

1. Какими существенными недостатками обладает функция `relu`?
Большой градиент, проходящий через `relu`, может привести к такому обновлению весов, что данный нейрон никогда больше не активируется.
Если это произойдет, то, начиная с данного момента, градиент, проходящий через этот нейрон, всегда будет равен нулю.
2. Что такое L1-регуляризация?
Отбор наиболее важных факторов (сильно влияют на результат)
3. С увеличением количества нейронов на слое, увеличивается ли степень нелинейности модели?
Не увеличивается.
4. Отчет крайне не подробный, за него сразу -0.5 балла
:((
5. Что происходит на строчках 23-27?
Нормировка данных, после этого среднее значение по каждому признаку равно 0, стандартное отклонение по каждому признаку = 1
6. Почему в качестве метрики используете `mae`?
Понятие точности не применимо для регрессии, поэтому используется оценка в виде абсолютного отклонения значения от ожидаемого результата.
7. Имеет ли смысл применять проверку по k блокам при большом размере выборки?
Это достаточно затратное мероприятие по вычислениям, так что если размер выборки действительно большой – нет.