



**Министерство науки и высшего образования Российской
Федерации Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Московский государственный
технический университет имени Н.Э.
Баумана (национальный
исследовательский университет)» (МГТУ
им. Н.Э. Баумана)**

ФАКУЛЬТЕТ Информатика и системы управления
КАФЕДРА Системы обработки информации и управления

**По курсу «Технологии машинного обучения»
Вариант 27**

Подготовил:
Студент группы ИУ5Ц-
84Б
Распашнов А.А.

Проверил:
Преподаватель кафедры ИУ5
Гапанюк Ю.Е.

Москва, 2023 г.

Тема: Технологии разведочного анализа и обработки данных.

ИУ5-64Б, ИУ5Ц-84Б

Линейная/логистическая регрессия

Градиентный бустинг

Задача №4.

Для заданного набора данных (по Вашему варианту) постройте модели классификации или регрессии (в зависимости от конкретной задачи, рассматриваемой в наборе данных). Для построения моделей используйте методы 1 и 2 (по варианту для Вашей группы). Оцените качество моделей на основе подходящих метрик качества (не менее двух метрик). Какие метрики качества Вы использовали и почему? Какие выводы Вы можете сделать о качестве построенных моделей? Для построения моделей необходимо выполнить требуемую предобработку данных: заполнение пропусков, кодирование категориальных признаков, и т.д.

Метод 1: Линейная/логистическая регрессия

Метод 2: Градиентный бустинг

Набор данных: <https://www.kaggle.com/datasets/fedesoriano/company-bankruptcy-prediction>

Данные были получены из Тайваньского экономического журнала за период с 1999 по 2009 год. Банкротство компании было определено на основании правил ведения бизнеса Тайваньской фондовой биржи.

Столбцы:

Y - Банкрот?: Ярлык класса

X1 - ROA(C) до вычета процентов и амортизации до вычета процентов:
Рентабельность общих активов(C)

X2 - ROA(A) до уплаты процентов и % после налогообложения: Рентабельность
общих активов(A)

X3 - ROA (B) до вычета процентов и амортизации после налогообложения:
Рентабельность всех активов(B)

X4 – Валовая прибыль от операционной деятельности: Валовая прибыль/чистая
выручка от продаж

X5 – Валовая прибыль от реализованных продаж: Реализованная валовая
прибыль/чистая выручка от продаж

X6 – Норма операционной прибыли: $\text{Операционная прибыль} / \text{Чистая выручка}$

X7 – Чистая процентная ставка до налогообложения: $\text{Доход} / \text{чистая выручка}$

X8 – Чистая процентная ставка после уплаты налогов: $\text{Чистая прибыль} / \text{Чистая выручка}$

X9 – Небизнесовые доходы и расходы/выручка: $\text{Коэффициент чистой внебизнесовой прибыли}$

X10 — Непрерывная процентная ставка (после налогообложения): $\text{Чистая прибыль} - \text{исключая прибыль или убыток от выбытия} / \text{Чистая выручка}$

X11 — Ставка операционных расходов: $\text{Операционные расходы} / \text{Чистая выручка}$

X12 — Ставка расходов на исследования и разработки: $(\text{Расходы на исследования и разработки}) / \text{Чистая выручка}$

X13 - Скорость движения денежных средств: $\text{Денежный поток от операционных/текущих обязательств}$

X14 - Процентная ставка по процентным долгам: $\text{Процентные долги} / \text{капитал}$

X15 - Ставка налога (A): $\text{Эффективная налоговая ставка}$

X16 - Чистая стоимость одной акции (B): $\text{Балансовая стоимость На акцию} (B)$

X17 - Чистая стоимость на акцию (A): $\text{Балансовая стоимость на акцию} (A)$

X18 - Чистая стоимость на акцию (C): $\text{Балансовая стоимость на акцию} (C)$

X19 - Постоянная прибыль на акцию за последние четыре сезона: $\text{прибыль на акцию} - \text{Чистый доход}$

X20 - Денежный поток на акцию

X21 — Доход на акцию (в юанях): Продажи на акцию

X22 — Операционная прибыль на акцию (в юанях): $\text{Операционный доход на акцию}$

X23 — Чистая прибыль на акцию до налогообложения (в юанях): Доход на акцию до налогообложения X24

— Валовой реализованный доход Темп роста прибыли

X25 - Темп роста операционной прибыли: Рост операционной прибыли

X26 - Темп роста чистой прибыли после уплаты налогов: Рост чистой прибыли

X27 - Темп роста обычной чистой прибыли: Продолжающийся рост операционной прибыли после налогообложения

X28 - Темп постоянного роста чистой прибыли: Чистая прибыль - Исключая рост прибыли или убытков от выбытия

X29 - Темп роста общих активов: Рост совокупных активов

X30 - Темп роста чистой стоимости: Рост совокупных активов X31 - Коэффициент темпов роста общей доходности активов: Прибыль к совокупным приростам активов

X32 - Реинвестирование денежных средств, %: Коэффициент реинвестирования денежных средств

X33 - Коэффициент текущей ликвидности

X34 - Коэффициент быстрой ликвидности: тест на кислотность

X35 - Коэффициент процентных расходов: процентные расходы/общий доход

X36 - Общий долг/Общий собственный капитал: Коэффициент общих обязательств/капитала

X37 - Коэффициент долга, % : Обязательства/Общие активы

X38 - Чистая стоимость/Активы: Собственный капитал/Общие активы

X39 - Коэффициент пригодности долгосрочного фонда (A): (Долгосрочные обязательства + Собственный капитал)/Основные активы

X40 - Зависимость от заимствования: Стоимость процентного долга

X41 — Условные обязательства/Собственный капитал: Условные

обязательства/Капитал

X42 – Операционная прибыль/Оплаченный капитал: Операционный доход/Капитал

X43 – Чистая прибыль до налогообложения/Оплаченный капитал: Доход/капитал до налогообложения

X44 — Товарно-материальные запасы и дебиторская задолженность/Чистая стоимость: (Запасы+Дебиторская задолженность)/Собственный капитал

X45 — Общий оборот активов

X46 — Оборачиваемость дебиторской задолженности

X47 — Среднее количество дней сбора: Дни непогашенной дебиторской задолженности

X48 — Коэффициент оборачиваемости запасов (в раз)

X49 — Оборачиваемость основных средств Частота

X50 - Коэффициент оборота чистой стоимости (раз): Оборачиваемость собственного капитала

X51 - Выручка на человека: Продажи на одного работника

X52 - Операционная прибыль на человека: Операционный доход на одного работника

X53 - Коэффициент распределения на человека: Основные средства на одного работника

X54 - Общий оборотный капитал Активы

X55 - Быстрые активы/Итого активы

X56 - Оборотные активы/Итого активы

X57 - Денежные средства/Итого активы

X58 - Оборотные активы/текущие обязательства

X59 - Денежные средства/текущие обязательства

X60 - Текущие обязательства по активам

X61 - Оборотные средства по обязательствам

X62 - Запасы/оборотные средства

X63 - Запасы/текущие обязательства

X64 - Текущие обязательства/обязательства

X65 - Оборотные средства/капитал

X66 - Текущие обязательства/капитал

X67 - Долгосрочные обязательства по текущим активам

X68 - Нераспределенная прибыль к общей сумме активов

X69 - Общая прибыль/общая сумма расходов

X70 - Общая сумма расходов/активов

X71 - Коэффициент оборачиваемости
текущих активов: Текущие активы к продажам

X72 - Быстрая оборачиваемость активов Отношение оборотных средств к
продажам

X73 - Коэффициент оборачиваемости оборотного капитала: оборотный капитал
к продажам

X74 - Норма денежного оборота: Денежные средства к продажам

X75 - Денежные потоки к продажам

X76 - Основные средства к активам

X77 - Текущие обязательства к обязательствам

X78 - Текущие обязательства к
капиталу

X79 - Капитал к долгосрочным обязательствам

X80 - Денежные потоки к общим активам

X81 - Денежный поток к обязательствам

X82 - Финансовый директор к активам

X83 - Денежный поток к собственному капиталу

X84 - Текущие обязательства к оборотным активам

X85 - Флаг пассивов-активов: 1, если общая сумма обязательств превышает общую сумму активов, 0 в противном случае

X86 - чистая прибыль к общей сумме активов

X87 - общая сумма активов к цене ВВП

X88 - Интервал без кредита

X89 - Валовая прибыль к продажам

X90 - Чистая прибыль к акционерному капиталу

X91 - Обязательства к капиталу

X92 — Степень финансового рычага (DFL)

X93 — Коэффициент покрытия процентов (процентные расходы к EBIT)

X94 — Флаг чистой прибыли: 1, если чистая прибыль отрицательная за последние два года, 0 в противном случае

X95 — отношение капитала к обязательствам