

Задание. Знакомство с инструментами сбора и подготовки данных

1. Импортировать датасет (Исходные данные можно выбрать самостоятельно на kaggle.com, либо взять один из файлов, прикрепленных к заданию)

2. Проверить данные на пропуски и наличие аномальных значений, использовать для этого статистические методы и средства графической визуализации(графики распределения, диаграммы, графики типа BoxPlot, тепловые карты и т.п.). Заполнить пропуски и устранить выбросы. Проверить, при помощи графических средств, что пропуски и аномальные значения отсутствуют.

3. Сгенерировать новые признаки (не менее двух), а так же применить кодирование категориальных признаков (применить ONE-кодирование, бинаризацию признаков и др.)

3. Выполнить стандартизацию и/или нормализацию данных.