**Признаковая инженерия**

# Домашнее задание № 3

**Цель работы**. В данном домашнем задании вы потренируетесь в очистке и обогащении данных.

|  |
| --- |
| Уважаемый студент!  Вы построили ваши торговые стратегии на основе технического анализа и провели их оценку. С удивлением вы обнаружили, что некоторые модели работают лучше, некоторые хуже, причем это неравномерно распределено во времени.  Поговорив с коллегами, вы решаете построить более сложные модели машинного обучения, но для их эффективной работы необходимо подготовить и очистить данные. Также постоянно поступают новые данные, и тратить время на их очистку и подготовку непозволительная роскошь. Вы понимаете, что на подготовленных данных вы сможете сформировать новые признаки, которые послужат основой для более сложных моделей.  На основании вышесказанного вам необходимо построить конвейер подготовки данных. Он будет получать сырые данные с информационных площадок, и автоматически готовить их в вид, пригодный для использования в моделях машинного обучения. |

**Вам предлагается на основе представленной информации:**

1. Создать код на Python, который очистит ваши данные от выбросов и заполнит пропуски в них (если они присутствуют).

2. Создать код Python, который сформирует новые признаки из данных для их дальнейшего использования в процессе моделирования.

3. Код должен в автоматическом режиме при получении новой порции данных запустить конвейер очистки и преобразования данных.

4. Подготовленные данные должны сохраняться в виде файлов или БД.

5. Новые данные не должны дублировать предыдущие и не должны их перезаписывать. Только добавлять новые.

6. Сформировать дашборд, показывающий процесс накопления данных и отображающий графики по ним.

**Для того, чтобы вашу работу зачли** необходимо выполнить первые четыре вышеприведенные задания.

***Желаем успехов!***