

Вопросы к экзамену по спецкурсу

«Байесовские методы машинного обучения», 2013

В билете три вопроса: два вопроса из основной части и один вопрос из секции ликбеза.

Основная часть

1. Различные постановки задач машинного обучения. Основные проблемы теории распознавания образов: переобучение, противоречивость информации, малый объем выборки.
2. Вероятностные модели. Основные задачи, решаемые с помощью вероятностных моделей. Примеры.
3. Байесовский подход к теории вероятностей. Примеры байесовских рассуждений.
4. Байесовский подход для игры «Акинатор».
5. Задача выбора модели на примере выбора коэффициента регуляризации, ядровой функции, настройки структурных параметров алгоритма обучения. Основные методы выбора модели.
6. Решение задачи выбора модели по Байесу. Обоснованность модели. Полный байесовский вывод.
7. Одномерное и многомерное нормальное распределение. Его основные свойства.
8. Вероятностная модель линейной регрессии. Метод релевантных векторов для задачи регрессии.
9. Логистическая регрессия и метод релевантных векторов для задачи классификации. Алгоритм обучения с использованием нижних оценок, зависящих от параметра.
10. EM-алгоритм в общем виде. EM-алгоритм как покоординатный подъем. Примеры применения.
11. Вариационный подход для приближенного Байесовского вывода.
12. Методы Монте Карло для оценки вероятностных интегралов в байесовском подходе. Методы генерации одномерной случайной величины. Теоретические свойства марковских цепей.
13. Методы Монте Карло по схеме марковских цепей. Схема Гиббса.
14. Задача уменьшения размерности в данных. Вероятностная модель главных компонент, ее обучение с помощью метода максимального правдоподобия и EM-алгоритма.
15. Задача уменьшения размерности в данных. Байесовская модель главных компонент для автоматического выбора размерности редуцированного пространства. Модель смеси главных компонент, примеры применения.
16. Тематическая модель LDA. Вариационный EM-алгоритм для настройки параметров LDA. Примеры использования.

Секция ликбеза

1. Основные понятия математической статистики, функция правдоподобия, метод максимального правдоподобия, его недостатки.
2. Условная вероятность. Правило суммы и произведения для вероятностей. Формула Байеса. Условная независимость случайных величин.
3. Теорема Шеннона и оптимальная длина описания.
4. Гамма-распределение. Сопряженные распределения, их свойства, примеры.
5. Распределение Дирихле, его основные свойства.
6. Матричные вычисления, основные матричные тождества. Примеры применения.
7. Тождество Вудбери и лемма об определителе матрицы.
8. Дивергенция Кульбака-Лейблера, ее использование для поиска аппроксимирующих распределений.