



Описание данных

Данные (Данные были взяты с конкурса Tinkoff):

* Некоторые значения в колонках могут быть изменены в превентивных целях плагиата

1. **credit_train.csv** - Датасет содержит данные о кредитах на покупку электроники, которые были одобрены
 - a. **client_id** - это уникальный идентификатор для каждого клиента в наборе данных. Этот признак может быть полезен для отслеживания кредитной истории конкретных клиентов.
 - b. **gender** - этот признак указывает пол клиента. Эта информация может использоваться для выявления любого потенциального дискриминационного отношения на основе пола в процессе одобрения кредита.
 - c. **age** - этот признак указывает возраст клиента. Возраст может быть фактором в кредитоспособности, с более взрослыми клиентами, возможно, имеющими более закрепленные кредитные истории.
 - d. **marital_status** - этот признак указывает семейное положение клиента. Эта информация может использоваться для выявления любого потенциального дискриминационного отношения на основе семейного положения в процессе одобрения кредита.
 - e. **job_position** - этот признак указывает должность клиента. Эта информация может использоваться для выявления любого потенциального дискриминационного отношения на основе занятости в процессе одобрения кредита.
 - f. **credit_sum** - этот признак указывает сумму кредита, которую запросил клиент. Эта информация может использоваться для выявления закономерностей в запрашиваемой сумме кредита и для идентификации

клиентов, которые могут быть более рискованными для невыплаты кредита.

- g. **credit_month** - этот признак указывает срок кредита в месяцах. Эта информация может использоваться для выявления закономерностей в длительности кредита и для идентификации клиентов, которые могут быть более рискованными для невыплаты кредита.
- h. **tariff_id** - этот признак указывает тарифный план клиента в банке. Эта информация может использоваться для выявления закономерностей в типах тарифных планов и для идентификации клиентов, которые могут быть более рискованными для невыплаты кредита.
- i. **score_shk** - этот признак является внутренним баллом, который использует банк для определения кредитоспособности. Эта информация может использоваться для выявления закономерностей в кредитоспособности и для идентификации клиентов, которые могут быть более рискованными для невыплаты кредита.
- j. **education** - этот признак указывает уровень образования клиента. Эта информация может использоваться для выявления закономерностей в уровне образования и кредитоспособности и для идентификации клиентов, которые могут быть более рискованными для невыплаты кредита.
- k. **living_region** - этот признак указывает регион, в котором живет клиент. Эта информация может использоваться для выявления закономерностей в региональной кредитоспособности и для идентификации клиентов, которые могут быть более рискованными для невыплаты кредита.
- l. **monthly_income** - этот признак указывает на ежемесячный доход клиента. Эта информация может использоваться для выявления закономерностей в доходе и кредитоспособности и для идентификации клиентов, которые могут быть более рискованными для невыплаты кредита.
- m. **credit_count** - этот признак указывает количество кредитов, которые клиент имеет на момент подачи заявки. Эта информация может использоваться для выявления закономерностей в количестве кредитов и для идентификации клиентов, которые могут быть более рискованными для невыплаты кредита.

п. **open_account_flg** - этот признак указывает на факт открытия кредитного счета клиентом в данном банке. Эта информация может использоваться для выявления закономерностей в истории кредитования клиента и для идентификации клиентов, которые могут быть более доверенными для банка.

Этот набор данных может использоваться для решения задачи бинарной классификации, такой как определение клиентов, которые могут не выплатить кредит или определение кредитоспособности клиента. Этот набор данных также может использоваться для просмотра и исследования данных о кредитах и кредитных историях клиентов