

# Классификация с учетом издержек (Cost-sensitive classification)

Синонимы: [Классификация, чувствительная к издержкам](#)

Разделы: [Бизнес-задачи](#), [Алгоритмы](#)

LogInom: [Логистическая регрессия \(обработчик\)](#), [Нейросеть \(классификация\)](#), [\(обработчик\)](#).

Классификация с учетом издержек — это случай бинарной классификации, когда издержки ошибок классификации являются неодинаковыми. Под издержками в данном случае понимаются возможные потери (убытки) — финансовые, материальные, временные, репутационные и т.д., вызванные неоптимальными управленческими решениями, принятыми из-за ошибочной классификации.

Типичным примером является классификация заемщиков на кредитоспособных и некредитоспособных. Пусть целью является распознавание некредитоспособных заемщиков (положительный класс). Тогда ложно-положительная (false positive) ошибка, или, в терминах статистики, ошибка I рода — кредитоспособный заемщик распознан как некредитоспособный и принято решение об отказе в кредите. Это грозит потерей только потенциальных процентов — реальных денег банк не теряет.

Ложно-отрицательная (false negative) ошибка, или ошибка II рода — некредитоспособный заемщик распознан как кредитоспособный и принято решение о выдаче ему кредита. Это грозит не только потерей потенциальных процентов, но и всей суммы кредита.

Очевидно, что в данном случае издержки ложно-отрицательных ошибок выше, чем ложно-положительных. Поэтому модель должна минимизировать вероятность ошибок именно II рода.

Обучение классификационных моделей в условиях неравенства издержек обычно связано с несбалансированностью выборки, когда примеров одного класса значительно больше, чем другого. Модели, обученные на несбалансированных выборках, имеют «склонность» относить любое предъявленное наблюдение к более «частому» классу.

Если это приводит к росту числа ошибок с большими издержками, то обучение модели нужно производить таким образом, чтобы минимизировать издержки классификации. Соответствующая технология получила название обучение, чувствительное к издержкам.