ДЗ\_01 Романовский А.Н.

1. Основные проблемы машинного обучения:

недостаточное количество данных, зашумленность данных, пропуски данных, переобучение, недообучение, непрезентативные данные (в данных нет зависимостей и закономерностей), несущественные признаки

2. Параметры модели - изменяются в течении обучения модели, данные параметры изменяются для получения максимально близкого ответа к верному. Гиперпараметры модели - это характеристики, которые влияют на саму модель, т.е. являются настройками модели и задаются в начале обучения.

3. Помеченный обучающий набор - это набор данных, который содержит желаемое решение для модели. Данные наборы, поставляются модели для обучения с учителем. Типичной задачей является классификации.

4. Динамическое обучение — это такое обучение модели, при котором модель обучается постепенно за счет постепенного/непрерывного предоставления ей данных.

5. Задачей кредитного скоринга - является задача классификации. Т.к изначально выборка данных является размеченной смею предположить, что данная является задачей обучения с учителем.

6. Задача определения спама - является задачей обучения с учителем. Модель обучается на размеченных данных, содержащих желаемый результат.

7. Для разделения схожих данных на несколько групп, подходит алгоритм кластерезации. Алгоритмы кластеризации разбивают множества на группы имеющие схожие свойства.

8. Машинное обучение - обучение компьютера, на основе математических алгоритмов, для получения закономерностей, прогнозирования, действия.

9. Для обучения ИИ играть или ориентироваться в пространстве, подходит обучение с подкреплением. Обучение с подкреплением может наблюдать за средой, выбирать и выполнять действия, выдавая в ответ награды или штраф. Затем сама система определяет наилучшую стретегию.

10. Две наиболее распространенные задачи обучения с учителем - это задачи классификации и регрессии

11. Разделение профессий в DS:

data analytic, data science, data engineer

12. Переменные среды - это переменные которые хранят в себе информацию, которую могут использовать запускаемые программы. Данные переменные содержат настройки: системы, графической оболочки, интерактивной оболочки.