

Нормализация входных векторов (Normalization)

Синонимы: Нормировка значений признаков

Разделы: [Алгоритмы](#)

Logintom: [Нормализация](#)

Нормализация — это процедура предобработки входных данных ([обучающих](#), [тестовых](#) и рабочих множеств), при которой значения признаков, образующих входной вектор, приводятся к некоторому заданному диапазону.

Нормализация необходима потому, что исходные значения [признаков](#) могут изменяться в очень большом диапазоне, и работа аналитических моделей ([нейронных сетей](#), [карт Кохонена](#) и др.) с такими данными может оказаться некорректной.

Так, в одном входном векторе могут содержаться значения, отличающиеся друг от друга на несколько порядков, например, возраст и доход клиента. Данная ситуация может иметь место и для значений одного признака, например, доходы клиентов могут различаться в десятки и сотни раз.

Такой дисбаланс между значениями признаков может вызвать неустойчивость работы модели, ухудшить результаты [обучения](#) и замедлить его процесс. После нормализации все значения входных признаков будут приведены к некоторому ограниченному диапазону (например, $[0...1]$ или $[-1...1]$), что позволит моделям работать с ними более корректно. Нормализация может быть [линейной](#) и [нелинейной](#).

Существует множество способов нормализации значений признаков. К числу наиболее популярных относятся десятичное масштабирование, минимаксная нормализация, нормализация [стандартным отклонением](#) и др.