

## Вопросы к экзамену по курсу «Машинное обучение»

1. Определение машинного обучения (ML). Взаимосвязь ML с AI и Data Science. Классификация методов ML. Примеры задач.
2. Этапы машинного обучения. Библиотеки для ML. Классификация методов обучения с учителем.
3. Определение задачи регрессии. Функция ошибки. Пример решения для полинома заданной степени.
4. Регуляризация. Цели регуляризации. Вывод формулы весов для гребневой регрессии.
5. Теорема Байеса в ML. Оценка правдоподобия для нормального закона распределения. Функция ошибки.
6. Задача регрессии для случая нормального шума.
7. Вывод уравнения регрессии для нормального априорного распределения коэффициентов полинома (MAP-вывод). Байесовская аппроксимация кривой.
8. Выбор модели в ML. No free lunch theorem. Перекрёстная проверка. Информационный критерий Акаике.
9. Проблемы высокой размерности в ML и пути их решения.
10. Принятие решений в машинном обучении. Пример оценки потерь в задаче классификации.
11. Достоинства и недостатки принятия решений в машинном обучении на основе генеративных моделей, дискриминантных моделей и дискриминантных функций.
12. Принятие решений в задаче регрессии. Виды функций потерь.
13. Определение энтропии. Условная энтропия. Дивергенция Кульбака — Лейблера. Взаимная информация.
14. Непараметрические методы оценки плотности распределения. Функции ядра. Оценка плотности ближайших соседей.
15. Метод K-ближайших соседей.
16. Модели линейной регрессии с базисными функциями. Определение. Разновидности базисных функций.
17. Обоснование метода наименьших квадратов через максимальное правдоподобие.
18. Регуляризованный метод наименьших квадратов. Сравнение лассо и ридж регрессии.
19. Разложение математического ожидания ошибки. Выбор коэффициента регуляризации.
20. Байесовская линейная регрессия. Эквивалентное ядро.

21. Сравнение байесовских моделей. Соображения по поводу выбора лучшей модели.
22. Определение линейной модели классификации. Подходы к классификации. Геометрическая интерпретация линейного классификатора.
23. Линейный дискриминант Фишера.
24. Персептрон. Функция ошибки. Процедура обучения.
25. Вероятностные генеративные модели в задачах классификации. Многоклассовое обобщение логистической сигмоидальной функции.
26. Вероятностная генеративная модель для нормального распределения. Решение методом максимального правдоподобия.
27. Вероятностные дискриминантные модели. Логистическая регрессия. Функция кросс-энтропии.
28. Поиск весов в логистической регрессии. Итеративный метод наименьших квадратов с пересчетом весов.
29. Пробит-регрессия.
30. Метрики оценки качества классификации.
31. Деревья решений в машинном обучении.
32. Смесь экспертов в машинном обучении.
33. Метод K-средних в машинном обучении. K-медоиды.
34. Метрика Силуэт. Метод Локтя.
35. EM алгоритм для гауссовых смесей.

Для каждого метода следует привести его достоинства и недостатки. В каждом вопросе следует привести пример задачи, в которой используется данная теория. При описании моделей, методов, алгоритмов, метрик следует использовать математические формализмы. Необходимо приводить обоснование методов если оно было дано в лекциях.

Каждый билет содержит 2 теоретических вопроса.