

Тестовое задание - анализ данных

Организация выдает займы физическим лицам.

Информация о ранее выданных организацией займах в файле [new_training_data_31_08_2022.csv](#)

Файл [all_reject_data.csv](#) содержит данные о клиентах, оставлявших заявки на займ, но в итоге не взявших его по причине отказа финансовой организации или нежеланию самого клиента.

Необходимо

1. выполнить анализ и охарактеризовать клиентский портфель организации
2. построить базовую модель прогнозирования банкротства, одобряющую не менее 35% клиентов при банкротстве среди одобренных не выше 15%.
3. подготовить рекомендации и предложения по изменению признакового пространства, использованию внешних данных и иному развитию базовой модели.

Описание признакового пространства

- **age** - Возраст заемщика
- **lastcredit** - Время в днях, которое прошло с момента открытия последнего кредитного продукта
- **time_to_lastcredit_closeddt** - Время в днях, которое прошло с момента закрытия последнего микрокредита (если есть активные кредиты, эта переменная будет равна 0)
- **close_loan_median** - Медиана, взятая по количеству дней между закрытием предыдущего и следующего микрокредита (считается по всем последовательно открытым микрозаймам), т.е. сколько в среднем проходит времени между закрытием предыдущего и следующего микрокредита
- **open_loan_median** - Медиана, взятая по количеству дней между открытием предыдущего и следующего микрокредита (считается по всем последовательно открытым микрозаймам), т.е. сколько в среднем проходит времени между открытием предыдущего и следующего микрокредита
- **is_active_100** - Количество активных кредитов, открытых за все время
- **isnt_active_100** - Доля не возвращенных кредитов относительно всех кредитов, взятых за все время
- **is_lost_100** - Невозвращенные кредиты, открытые за последний месяц (например, переданные по цессии)
- **micro_loans_active_100** - Активная сумма микрокредитов, открытых за всё время
- **is_active_12** - Количество всех активных кредитов, открытых за последние 12 месяцев
- **open_sum_12** - Активная сумма кредитов, взятых за последние 12 месяцев
- **isnt_active_12** - Количество закрытых кредитов, которые были открыты за последние 12 месяцев
- **is_lost_12** - Невозвращенные кредиты, открытые за последние 12 месяцев
- **overdue_loans_12** - Количество просроченных кредитов, открытых за последние 12 месяцев

- **micro_loans_active_12** - Активная сумма микрокредитов, открытых за последние 12 месяцев
- **is_active_3** - Количество всех активных кредитов, открытых за последние 3 месяца
- **open_sum_3** - Активная сумма кредитов, взятых за последние 3 месяца
- **isnt_active_3** - Количество закрытых кредитов, которые были открыты за последние 3 месяца
- **is_lost_3** - Невозвращенные кредиты, открытые за последние 3 месяца
- **overdue_loans_3** - Количество просроченных кредитов, открытых за последние 3 месяца
- **micro_loans_active_3** - Активная сумма микрокредитов, открытых за последние 3 месяца
- **is_active_1** - Количество всех активных кредитов, открытых за последний месяц
- **open_sum_1** - Активная сумма кредитов, взятых за последний месяц
- **isnt_active_1** - Количество закрытых кредитов, которые были открыты за последний месяц
- **is_lost_1** - Невозвращенные кредиты, открытые за последний месяц (например, переданные по цессии)
- **micro_loans_active_1** - Активная сумма микрокредитов, открытых за последний месяц
- **ratio_all_microloans_3_to_12** - Отношение количества микрокредитов, взятых за последние 3 месяца, к количеству микрокредитов, взятых за последние 12 месяцев
- **ratio_overdue_loans_3_to_12** - Отношение количества просроченных микрокредитов, взятых за последние 3 месяца, к количеству просроченных микрокредитов, взятых за последние 12 месяцев
- **ratio_history_100** - Доля не возвращенных кредитов относительно всех кредитов, взятых за все время
- **ratio_history_12** - Доля не возвращенных кредитов относительно всех кредитов, взятых за последние 12 месяцев
- **fraction_last_x_12** - Доля кредитов, взятых за последние 12 месяцев, относительно всех кредитов истории
- **ratio_history_3** - Доля не возвращенных кредитов относительно всех кредитов, взятых за последние 3 месяца
- **fraction_last_x_3** - Доля кредитов, взятых за последние 3 месяца, относительно всех кредитов истории
- **ratio_history_1** - Доля не возвращенных кредитов относительно всех кредитов, взятых за последний месяц
- **fraction_last_x_1** - Доля кредитов, взятых за последний месяц, относительно всех кредитов истории
- **mean_delay_100_with_lag** - Средняя просрочка за всё время (с лагом по времени в 2 месяца)
- **mean_delay_12_with_lag** - Средняя просрочка за последние 12 месяцев (с лагом по времени в 2 месяца)
- **mean_delay_3_with_lag** - Средняя просрочка за последние 3 месяца (с лагом по времени в 2 месяца)
- **mean_delay_1_with_lag** - Средняя просрочка за последний месяц (с лагом по времени в 2 месяца)
- **ratio_mean_delay_3_to_12** - Отношение средней просрочки за последние 3 месяца к средней просрочке за последние 12 месяцев (в днях, с лагом по времени в 2 месяца)
- **count_all_credits** - Количество всех кредитов в истории
- **ratio_pattern_len_to_pattern_1** - Отношение количества платежей в платежном паттерне к общему количеству запланированных платежей на данный момент

- **ratio_pattern_len_to_pattern_2** - Отношение количества просрочек в 0-5 дней в платежном паттерне к общему количеству запланированных платежей на данный момент
- **ratio_pattern_len_to_pattern_3** -
- **ratio_pattern_len_to_pattern_4** -
- **ratio_pattern_len_to_pattern_bad_len** - Отношение количества символов сильной просрочки (> 60 дней) в платежном паттерне к общему количеству символов в строке
- **last_microloan_openeddt** - Время в днях, которое прошло с момента открытия последнего микрокредита
- **is_type_credit_card_100** - Количество кредитов типа "кредитная карта", открытых за всё время
- **is_type_consumer_100** - Количество кредитов типа "потребительский кредит", открытых за всё время
- **is_type_micro_100** - Количество кредитов типа "микрокредит", открытых за всё время
- **is_active_type_credit_card_100** - Количество активных кредитов, открытых за всё время с типом займа "кредитная карта"
- **is_active_type_consumer_100** - Количество активных кредитов, открытых за всё время с типом займа "потребительский кредит"
- **is_active_type_micro_100** - Количество активных кредитов, открытых за все время с типом займа "микрокредит"
- **is_type_credit_card_12** - Количество кредитов типа "кредитная карта", открытых за последние 12 месяцев
- **is_type_consumer_12** - Количество кредитов типа "потребительский кредит", открытых за последние 12 месяцев
- **is_type_micro_12** - Количество кредитов типа "микрокредит", открытых за последние 12 месяцев
- **is_active_type_credit_card_12** - Количество активных кредитов, открытых за последние 12 месяцев с типом займа "кредитная карта"
- **is_active_type_consumer_12** - Количество активных кредитов, открытых за последние 12 месяцев с типом займа "потребительский кредит"
- **is_active_type_micro_12** - Количество активных кредитов, открытых за последние 12 месяцев с типом займа "микрокредит"
- **is_type_credit_card_3** - Количество кредитов типа "кредитная карта", открытых за последние 3 месяца
- **is_type_consumer_3** - Количество кредитов типа "потребительский кредит", открытых за последние 3 месяца
- **is_type_micro_3** - Количество кредитов типа "микрокредит", открытых за последние 3 месяца
- **is_active_type_credit_card_3** - Количество активных кредитов, открытых за последние 3 месяца с типом займа "кредитная карта"
- **is_active_type_consumer_3** - Количество активных кредитов, открытых за последние 3 месяца с типом займа "потребительский кредит"
- **is_active_type_micro_3** - Количество активных кредитов, открытых за последние 3 месяца с типом займа "микрокредит"
- **is_type_credit_card_1** - Количество кредитов типа "кредитная карта", открытых за последние 12 месяцев
- **is_type_consumer_1** - Количество кредитов типа "потребительский кредит", открытых за последний месяц
- **is_type_micro_1** - Количество кредитов типа "микрокредит", открытых за последний месяц

- **is_active_type_credit_card_1** - Количество активных кредитов, открытых за последние 12 месяцев с типом займа "кредитная карта"
- **is_active_type_consumer_1** - Количество активных кредитов, открытых за последний 1 месяц с типом займа "потребительский кредит"
- **is_active_type_micro_1** - Количество активных кредитов, открытых за последний 1 месяц с типом займа "микрокредит"
- **overall_worst_overdue_state_12** -
- **ratio_sum_outstanding_to_open_sum** - Отношение суммы просрочки по всем кредитам к сумме взятых кредитов за всю историю