

1. Какой вид ансамблирования Вы использовали, что для ансамбля точность выше, чем отдельно для сетей?

Ансамблирование было выполнено как среднеарифметическое предсказаний тестовых данных. Суть метода ансамблирования заключается в объединении прогнозов, полученных набором разных моделей, для получения лучшего прогноза, поэтому для ансамбля точность выше, чем отдельно для сетей.

2. Что такое L_2 регуляризация, в чем отличие от L_1 ?

L_2 регуляризация (регуляция Тихонова) решает некорректно поставленную задачу или предотвращает переобучение модели путём запрета на несоразмерно большие весовые коэффициенты. При L_2 регуляризации дополнительный член является квадратичной функцией (при L_1 регуляризации – модулем).

$$L_2 = \sum_i^n (y_i - y(t_i))^2 + \lambda \sum_i^n a_i^2$$

λ в уравнении является гиперпараметром, который контролирует интенсивность штрафа.

Когда $\lambda \rightarrow 0$, результаты аналогичны линейной регрессии

Когда $\lambda \rightarrow \infty$, все особенности уменьшены до 0.

Отличия:

- 1) L_1 штрафует сумму абсолютных значений весов, а L_2 штрафует сумму квадратичных весов.
- 2) L_1 имеет разреженное решение, L_2 имеет не разреженное решение.
- 3) L_1 имеет несколько решений, а у L_2 есть одно решение.
- 4) L_1 имеет встроенный выбор функций, L_2 не имеет выбора функций.
- 5) L_1 устойчив к выбросам, L_2 не устойчив к выбросам.

- 6) L_1 генерирует модели, которые просты и понятны, но не могут выучить сложные шаблоны, а L_2 дает лучший прогноз, когда выходная переменная является функцией всех входных функций.

3. Для чего может добавляться шум к весам в ходе обучения?

Выходные метки набора данных могут содержать некоторое количество ошибок, из-за которых сеть может обучаться не верно. Добавление шума к меткам может повысить устойчивость сети к таким ошибкам.