1. Назовите основные проблемы машинного обучения?
2. Недостаточно обучающих данных
3. Нерепрезентативность обучающих данных
4. Переобучение
5. Недообучение
6. Мусор в обучающих данных
7. Выбор существенных признаков
8. В чем разница между параметрами модели и гиперпараметрами?

Параметры модели характеризуют данные, на которых обучается модель. Гиперпараметры характеризуют саму модель, определяют ее архитектуру, скорость обучения и т.д.

1. Что такое помеченный обучающий набор? В каких моделях используется такой набор?

Помеченный обучающий набор используется в обучении с учителем и частичном обучении. Такой набор сразу содержит «ответы» для каждой записи, которые должна дать модель для конкретной записи.

1. Что такие динамическое обучение?

Суть динамического обучения заключается в одновременной работе и обучении. Видится использование такого обучения в предсказании каких-то результатов с последующей сверкой предсказанного результата с тем, который в итоге получился.

1. Как вы думаете задача кредитного скоринга эта задача обучения с учителем или без учителя?

Я думаю, что это задача обучения с учителем. Есть параметры, по которым можно представить клиента, и нужно получить какую-то кредитную оценку этого клиента, т.е. фактически это задача регрессии.

1. Как вы думаете задача определения спама эта задача обучения с учителем или без?

Я думаю, что это тоже задача обучения с учителем. Здесь задача классификации. Данными могут быть разного рода письма, комментарии; а меткой для данных будет выступать флаг спам/не спам.

1. Какой алгоритм можно применить что бы разделить данные на несколько групп, которые обладают схожими свойствами?

Обучение без учителя.

1. Как вы понимаете определение «машинное обучение». Напишите своими словами.

Сфера, задачами которой являются написание алгоритмов, не напрямую решающих поставленную задачу, а «обучаться» решать эту задачу, т.е. алгоритмы имеющие свойство со временем улучшать качество решения задачи.

1. Какой алгоритм обучения вы бы использовали если бы хотели научить ИИ играть в Доту2\ориентироваться робота на неизвестной территории?

Обучение с подкреплением.

1. Каковы две наиболее распространенные задачи обучения с учителем?

Классификации и регрессии.

1. Какое разделение профессий в области DS вы знаете?

Data analytic, data scientist, data engineer.

1. Что такое переменные окружения?

Переменные окружения – это специальный набор переменных, инициализирующиеся вместе с сессией системы(окружения).