

1. Какие виды перекрестных проверок есть при обучении моделей ИНС?

Ответ: я нашел информацию о 3 видах.

Первый способ – кросс-валидация по K блокам, собственно, то, чем мы и занимались в этой лабораторной. Имеющийся датасет разбивается на K блоков, один из которых берется для тестирования, а остальные K-1 – для обучения. И так повторяется K раз, чтобы каждый блок оказался в качестве тестового. В итоге мы получаем K результатов, которые мы усредняем и получаем конечную оценку.

Второй способ - валидация последовательным случайным сэмплированием. В отличие от первого способа, разбиение на обучающие и тестовые блоки каждый раз производится случайным образом.

Третий способ – поэлементная кросс-валидация. По сути это частный случай обычной кросс-валидации при K равном числу образцов в датасете. Отдельный образец берется в качестве тестового, а остальные идут на обучение сети. Процесс продолжается, пока каждый из образцов не будет использован, как тестовый.

2. Какие ограничения есть на скалярное произведение тензоров?

Ответ: размеры последнего измерения левого тензора и предпоследнего измерения правого должны совпадать.

3. Что такое коэффициент регуляризации?

Ответ: Регуляризация — это способ уменьшения сложности модели для предотвращения переобучения. Отличительной чертой переобучения являются большие веса. И для борьбы с этим вводится регуляризатор:

$$\|w\|^2 = \sum_{i=1}^d w_i^2 \text{ или } \|w\| = \sum_{i=1}^d |w_i|.$$

И теперь стоит задача не:

$$Q(w, X) \rightarrow \min,$$

а

$$Q(w, X) + \alpha R(w) \rightarrow \min,$$

где $R(w)$ – регуляризатор, а α – коэффициент регуляризации.

Необходимо найти баланс нового коэффициента, потому что, если взять α слишком маленьким, то модель может получиться слишком сложной, что опять приведет к переобучению, а, если α сделать слишком большим, то это приведет к слишком простой модели, т.к. если делать α все больше и больше, то в какой-то момент станет оптимальным просто занулить все веса.

4. Почему в задачах регрессии в качестве метрики нельзя использовать mse?

Ответ: вообще mse использовать вроде как можно, но она более чувствительна к выбросам, чем mae, т.е. большее отклонение в данных сильнее скажется на общей ошибке, и не позволяет сделать вывод о точности обученной модели. Поэтому mse чаще используют для сравнения для сравнения двух моделей или в качестве функции потерь при обучении.

5. Правильно ли я понимаю, что у Вас с ходом обучения точность падает?

Ответ: если я правильно понял вопрос, то ассигасу – это условная точность, не метрика. В качестве метрики используется mae, и соответственно, чем выше «точность», тем меньше абсолютная ошибка, что и можно наблюдать на скриншотах в отчете. если вопрос заключается в переобучении, то на некоторых блоках, действительно, наблюдается небольшой эффект переобучения, но, если посмотреть на усредненные данные, результат остается приемлемым.

6. Что такое стандартизация данных?

Ответ: это такой процесс обработки наборы данных, что каждый признак имеет среднее 0 и дисперсию 1. Стандартизация проводится по причине того, что большинство градиентных методов чувствительны к разбросу значений данных. И приведение данных к единому диапазону изменения обеспечивает более корректную работу вычислительных алгоритмов.

7. В каких случаях имеет смысл использовать перекрестную проверку по k блокам?

Ответ: кросс-валидацию применяют в ситуациях, когда нужно проверить работу модели на практике, имея ограниченный набор данных. Если обучать и оценивать модель на одних и тех же данных, то модель будет плохо работать с неизвестными ей данными(переобучение).