|  |  |
| --- | --- |
| 1 | Изучение PySpark Установить pyspark, запустить сессию и создать табличку. Выполнять можно в гугл колаб |
| 2 | Задание в файле: Урок\_2\_HW\_rdd.ipynb Выполнять можно в гугл колаб |
| 3 | Повторить код из урока и еще раз прислать Урок\_2\_HW\_rdd.ipynb |
| 4 | Задание в файле: Домашнее\_задание\_4.ipynb |
| 5 | Доделать Домашнее задание 4 |
| 6 | Изучить нотебуки: работа на уроке , custom estimators, custom transformers Домашняя работа (апдейт): <https://colab.research.google.com/drive/1lpG6gnANFMHA7Kqx2gt9UCpTEiHEOD7U?usp=sharing> |
| 7 | Курсовая: - Написать TfIdf на спарке (потом сделать это же со SparkMl) - Обучить свою линейную регрессию  Титаник: <https://colab.research.google.com/drive/1_LLKMKGT9NuaR64U9370NuGer2DnXeH9?usp=sharing> HW: <https://colab.research.google.com/drive/1QltsS5jJsqfMtHJCyC-PXNM5eenuZK_t?usp=sharing>  Сюда присылаем заглушку, и делаем курсовую.  Данные кредитного скоринга в прикрепленном архиве Данные титаника в csv файлах |
| 8 | Курсовая: - Написать TfIdf на спарке (потом сделать это же со SparkMl) - Обучить свою линейную регрессию  Титаник: <https://colab.research.google.com/drive/1_LLKMKGT9NuaR64U9370NuGer2DnXeH9?usp=sharing> HW: <https://colab.research.google.com/drive/1QltsS5jJsqfMtHJCyC-PXNM5eenuZK_t?usp=sharing>  Данные кредитного скоринга в прикрепленном архиве (7 урок) Данные титаника в csv файлах (7 урок) |