

# Одноклассовая классификация (One-class classification)

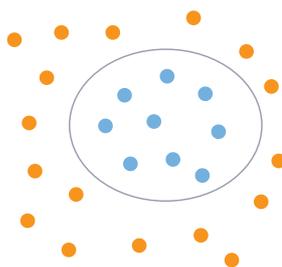
Синонимы: Классификация с одним классом, OCC, Unary classification, Class-modelling

Разделы: [Алгоритмы](#)

Классификация с одним классом (One-class classification — OCC) — задача в машинном обучении, где на множестве наблюдений задан только один класс, объекты которого и требуется обнаружить. Классовая принадлежность остальных объектов не задана или не имеет значения с точки зрения анализа. Данная задача коренным образом отличается от бинарной, мультиклассовой и мультиметочной классификации, где требуется различить два или более класса.

Термин «одноклассовая классификация» был введен Мойей и Хашем в работе «Network constraints and multi-objective optimization for one-class classification».

Главной особенностью OCC является то, что она использует для обучения наблюдения только одного класса, поэтому репрезентативность должна быть обеспечена только для него и необязательна для остальных. Поскольку рассматривается единственный класс, его объекты образуют область, отделенную от других наблюдений некоторой границей, возможно, замкнутой.



В процессе практической работы одноклассовый классификатор получает на вход смесь наблюдений, принадлежащих (имеющих метку) и не принадлежащих (не имеющих метку) классу, на котором он обучался. Задача заключается в том, чтобы ответить на вопрос: относится ли наблюдение к классу, на котором он обучался. Его называют положительным или целевым.

Задача возникает, когда неизвестный объект не принадлежит ни к одному из известных классов, когда объектов нецелевого класса вообще нет, либо их очень мало, или они не образуют репрезентативной выборки.

Одноклассовую классификацию успешно используют для решения следующих задач:

- обнаружение выбросов и аномалий, из которых можно сформировать класс «ненормальных» наблюдений, полагая, что все остальные являются «нормальными»;
- классификация при сильном дисбалансе классов, когда один из них представлен настолько бедно, что его просто нет смысла включать в обучающую выборку.

Типичным примером использования ОСС является классификация катастрофических состояний объектов и систем. Например, в кредитном скоринге популярна задача разделения клиентов на «хороших» и «плохих». Хорошими считаются те, кто не допускает просрочки по кредитам, плохими — те, кто допускает просрочку, но все же возвращает деньги. Кроме этого можно определить класс «заемщики-катастрофы», которые безнадежны. Это единственный класс, который требуется выделить среди всех остальных видов клиентов. При этом сведения о безнадежных заемщиках практически отсутствуют, поскольку в большинстве случаев кредиты им не выдают.

Другой пример — классификация опасных эксплуатационных состояний вредных производств (например, АЭС). Состояний может быть много, но интересовать может только катастрофические для выработки соответствующих решений. Особенно это важно, когда статистика катастрофических состояний неизвестна и такие данные недоступны для анализа.

Есть три варианта обучения одноклассовых классификаторов:

1. Обучение только на положительных примерах.
2. Обучение на положительных примерах и некотором нерепрезентативном количестве отрицательных.
3. Обучение на положительных и неразмеченных примерах.

Существуют несколько методов решения задач ОСС.

**Методы, основанные на плотностях** — оценивают плотность распределения вероятности целевого класса и рассматривают наблюдения, попадающие в области с низкой плотностью, как «аномалии».

Одной из популярных моделей является **одноклассовый Гауссиан**, в которой предполагается, что целевой класс представляет собой выборку из нормального распределения. Вероятность принадлежности к нему можно рассчитать по формуле:

$$p(X) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right\},$$

где  $\mu$  и  $\Sigma$  — математическое ожидание и ковариационная матрица соответственно,  $n$  — размерность задачи. Затем определяется порог вероятности, ниже которого все наблюдения считаются отрицательными.

**Методы, основанные на расстоянии** — используют метрику расстояния как меру сходства между наблюдениями и определяют аномалии на основе их расстояний до целевого класса. Типичным примером является одноклассовый метод К-ближайших соседей (One-Class KNN). Он производит классификацию следующим образом.

Обозначим расстояние между примером целевого класса  $x$  и произвольным объектом  $y$  как  $d(x, y)$ , а расстояние от  $y$  до его ближайшего соседа  $\hat{y}$  как  $d(y, \hat{y})$ . Тогда наблюдение  $\hat{y}$  относится к целевому классу, если выполняется соотношение:

$$\frac{d(x, y)}{d(y, \hat{y})} < \delta,$$

где  $\delta$  по умолчанию равно 1, но может подбираться экспериментально для достижения лучшего результата.

**Методы на основе границ** — формируют границы, которые отделяют наблюдения целевого класса от остальных. Типичным примером являются одноклассовые машины опорных векторов (One-Class SVM). One-Class SVM — это одна из форм классического алгоритма, однако, как следует из названия, для обучения достаточно иметь только один класс.

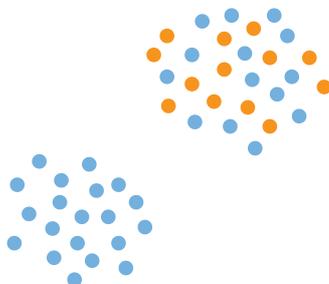
Одноклассовая SVM формирует границу, отсекающую в пространстве признаков целевой класс от остальных наблюдений.

**Методы на основе ансамбля** — объединяют несколько ОСС-классификаторов для повышения общей производительности и точности модели.

**ПУ-обучение** (Positive Unlabeled learning — обучение с положительными примерами без разметки), при котором бинарный классификатор строится с помощью обучения с частичным применением учителя, т.е. только с использованием положительных и неразмеченных наблюдений выборки.

ПУ-алгоритм предполагает, что для обучения доступны два набора примеров: положительный  $P$  и смешанный  $U$ , который содержит как положительные, так и отрицательные примеры, но без соответствующей разметки.

ПУ-алгоритм — это аналог бинарной классификации для случаев, когда имеются размеченные данные только одного из классов, но доступна неразмеченная смесь для обоих. При этом неизвестно, сколько объектов в смеси соответствует положительному классу, а сколько — отрицательному. На основе таких наборов требуется построить бинарный классификатор, аналогичный тому, что строится при наличии чистых данных обоих классов.



На рисунке синими точками представлены размеченные примеры положительного класса, а оранжевыми — неразмеченные. Хотя заранее не известно, как сформирован каждый конкретный неразмеченный объект, можно примерно их разметить, сравнив с

положительными: оранжевые точки, расположенные близко к синим, вероятно, являются положительными. И наоборот, находящиеся далеко — скорее всего отрицательные. Таким образом, несмотря на то что «чистые» отрицательные данные отсутствуют, информацию о них можно получить из неразмеченной смеси и использовать для более точной классификации.

Таким образом, одноклассовая классификация оказывается наиболее полезной в тех задачах интеллектуального анализа данных, где получить данные, связанные с некоторыми классами, сложно, дорого или невозможно вообще.