

Машинное обучение, осень 2014/2015

Вопросы по семинарской части

1. Метрика. Определение и примеры. Метрика Минковского, Махалонобиса, Джаккарда, косинусное расстояние. Примеры метрик на строках.
2. Locality-sensitive hashing. Определение. Семейства хэш-функций для меры Джаккарда, косинусного расстояния и евклидовой метрики.
3. Locality-sensitive hashing. Определение. Построение композиций хэш-функций.
4. Критерии информативности для решающих деревьев. Критерий ошибки классификации, критерий Джини, энтропийный критерий.
5. Градиент и его свойства. Примеры дифференцирования по вектору и по матрице. Дифференцирование логарифма определителя матрицы.
6. Регуляризация. Примеры регуляризаторов. Разреживающие регуляризаторы. Байесовская интерпретация регуляризации.
7. Задачи условной оптимизации. Лагранжиан и двойственная задача. Теорема Куна-Таккера.
8. Метод опорных векторов. Многоклассовый случай, способы обобщения.
9. Метод опорных векторов. Вывод двойственной задачи. Связь метода ближайших соседей и SVM с гауссовским ядром.
10. Ядровые методы. Использование ядер в линейной регрессии и методе опорных векторов.
11. Ядровые методы. Способы построения ядер. Полиномиальные и гауссовские ядра.
12. Метод главных компонент как поиск проекционной плоскости.
13. Байесовские методы. Оптимальные решающие правила для бинарной и квадратичной функций потерь.
14. Нормальное распределение. Одномерное нормальное распределение, центральная предельная теорема и «правило трех сигм». Многомерное нормальное распределение, его геометрия. Оценки максимального правдоподобия для среднего и ковариационной матрицы многомерного нормального распределения.
15. Линейный дискриминант Фишера как поиск одномерного представления выборки.