**Контрольная работа 1 – Вариант 9**

1. Набор данных: diamonds

2. Независимая переменная: features/carat

3. Зависимая переменная: features/y

4. Визуализация для независимой переменной – эмпирическая плотность распределения

5. Визуализация для зависимой переменной – столбчатая диаграмма

6. Показатель качества регрессии – MAE (mean absolute error)

В соответствии с индивидуальным заданием, указанным в записной книжке команды Teams, выполните следующее:

1. Загрузите заданный в индивидуальном задании набор данных из Tensorflow Datasets и оставьте в наборе данных признаки, принимающие непрерывные числовые значения, включая указанные в индивидуальном задании независимую и зависимую переменные. Вычислите матрицу корреляции признаков и выведите названия (номера) пар признаков с наиболее низкой и наиболее высокой корреляцией.
2. Выполните визуализацию независимой и зависимой переменных в соответствии с индивидуальным заданием, подписывая оси и рисунок.
3. Постройте диаграмму рассеяния для независимого и зависимого признаков, подписывая оси и рисунок, определите наличие одиноко расположенные точек и, при наличии, удалите их.
4. Постройте парную линейную регрессию для независимого и зависимого признаков при помощи точного подхода и при помощи нейронной сети с одним нейроном. Определите лучший из двух подходов по показателю качества R^2 (коэффициенту детерминации).
5. Постройте диаграмму рассеяния для независимого и зависимого признаков и изобразите линии двух построенных парных регрессий, подписывая оси и рисунок и создавая легенду для линий регрессии.
6. Разбейте набор признаков на обучающую и контрольную выборки. Создайте и адаптируйте нормализующий слой Tensorflow для всех признаков набора данных (за исключением зависимого признака). Нормализуйте зависимый признак.
7. Используя созданный нормализующий слой и нормализованный зависимый признак, постройте регресоры на базе следующих моделей множественной регрессии:

* линейной регрессии
* гребневой регрессии (L2)
* лассо регрессии (L1)   
    
  Выберите коэффициенты регуляризации l1 и l2 так, чтобы нейронные сети для всех трех моделей обучались (значение ошибки уменьшалось в процессе обучения).

1. Определите на контрольной выборке (с нормализованным зависимым признаком) модель множественной регрессии с наиболее высоким качеством по показателю, указанному в индивидуальном задании, среди построенных моделей.
2. Для лучшего регрессора визуализируйте кривые обучения (в зависимости от эпохи обучения).
3. Определите медианные значения признаков (кроме независимого и зависимого признаков) и для построенных медианных значений визуализируйте на плоскости с независимым признаком в качестве оси абсцисс и зависимым признаком в качестве оси ординат точки тестовой выборки и линии (графики) различных моделей множественной регрессии разными цветами. Подпишите оси и создайте легенду и заголовок для рисунка.

Результат контрольной работы оформить в виде отчета в формате файла Jupiter Notebook (шаблон отчета находится в учебных материалах команды в формате .ipynb). Включите в отчет номер варианта, текст индивидуального задания, пункты 1-10 задания, указанные выше, и программный код для решения этих пунктов. Сопроводите программный код необходимыми комментариями. Дополнительно (кроме файла расширением .ipynb) представить распечатку файла с отчетом в формате PDF. Не архивировать файлы.

Отчет по контрольной работе представить как результат выполнения задания MS Teams (представить файлы PDF и ipynb). **Не забыть нажать на кнопку сдачи задания**.