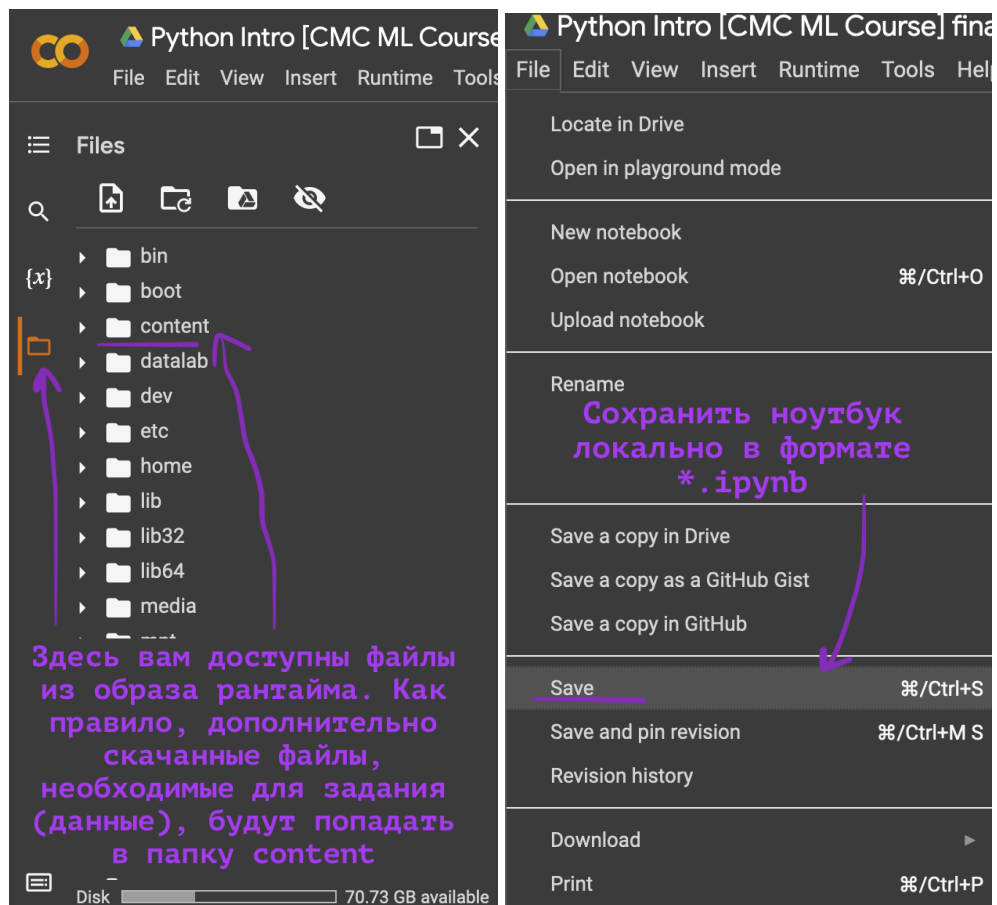
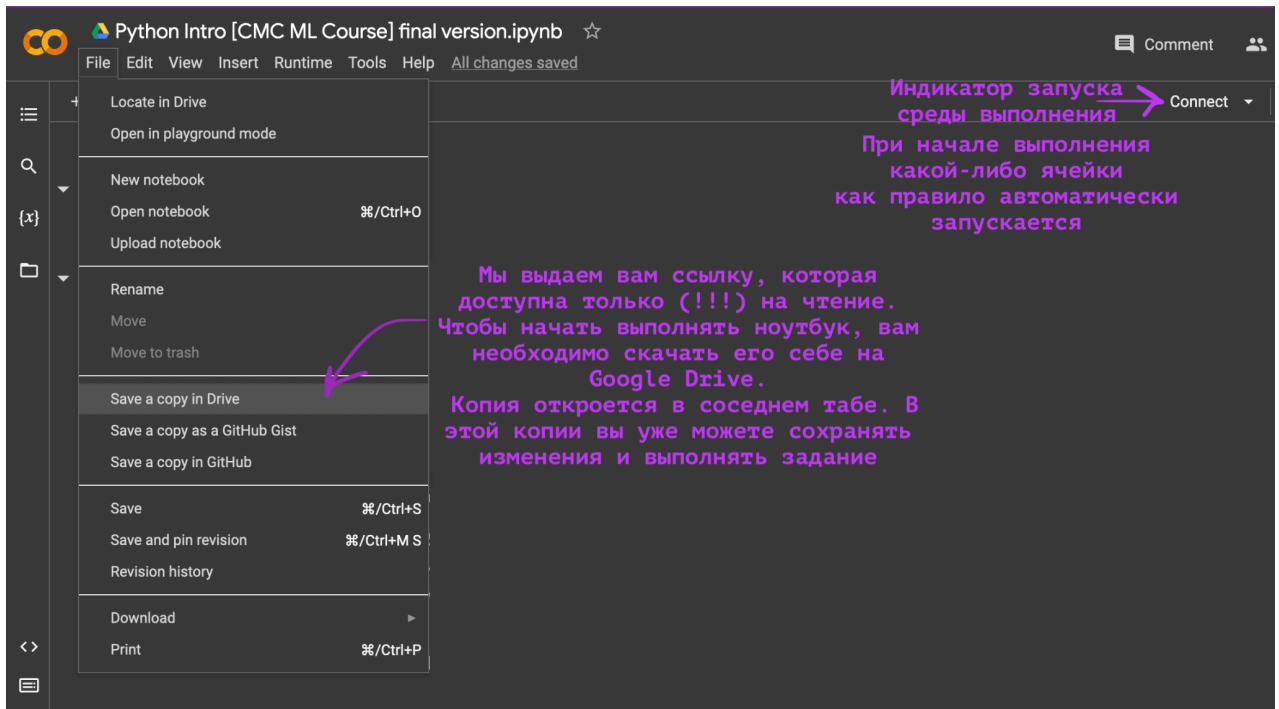


Краткий tutorial по использованию Google Colab и подготовке рабочей среды

В нашем курсе машинного обучения предполагается выполнение заданий с использованием языка Python3 и связанных с ним инструментов. Чтобы облегчить вам выполнение практических заданий, а также предоставить вам вычислительные ресурсы (если у вас слабый ноутбук, например), мы предлагаем вам выполнять некоторые части заданий в среде Google Colab.

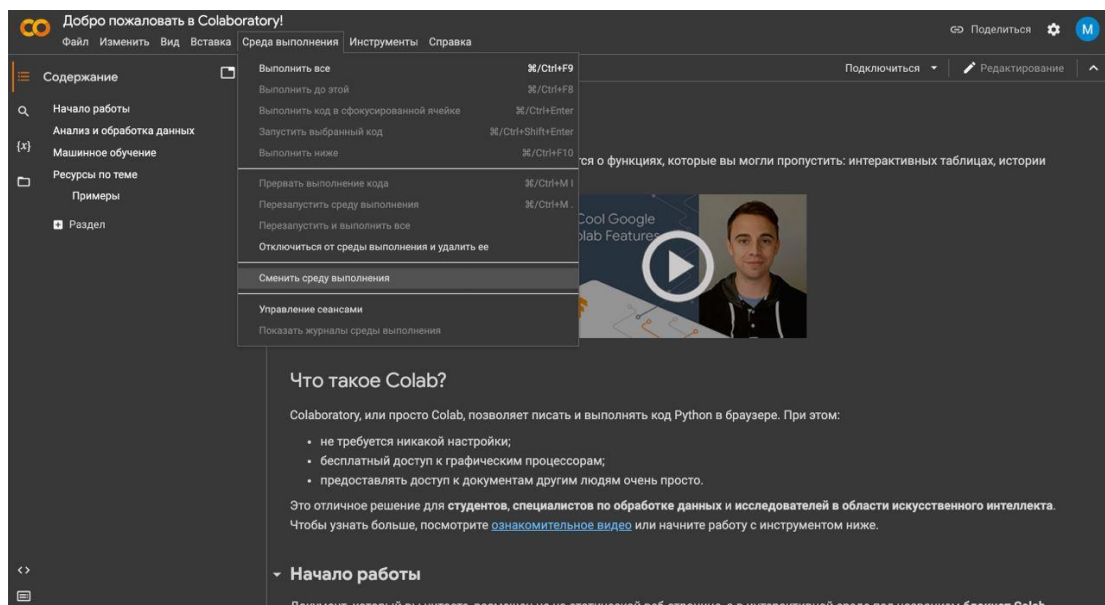
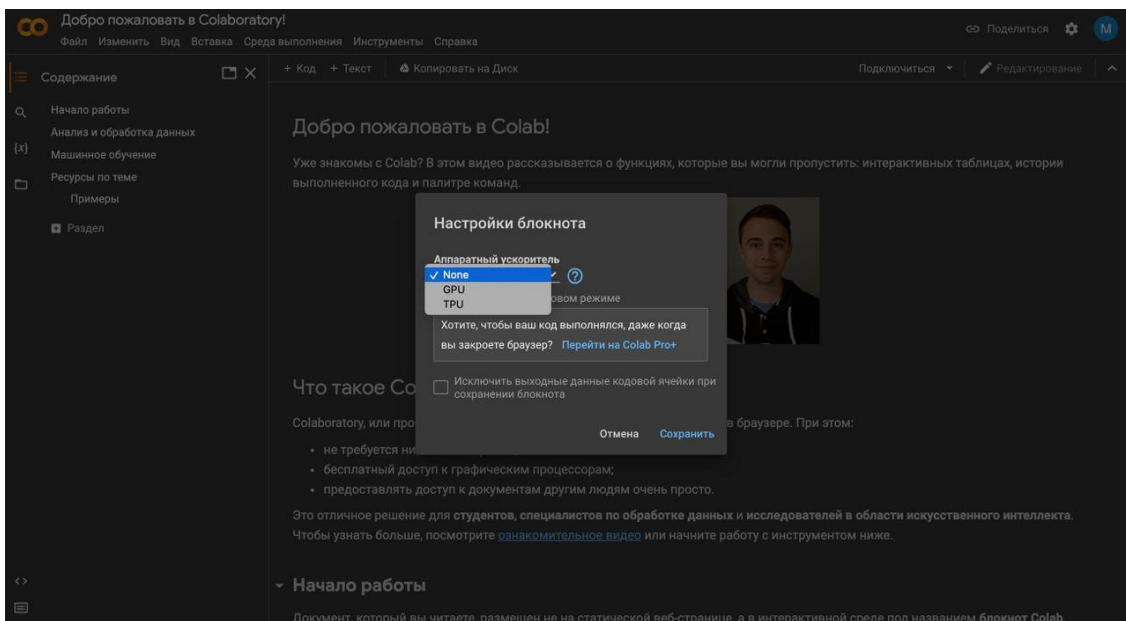
- Сайт проекта - <https://colab.research.google.com/>
- Вам необходим доступ к Google Drive. Скорее всего, ваша почта *gse.cs.msu.ru* уже имеет этот доступ. Если доступа по каким-то причинам нет, то можно зарегистрироваться в обычной почте *gmail.com*. Google Colab умеет работать с Google Drive как с файловой системой или хранилищем, что будет использоваться при выполнении вами практических заданий
- Google Colab представляет из себя интерактивную среду выполнения кода на Python наподобие среды Jupyter Notebook/Jupyter Lab. Если вы знакомы с последними, то перейти на Google Colab вам не составит труда.
- В Google Colab можно писать в ячейках код, выполнять его, смотреть на результат, писать текст с разным форматированием, вставлять картинки и прочее
- Хороший tutorial по ознакомлению с Google Colab можно найти вот тут <https://colab.research.google.com/notebooks/intro.ipynb?hl=ru-ru>, <https://blog.skillfactory.ru/chto-takoe-google-colaboratory-i-komu-on-nuzhen/>. Также не воспрещается искать дополнительную информацию в интернете :)

Здесь мы приведем функционал, который понадобится вам при выполнении заданий



Работа с GPU! (Актуально для курса по DL)

- Для обучения нейронных сетей и некоторых других алгоритмов (например, разных реализаций градиентных бустингов: xgboost, catboost, lightgbm) предпочтительнее использовать графические процессоры gpu. Их можно использовать, сменив runtime_type так, как показано ниже.
- Обратите внимание, что включить gpu в runtime_type недостаточно; нужно каким-то образом указать в коде (по-разному для каждого метода), что вы хотите использовать именно gpu. (например, для catboost – это параметр task_type = 'cuda', а для torch – torch.device("cuda").



Важное про использование runtime_type:

- При смене runtime_type ноутбук полностью перезапустится!

Важно!

- Google Colab имеет ограничения на используемую в сессии оперативную память и CPU, а также имеет ограничения на продолжительность выполнения одной сессии
- Хотя официально Google Colab не озвучивает ограничения <https://research.google.com/colaboratory/faq.html#usage-limits>, обычно часов через 12 активная запущенная сессия ноутбука прерывается <https://stackoverflow.com/questions/54057011/google-colab-session-timeout>
- Для нас это означает следующее: если вы хотите оставить ноутбук считаться (например на час) и пойти заниматься своими делами, то не закрывайте браузер и не переводите компьютер в спящий режим. Тогда ваши вычисления, скорее всего, не должны прерваться. Если все же вычисления прервались, можно попробовать сделать один из хаків, описанных тут <https://stackoverflow.com/questions/54057011/google-colab-session-timeout> или поискать еще варианты решений в интернете.
- После сохранения сделанной работы на локальный компьютер, откройте его на своем локальном Jupyter Notebook/Jupyter Lab/Visual Studio Code или пр. умеющим отображать *.ipynb приложением. Проверьте, что все требуемые выводы, картинки, код, вычисления отображаются -- без этого ваш ноутбук не может быть оценен корректно!
- В Google Colab вам предлагается делать только те части задания, которые идут на кросс-проверку (формата ipynb; проверяются вручную другими вашими однокурсниками). Задания типа юнит-тестирование или ML-решение предлагается делать локально, так как решения не требуют большого запаса вычислительных ресурсов. Также мы выдаем вам тесты и скрипт их запуска, на котором можно будет проверить ваше решение, что тоже гораздо удобнее делать на своем личном компьютере
- Если вы хотите делать задание формата ipynb также на своем личном компьютере, то вам предлагается установить виртуальное окружение с соответствующими версиями библиотек, *которое вам выдадут вместе с первым заданием*. Скорее всего, каких-то проблем с совместимостью в этом случае при выполнении практического задания быть не должно.