Федеральное государственное бюджетное   
образовательное учреждение высшего   
профессионального образования

Московский государственный технический   
университет имени Н.Э. Баумана   
(МГТУ им. Н.Э. Баумана)

**ЛАБОРАТОРНАЯ РАБОТА 4  
«Подготовка обучающей и тестовой выборки, кросс-валидация и подбор гиперпараметров на примере метода ближайших соседей»**

**ПО ДИСЦИПЛИНЕ «Технологии машинного обучения»**

Выполнил(а): Орлова С. М.   
 (Фамилия И.О. студента)

РТ5-61   
 (Индекс группы)

Проверил(а): Гапанюк Ю. Е.   
 (Фамилия И.О. преподавателя)

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_ 2020 г.

ЗАЧТЕНО / НЕ ЗАЧТЕНО \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
 (нужное выделить) (подпись)

Москва, 2020

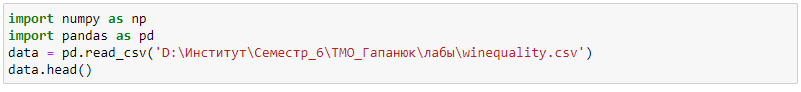
1. Описание задания лабораторной работы

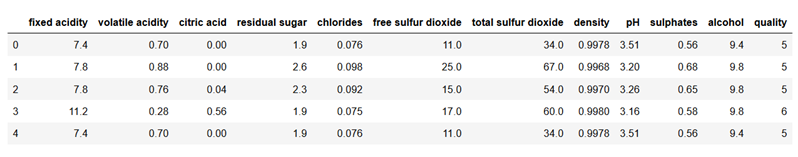
Цель лабораторной работы:  
Изучение сложных способов подготовки выборки и подбора гиперпараметров на примере метода ближайших соседей.

Задание:

1. Выберите набор данных (датасет) для решения задачи классификации или регрессии.
2. С использованием метода train\_test\_split разделите выборку на обучающую и тестовую.
3. Обучите модель ближайших соседей для произвольно заданного гиперпараметра K. Оцените качество модели с помощью подходящих для задачи метрик.
4. Постройте модель и оцените качество модели с использованием кросс-валидации.
5. Произведите подбор гиперпараметра K с использованием GridSearchCV и кросс-валидации.
6. Текст программы и экранные формы с примерами выполнения программы

Загрузка и первичный анализ данных:





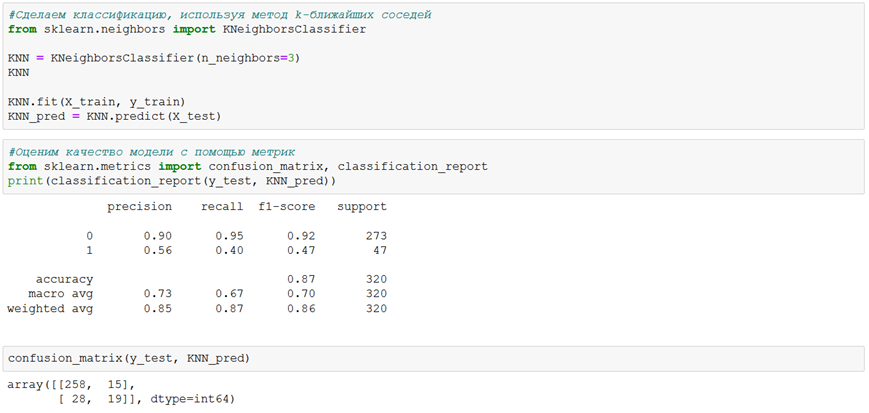




Разделение выборки на обучающую и тестовую:



Обучение модели ближайших соседей для заданного гиперпараметра K:



Построение модели с использованием кросс-валидации:



Подбор гиперпараметра K с использованием GridSearchCV и кросс-валидации:



