

1. Выбрать один из популярных/любимых алгоритмов машинного обучения (можно взять алгоритм с предыдущей домашки)
  2. Скачать датасет, для которого выбранный алгоритм достаточно эффективен (можно взять алгоритм с предыдущей домашки)
  3. Реализовать выбранный алгоритм с помощью Spark MLlib
    - подготовьте данные для алгоритма, если это необходимо
    - воспользуйтесь стандартной конфигурацией алгоритма, выведите её на экран и посчитайте точность алгоритма
    - с помощью гиперпараметризации подберите более удачную конфигурацию алгоритма, выведите её на экран и посчитайте точность алгоритма в этом случае
    - воспользуйтесь Pipeline механизмом для препроцессинга и работы алгоритма со стандартной конфигурацией / поиска лучших параметров
    - замерьте время работы стандартной конфигурации алгоритма и поиска лучших параметров
  4. Повторить предыдущий пункт с использованием библиотек машинного обучения, не основанных на Spark.
- Хинт: можно взять данные из предыдущей домашки \*roll\_safe\_mem.png\*
5. Сравнить реализации алгоритмов с точки зрения выходных результатов, статистик и полученного удовольствия при написании :)