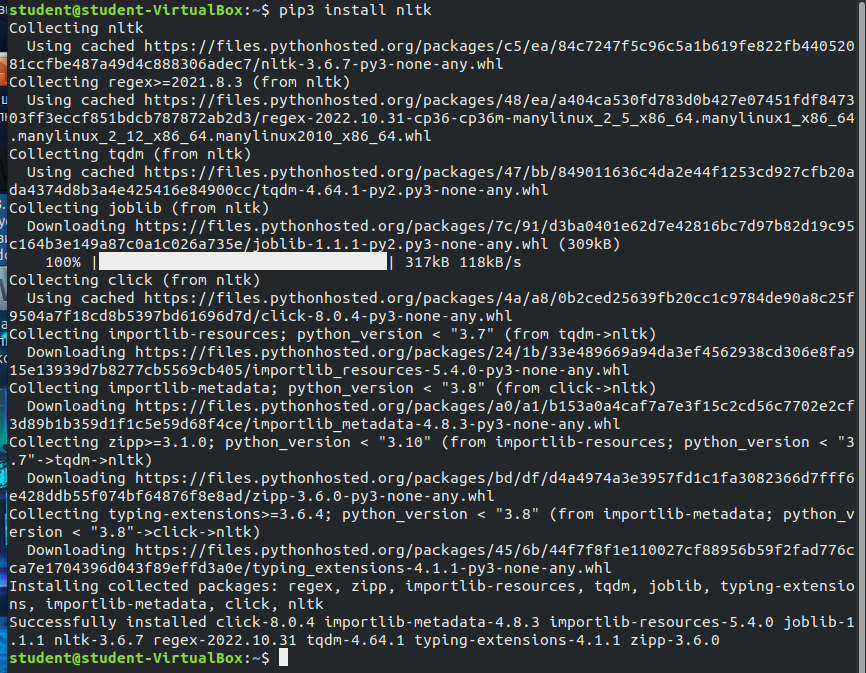
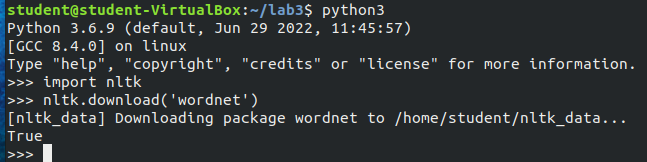
|  |  |  |
| --- | --- | --- |
| Лабораторная работа № 3  «Корпусная лингвистика. Машинное обучение»  Компьютерная лингвистика | Ф.И.О. | Самороков Н. Н. |
| Группа | ИВТ-363 |
| Преподаватель | Руженников А. А. |
| Дата сдачи |  |

**Ход работы:**

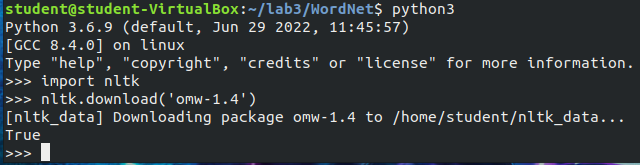
Установка пакета nltk (*Natural Language Toolkit*)



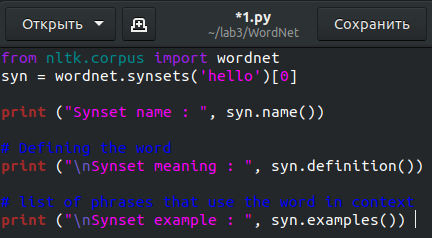
Загрузка WordNet через nltk (*WordNet -* *лексическая база данных английского языка. Представляет собой словарь-тезаурус*)



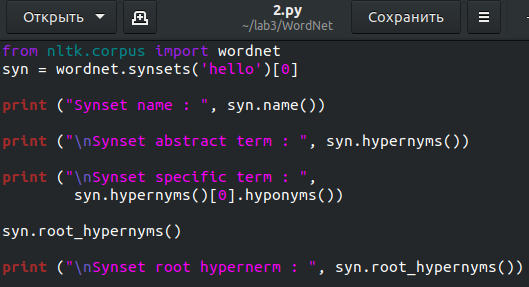
Загрузка omw-1.4, для избавления от дальнейших ошибок



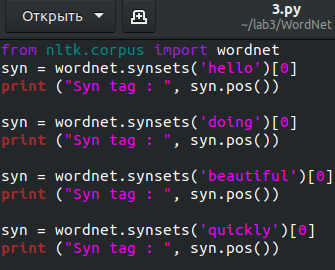
Редактируем файл 1.py



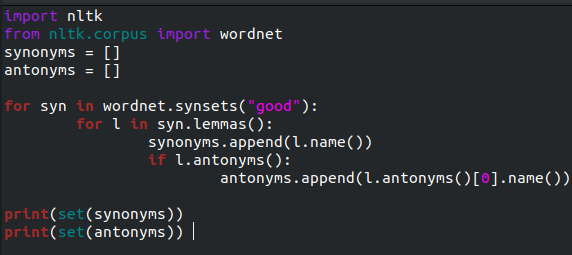
Редактируем файл 2.py



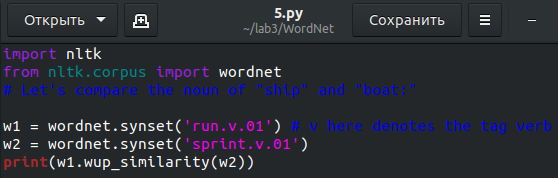
Редактируем файл 3.py



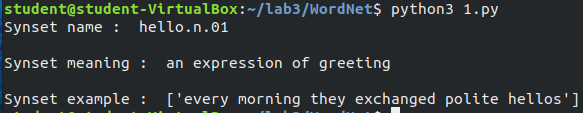
Редактируем файл 4.py



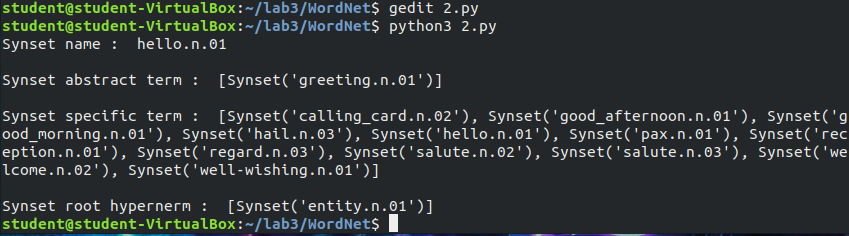
Редактируем файл 5.py



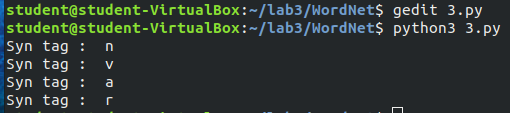
Запустим файлик 1.py и просмотрим результат



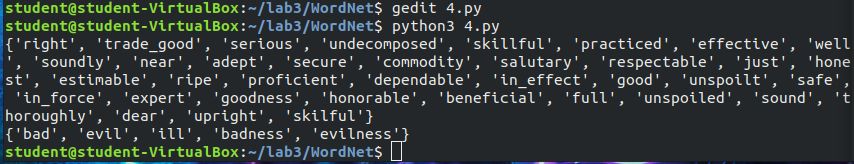
Запустим файлик 2.py и просмотрим результат



Запустим файлик 3.py и просмотрим результат



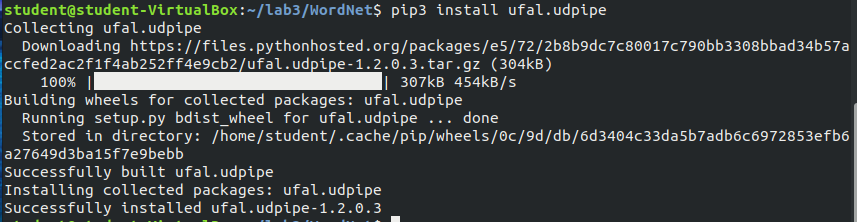
Запустим файлик 4.py и просмотрим результат



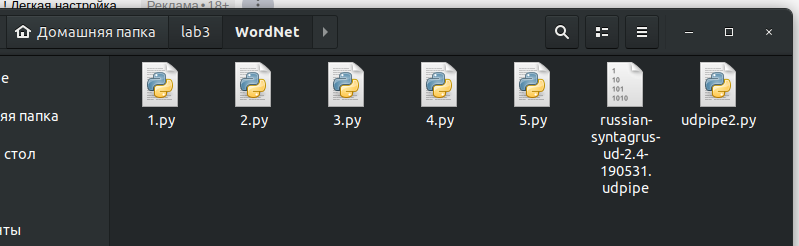
Запустим файлик 5.py и просмотрим результат (в данном примере он рассчитывает коэф. «схожести», т.е. на сколько 2 слова схожи с друг-другом)



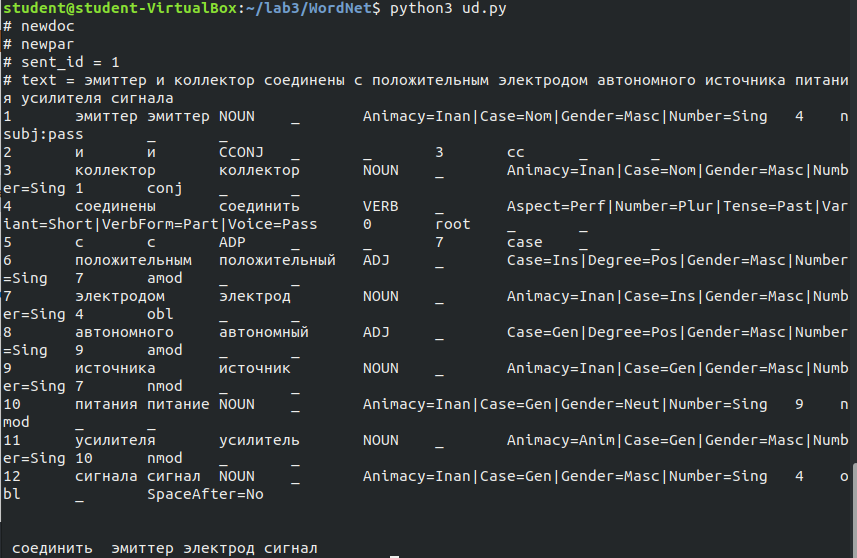
Предварительно установим ufal.udpipe



Скачаем модель «russian-syntagrus-ud-2.4-190531.udpipe» (<https://yadi.sk/d/b19-KhrAjMHgmg>). (Переместим в директорию к остальным файлам)

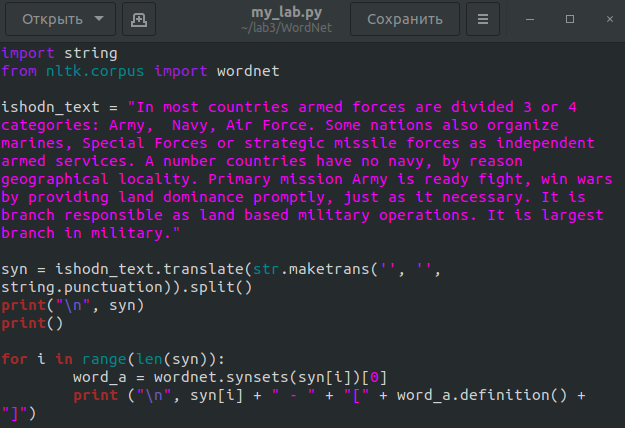


Скачаем пример с eos2.vstu и запустим ud.py, предварительно все перекинув в рабочую папку



Код для выполнения задания:

* + - 1. Для каждого слова в тексте вывести в скобках его WordNet-определение.



Переменная ishodn\_text – хранит анализируемый текст

syn – хранит массив слов, выделенных в тексте и заключенных в кавычки и преобразованных в строковой тип данный

word\_a – хранит «синсет» анализируемого слова, при проходе исходного текста по словам, соответственно

Результат работы кода

