

Теоретические вопросы по решающим деревьям и методу ближайших соседей  
Решающее дерево.

1. Что такое решающее дерево? Как по построенному дереву найти прогноз для объекта?
2. Зачем в вершинах нужны предикаты? Какие типы предикатов вы знаете? Приведите примеры.
3. Почему для любой выборки можно построить решающее дерево, имеющее нулевую ошибку на ней?
4. Почему не рекомендуется строить небинарные деревья (т.е. имеющие больше двух потомков у каждой вершины)?
5. Как устроен жадный алгоритм построения дерева? Какие у него параметры?
6. Зачем нужны критерии информативности?
7. Как задается критерий ошибки классификации? Критерий Джини? Энтропийный критерий? Какой у них смысл?
8. Как задается критерий информативности, основанный на среднеквадратичной ошибке, в задачах регрессии
9. Какие критерии останова вы знаете?
10. Что такое стрижка дерева?

Метод ближайших соседей

1. В чем суть метода ближайших соседей? Как данный метод решает задачи классификации и регрессии? Приведите пример
2. Какие гиперпараметры, есть у метода ближайших соседей? Назовите метрики, которые можно применить в данном методе?
3. Расскажите, что такое ядра в методе ближайших соседей.