В этом проекте создается чат-бот TermNormForm (@term\_norm\_bot), который принимает на вход текст и возвращает термины, которые можно выделить в этом тексте, в нормальной форме. Проект содержит:

* скрипт «код1\_бот.py», в котором написан основной код
* файл «for\_bot.env», где записана переменная окружения (токен для бота)
* файл «requirements.txt», где перечислены зависимости и их версии:

certifi==2024.12.14

charset-normalizer==3.4.1

click==8.1.8

colorama==0.4.6

DAWG-Python==0.7.2

docopt==0.6.2

idna==3.10

joblib==1.4.2

nltk==3.9.1

numpy==2.0.2

pandas==2.2.3

pymorphy2==0.9.1

pymorphy2-dicts-ru==2.4.417127.4579844

pyTelegramBotAPI==4.25.0

python-dateutil==2.9.0.post0

python-dotenv==1.0.1

pytz==2024.2

regex==2024.11.6

requests==2.32.3

six==1.17.0

tqdm==4.67.1

tzdata==2024.2

urllib3==2.3.0

Бот был создан с помощью библиотеки «pyTelegramBotAPI». Токен для бота (в виде строки) был сгенерирован через BotFather. Токен хранится не в самом коде, а в файле с переменными окружения для безопасности и удобства (чтобы токен было легко вызывать в любой части программы (чтобы не нужно было его долго искать) и чтобы он не был виден посторонним). Переменные окружения загружаются в основной код с помощью модуля «python-dotenv».

Для токенизации текстовых сообщений от пользователей используется библиотека «nltk» (RegexpTokenizer). Токенизация производится с помощью регулярных выражений. Для морфологического анализа текстовых сообщений используется библиотека «pymorphy2».

**Как запускается бот:**

1. Запустить код бота (файл «код1\_бот.py») в PyCharm
2. Отправить команду «/start» в чат (@term\_norm\_bot)
3. В ответ на эту команду придет сообщение «Привет! Напиши предложение, я найду термины».
4. Написать текст, отправить его в чат.
5. В ответных сообщениях появятся термины, которые удалось найти в этом тексте.