

Группа: 674, магистратура	Курс: Выбор моделей в машинном обучении
Дата: 07.11.2011, 14:20–16:00	Преподаватель: к.ф.-м.н., доц. В.В. Стрижов

Студент: _____

При ответе на вопросы желательно сразу писать формулы, а потом пояснения к ним. Список обозначений прилагается. Время на экзамен — лекционная пара. Каждый ответ дает два балла из ста.

1. Дать определение модели.
2. Дать определение многоуровневой модели.
3. Что такое гипотеза порождения данных?
4. Вид нормального уравнения для случая существенно-нелинейной регрессионной модели.
5. Поставить задачу выбора модели линейной регрессии, используя при постановке понятие внешнего и внутреннего критериев.
6. Поставить задачу выбора признаков обобщенно-линейной регрессии
7. Поставить задачу выбора признаков и объектов обобщенно-линейной регрессии
8. Поставить задачу прогнозирования следующего отсчета периодического временного ряда, (периодика одна).
9. Представить задачу прогнозирования k последовательных значений периодического временного ряда в виде задачи выбора многоуровневой модели.
10. Поставить задачу сэмплированного пространства параметров модели на функцию ее ошибки.
11. Парето-оптимальный фронт, определение.
12. Поставить задачу Парето-расслоения, используя понятие недоминируемого объекта.
13. Что является точкой на ROC-кривой в терминах регрессионной выборки?
14. Поставить задачу линейного согласования экспертных оценок, выставленных в линейных шкалах.
15. Поставить задачу квадратичного согласования экспертных оценок, выставленных в линейных шкалах.

16. Поставить задачу отыскания наименьшего угла, принадлежащего соответствующим многогранным выпуклым конусам (задача согласования экспертных оценок, выставленных в ранговых шкалах).
17. Зафиксируем исходное множество признаков в линейной регрессии. Требуется описать в индексной нотации многоуровневый алгоритм МГУА (на каждом уровне порождаются линейные комбинации пар моделей, выбираются лучшие).
18. Перечислите основные критерии выбора моделей, желательно с формулами.
19. Как вычисляется число обусловленности матрицы?
20. Показать, что число обусловленности матрицы $X^T X$ есть квадрат числа обусловленности матрицы X .
21. Поставить задачу отыскания регрессии времени до исполнения и цены исполнения опциона на подразумеваемую волатильность $\sigma_{t,K}^{\text{imp}} = \arg \min_{\sigma} (C_{t,K}^{\text{hist}} - C(\sigma, P_t, B, K, t))$.
22. Поставить задачу нахождения эмпирического распределения параметров модели на предполагаемую плотность распределения вектора параметров.
23. Аппроксимация Лапласа, кратко — принцип, желательно с формулами.
24. Апостериорное распределение параметров регрессионной модели на первом уровне байесовского вывода.
25. Функция ошибки общего вида для нормального распределения зависимой переменной.
26. Функция ошибки общего вида для биномиального распределения зависимой переменной.
27. Сравнение моделей на втором уровне байесовского вывода, формула с пояснениями.
28. Правдоподобие модели, формула.
29. Отличия наиболее вероятных, наиболее правдоподобных и оптимальных параметров модели.
30. Что такое множитель Оккама?
31. Выполнение каких условий предполагается для получения состоятельных и несмещенных оценок параметров методом наименьших квадратов для линейных моделей?
32. Алгоритм Левенберга-Марквардта, формула вычисления параметров.

33. Перечислите предположения о видах ковариационных матриц, используемых в функции ошибки общего вида при гипотезе нормального распределения зависимой переменной.
34. Как используется фактор инфляции дисперсии для снижения мультиколлинеарности признаков?
35. Основная идея ступенчатой регрессии.
36. Основная идея добавления признаков с ортогонализацией.
37. Основная идея шаговой регрессии.
38. Основная идея лассо.
39. Основная идея LARS.
40. Основная идея оптимального прореживания.
41. Описать стохастический (генетический) алгоритм отыскания оптимального набора признаков.
42. Поставить задачу группировки категорий при использовании номинальных признаков.
43. Перечислить внешние критерии МГУА, желательно с формулами.
44. В поперечном сечении трубы равномерно сделали измерения координат стенок. Считаем, что число измерений достаточно для решения задачи. Нужно найти центр и радиус трубы. Требуется поставить задачу линейной регрессии.
45. Для предыдущего случая — сечение не поперечно и не продольно (другими словами труба — эллипс). Требуется поставить задачу.
46. Для случая круглой трубы — требуется отличить стандартную трубу от деформированной.
47. Поставить регрессионную задачу для случая квадратной трубы.