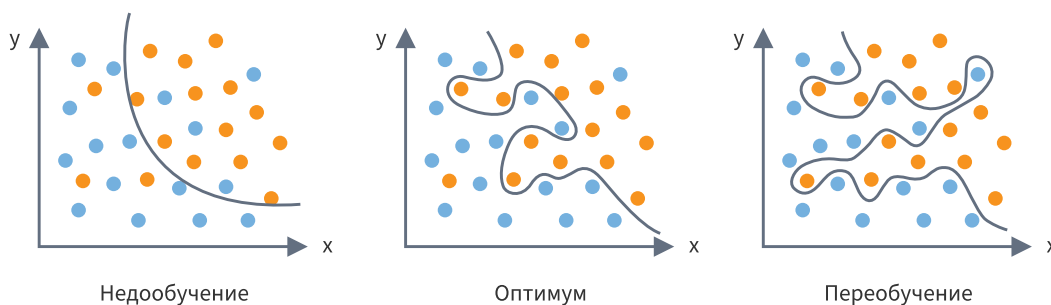


Переобучение (Overtraining)

Loginom: [Нейросеть \(классификация\)_\(обработчик\)](#), [Нейросеть \(регрессия\)_\(обработчик\)](#).

Переобучение — это явление, когда обучаемая модель хорошо распознает примеры из обучающего множества, но при этом не распознает или плохо распознает любые другие примеры, не участвовавшие в процессе обучения (т.е. предъявляемые ей в процессе практического использования).

Переобучение — это результат чрезмерной подгонки параметров модели к зависимостям, содержащимся в обучающем множестве. Если происходит переобучение, то модель не приобретает способности к обобщению — возможности распространять обнаруженные на обучающем множестве зависимости и закономерности на новые данные. Такая модель на практике оказывается бесполезной, даже если ее ошибка обучения мала.



Проблема переобучения в той или иной степени характерна для всех видов обучаемых моделей, но особенно остро она стоит для нейронных сетей. В процессе обучения производится подгонка весов нейронной сети таким образом, что сеть преобразовывала входы к желаемым выходам в соответствии с зависимостями, обнаруженными в обучающих данных.

Однако, если сеть обучать слишком долго или с использованием слишком большого числа параметров (весов), то произойдет переобучение. Это связано с тем, что с определенного момента сеть начинает «подстраиваться» не под общие зависимости в данных, а под особенности отдельных примеров, которые могут содержать аномальные значения, ошибки и т.д.

Как следствие, сеть начнет проверять новые, предъявляемые ей наблюдения, не на соответствие зависимости, а на соответствие отдельным примерам из обучающего множества. В итоге модель сможет распознать новое наблюдение только в том случае,

если оно совпадет с одним из обучающих примеров.

Для того, чтобы избежать переобучения, можно использовать следующие несложные правила.

- 1. Применение тестового множества.** Тестовое множество формируется из обучающего набора данных случайным образом и его примеры подаются на вход сети между обучающими примерами, но корректировку весов сети на вызывают. Вначале обучения ошибка на обучающем, и на тестовом множестве уменьшаются. Но начиная с какого-то момента ошибка на тестовом множестве начинает расти. Это сигнализирует о начале переобучения и необходимости принудительно остановить процесс обучения.
- 2. Использование перекрестной проверки.** Все данные, на которых строится модель, разделяются на k блоков равного размера. При этом обучение производится на $k-1$ блоках, а тестирование — на k -м. Процедура повторяется k раз, при этом для тестирования каждый раз выбирается другой блок. В результате все блоки оказываются используемыми и как обучающие, и как тестирующие.
- 3. Выбор конфигурации нейронной сети.** Нужно выбирать конфигурацию сети (число слоев и нейронов), чтобы количество параметров модели (весов) было в 2-3 раза меньше числа примеров обучающего множества. Если число параметров модели и обучающих примеров окажется соизмеримым, то сеть просто «запомнит» все комбинации вход-выход в примерах обучающего множества, и будет воспроизводить только их, а на новых данных допускать ошибки.