

ПОДГОТОВКА ОКРУЖЕНИЯ ДЛЯ ВЫПОЛНЕНИЯ ЛАБОРАТОРНЫХ РАБОТ ПО ДИСЦИПЛИНЕ «МАШИННОЕ ОБУЧЕНИЕ И БОЛЬШИЕ ДАННЫЕ»

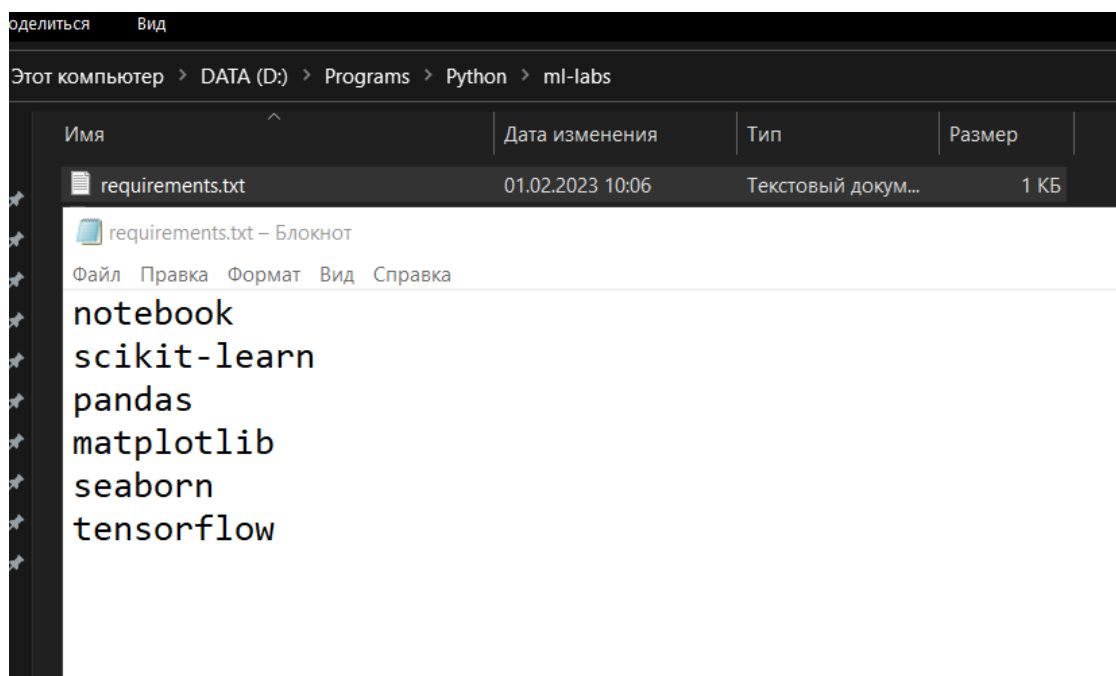
Все лабораторные работы выполняются локально, с использованием среды Jupyter Notebook. Для выполнения лабораторных работ рекомендуется использовать актуальную версию Python (например, Python 3.9). Выполните установку Python на ваш компьютер. Для проверки установки вызовите **cmd** из любой папки на вашем компьютере и выполните команду `python --version`. Если нет ошибок и на экране Вы видите версию установленного Python – все хорошо и можно следовать инструкции. Если нет – необходимо исправить ошибки установки.

Рассматривается настройка окружения для пользователей операционной системы (ОС) семейства Windows.

1. Создайте в удобном месте на компьютере папку, в которой будет располагаться окружение для работы (будут храниться ноутбуки, наборы данных, сохраненные модели, требуемые зависимости и др.). Рекомендуется, чтобы в названии папки и пути до нее не было кириллических символов.

2. Создайте в папке виртуальное окружение. Для этого:

а. Создайте файл *requirements.txt* (в нем перечисляются все зависимости, которые необходимо установить в окружение для работы кода); рекомендуется сразу же добавить несколько библиотек, с которыми мы будем работать (если при выполнении работ потребуются специфические библиотеки – об этом будет написано).



б. Далее рекомендуется (не обязательно, для вашего удобства) создать структуру папок; в папке notebooks предлагается хранить ноутбуки (точно такие же, как в Google Colab); в папке data – датасеты; в папке models – сохраненные модели.

Имя	Дата изменения	Тип	Размер
data	01.02.2023 10:09	Папка с файлами	
models	01.02.2023 10:09	Папка с файлами	
notebooks	01.02.2023 10:09	Папка с файлами	
requirements.txt	01.02.2023 10:06	Текстовый докум...	1 КБ

с. Создайте виртуальное окружение. Для этого откройте из рабочей папки **cmd** и выполните команду `python -m venv venv`. Активируйте окружение командой `venv\Scripts\activate`. **Важно:** перед началом выполнения работ, а также при необходимости установки зависимостей не забывайте каждый раз активировать окружение! Обратите внимание, что при активированном виртуальном окружении в командной строке перед командой вы видите префикс.

```
Администратор: C:\WINDOWS\System32\cmd.exe
Microsoft Windows [Version 10.0.19044.2486]
(c) Корпорация Майкрософт (Microsoft Corporation). Все права защищены.

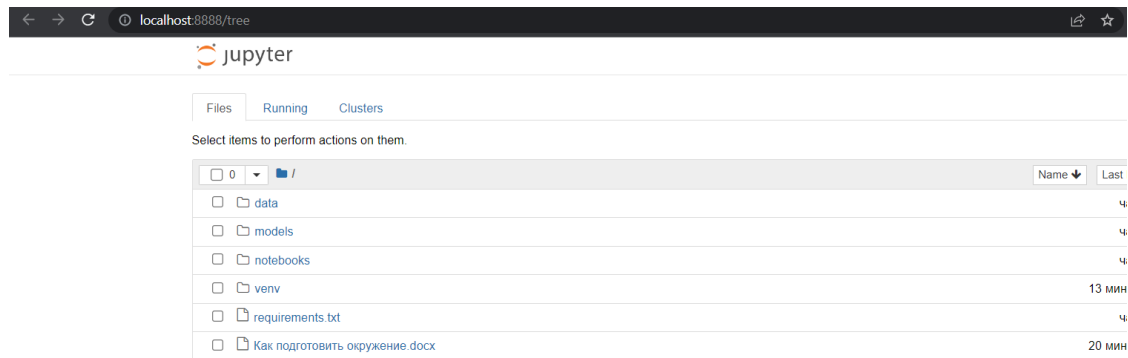
D:\Programs\Python\ml-labs>python -m venv venv

D:\Programs\Python\ml-labs>venv\Scripts\activate

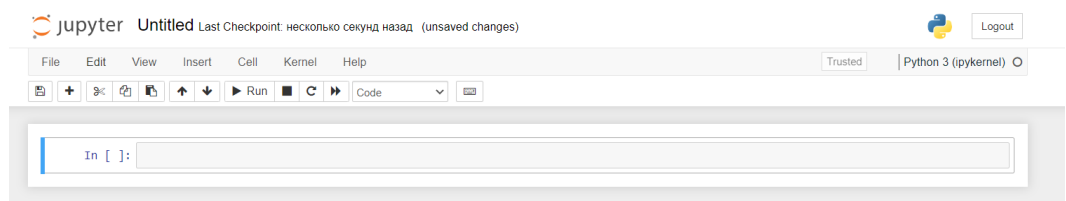
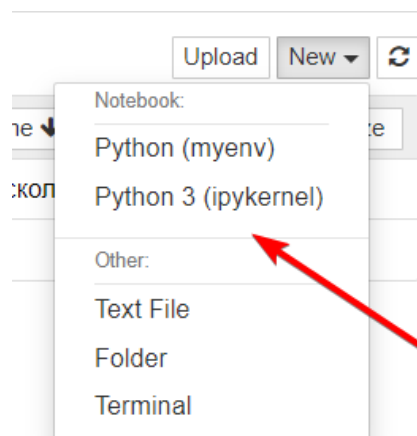
(venv) D:\Programs\Python\ml-labs>
```

д. Находясь в окружении, выполните команду **`pip install -r requirements.txt`** (эта команда запустит установку зависимостей которые описаны в файле; можно устанавливать их по отдельности, но обычно так удобнее).

е. После окончания установки зависимостей выполните команду `jupyter-notebook`. В результате автоматически должна открыться новая вкладка браузера.



ф. Самостоятельно изучите интерфейс приложения Jupyter Notebook. Во многом, он очень похож на знакомый вам Google Colab. Только все вычисления производятся на вашем компьютере. Важно: при создании нового ноутбука (если вам надо будет), выбирайте этот вариант



В этом видео рассказывается о том, как работать в приложении Jupyter Notebook:

https://www.youtube.com/watch?v=bfCfBuIhlB8&ab_channel=%D0%90%D0%BD%D0%B4%D1%80%D0%B5%D0%B9DataScience%D0%9A%D0%B0%D0%BD%D0%B0%D1%88%D0%BE%D0%B2

g. Важно: предыдущая команда запустила на вашем компьютере (порт 8888) сервер с приложением Jupyter Notebook. Если вызакроете окно cmd, в котором запустили сервер – он остановится. Выполнение всех ноутбуков прекратиться.

h. Перезапустить сервер можно всегда. Для этого не забудьте снова активировать виртуальное окружение и выполнить команду **jupyter-notebook**.

3. Лабораторные работы выдаются в виде тетрадок, или ноутбуков в формате *.ipynb. Они открываются, редактируются и сохраняются в приложении Jupyter Notebook. Вы можете выполнять задания (писать код) прямо в этих файлах. Можете создавать свои. Это ваш рабочий репозиторий, в котором можно формировать свое ML-портфолио. Только не забудьте в этом случае создать **файл .gitignore** и добавить туда папку **venv**.

ДОПОЛНИТЕЛЬНЫЕ ССЫЛКИ

<https://practicum.yandex.ru/blog/что-такое-jupyter-notebook/>

<https://habr.com/ru/companies/vk/articles/662734/>