

Задача А. Нефтегазовый коллапс

При экстренном отключении света, произошел сбой в центральном хранилище данных одной крупной нефтегазовых компаний. К сожалению, информация о типе месторождения по расположению (на берегу или в открытом море) была удалена. Поскольку эта информация крайне важна, Вам необходимо восстановить ее в пострадавшей «test.csv».

В качестве тренировочного датасета используйте неиспорченную базу данных «train.csv».

Входные данные

Тренировочный датасет train.csv, в котором поля обозначают следующее:

- Tectonic regime – тектонический режим месторождения
- Hydrocarbon type – тип углеводорода
- Reservoir status – статус разработки месторождения
- Structural setting – структурные особенности месторождения
- Depth – средняя глубина кровли резервуара
- Period – геологическая система образования резервуара
- Lithology – основной литологический состав коллектора
- Gross
- Netpay
- Porosity – пористость породы
- Permeability – проницаемость породы

Целевая переменная – Onshore/Offshore – тип месторождения по расположению

Выходные данные

Выходной файл, представляет собой текстовый файл из 133 строк, в каждой строке содержится значение типа месторождения по расположению.

Система оценивания

Для оценки качества работы алгоритма используем метрику Accuracy. Итоговый балл за задачу вычисляется по формуле:

$$\text{Балл за задачу} = \text{Accuracy} * 100$$

Обращаем Ваше внимание, что если выходные данные не соответствуют требованиям, участники получают 0 баллов и статус RE.