



Отчет по регрессии

Визуализатор *Отчет по регрессии* для компонента Логистическая регрессия отображает статистические параметры и результаты статистических тестов для анализа регрессионных моделей.

Данный визуализатор состоит из трех областей:

- Информации о модели.
- Коэффициентов регрессии.
- Шагов построения.

Подробно о том, как включить данный визуализатор, можно прочитать в статье.

Интерфейс

Операции

Область Информация о модели

-  **Информация о модели** — отображение/скрытие информации о модели.

Область Коэффициенты регрессии

-  **Таблица** — переключение отображения коэффициентов регрессии в вид *Таблица*.
-  **Дерево** — переключение отображения коэффициентов регрессии в вид *Дерево*.
-  **Нулевые значения** — отображение/скрытие нулевых значений.

Область Шаги построения

-  **Шаги построения** — отображение/скрытие шагов построения модели.

Области визуализатора

Информация о модели

Описание показателей модели представлено в таблице:

Показатель	Описание
0/1 Константа	Информация о том содержит модель константу или нет
9.0 Deviance нулевой модели	Показывает степень соответствия нулевой модели набору данных, на котором она обучается
9.0 Deviance	Показывает степень соответствия модели набору данных, на котором она обучается
9.0 <u>Псевдо</u> <u>R²</u> <u>МакФаддена</u>	Коэффициент детерминации МакФаддена определяется на основе логарифма функции правдоподобия и применяется для оценки степени соответствия модели регрессии реальным данным (изменяется от 0 до 1)
9.0 <u>Псевдо</u> <u>R²</u> <u>МакФаддена</u> (скорр.).	Используется для сравнения моделей с различным числом факторов (изменяется от 0 до 1)
12 Число степеней свободы ошибки	Разность между размером выборки и числом оцениваемых параметров модели
12 <u>Число</u> <u>степеней</u> <u>свободы</u> <u>модели</u>	Число независимо варьируемых переменных модели
9.0 <u>Хи-</u> <u>квадрат</u>	Статистика отношения правдоподобия
9.0 <u>Р-</u> <u>значение</u>	Оценка <u>значимости</u> модели. Чем меньше Р-значение, тем меньше вероятность нулевой гипотезы о том, что нет статистически значимой связи между зависимой и независимыми переменными. Вычисляется по статистике отношения правдоподобия.
9.0 <u>Критерий</u> <u>Акаике</u>	Критерий используется для сравнения моделей с разным числом параметров, когда требуется выбрать наилучший набор объясняющих переменных
9.0 <u>Критерий</u> <u>Акаике</u> (скорр.).	Модифицированный критерий Акаике, применяемый для <u>выборок</u> малого размера, когда отношение числа содержащихся в ней примеров к числу параметров модели меньше 40

Показатель	Описание
9.0 <u>Критерий Байеса</u>	Критерий основан на использовании <u>функции правдоподобия</u> и тесно связан с информационным критерием Акаике
9.0 <u>Критерий Ханнана-Куинна</u>	Критерий основан на использовании функции правдоподобия и является альтернативой критериям Акаике и Байеса

Коэффициенты регрессии

Описание коэффициентов представлено в таблице:

Коэффициент	Описание
Коэффициент	Оценки значений коэффициентов при независимых переменных
Стандартная ошибка	Стандартная ошибка оценки коэффициента
<u>Статистика Вальда</u>	Статистика Вальда для оценки значимости переменной
Р-значение	Оценка значимости переменной. Чем меньше Значимость (Р-значение), тем меньше вероятность нулевой гипотезы о том, что коэффициент при независимой переменной равен 0. Вычисляется по z-статистике.
Отношение шансов	Отношение шансов наступления события к его не-наступлению для каждого увеличения переменной на одну единицу, при условии, что все другие переменные остаются постоянными
Нижняя граница ДИ	Нижняя граница <u>доверительного интервала</u> отношения шансов
Верхняя граница ДИ	Верхняя граница доверительного интервала отношения шансов

Коэффициенты регрессии могут быть представлены в виде  *Таблицы* (см. рисунок 1) или  *Дерева* (см. рисунок 2).

Атрибут
9.0 Константа
9.0 Сумма кредита
9.0 Стоимость кредита
12 Срок кредита
ab Цель кредитования (Ин...
ab Цель кредитования (Оп...
ab Цель кредитования (Оп...
ab Цель кредитования (По...
ab Цель кредитования (Ту...
12 Количество
12 Возраст

Рисунок 1. Режим отображения Таблица

Атрибут
9.0 Константа
9.0 Сумма кредита
9.0 Стоимость кредита
12 Срок кредита
ab Цель кредитования
ab Иное
ab Оплата за образо...
ab Оплата услуг (ме...
ab Покупка и ремон...
ab Турпоездки, разв...
12 Количество

Рисунок 2. Режим отображения Дерево

Шаги построения

Область  Шаги построения имеет смысл включать только в том случае, если при настройке узла Логистическая регрессия в параметре Отбор факторов и защита от переобучения был выбран алгоритм, который подразумевает режим обработки фактов.

Область  Шаги построения состоит из следующих полей:

Поле	Описание
Модель	Дерево, в котором отображается пошаговый процесс отбора факторов для построения финальной модели
Показатель	Значение выбранного показателя для текущей модели
Изменения поля	Отображает добавлен или удален был факт ("+" показывает, что факт был добавлен в модель, а "-" означает, что факт удален из модели)
Поля	Текущие поля модели

Существует три вида моделей:

- **Нулевая модель** — начальная модель работы алгоритма.
- **Промежуточная модель** — модель, которую создал алгоритм в процессе обучения.
- **Финальная модель** — модель, которую алгоритм посчитал наилучшей, и дальнейшие действия не приведут к улучшению модели.

Доступные для выбора показатели:

- Deviance.
- Критерий Акаике.
- Критерий Акаике (скорр.).
- Критерий Байеса.
- Критерий Ханнана-Куинна.