**1 Что такое переобучение? Из-за чего оно возникает? Как его померить?**

Переобучение — это такое явление, когда модель начинает искать смысл там, где его нет.

Оно возникает из-за того, что при обучении модели происходит «натаскивание» на тренировочный набор данных, тем самым модель начинает выдавать правильные ответы на тренировочных данных, а на тестовом наборе все очень плохо

Померить его можно с помощью процента верных ответов на тестовом наборе данных

**Как бороться с переобучением?**

Уменьшать вероятность верного определения на тренировочной выборке

**Что такое регуляризация? Какие знаешь?**

**Почему l1 l2 регуляризации помогают с переобучением?**

**Что такое локальный минимум? Чем отличается от глобального?**

Локальный минимум — это область функции, где производная равна нулю

Отличие от глобального заключается в том, что глобальный минимум является самым меньшим на области существования функции потерь, а локальный только в окрестности своего минимума

**Необходимое условие минимума функции одной переменной? Достаточное?**

**Может ли у функции быть нулевая производная в точке, но нет экстремума? Пример**

Может быть. Пример y=5.

**Когда применяем градиентный спуск, в общем случае к какому минимуму сходится модель?**

В общем случае к локальному

**Какие плюсы/минусы стохастического градиентного спуска, относительно градиентного?**

**Приведите примеры алгоритмов глобальной оптимизации**

Adam, Nesterenko

**Для каких функций, находя локальный минимум, можно быть уверенным, что он глобальный?**

Парабола с ветками вверх, то есть квадратичная функция,

**Что такое кросс-валидация, чем она лучше, чем разбиение на обучение и валидацию?**

**Получили метрики двух моделей на кросс-валидации, как отличить значимые улучшения от шума?**

**Что такое precision, recall, F-мера?**

**Напишите формулу MSE, MAE**

MSE=sqrt((sum((Xi-Xm)^2)/i)

**Напишите формулу кросс-энтропии для k классов**

**Что такое SoftMax и зачем он нужен?**

Это функция активации. Софтмакс применяется для решения задачи классификации. С помощью данной функции активации мы имеем на выходе вероятность того, что входная информация относится к данному классу.