

Контрольные вопросы для подготовки к опросу по теме

Классификация

1. Объясните постановку задачи классификации. Объясните типовой процесс классификации и его этапы: обучение модели, оценка модели, применение модели.
2. Объясните идею и способ построения деревьев решений для классификации данных. Дайте определения следующих мер оценки доли примесей в узле дерева решений: индекс Джини, энтропия. Объясните схему работы алгоритмов классификации ID3, C4.5, CART.
3. Сформулируйте теорему Байеса. Объясните, почему на практике, как правило, Байесовская классификация дает невысокую точность классификации. Сформулируйте проблему нулевой вероятности в Байесовской классификации и ее решение.
4. Объясните идею и схему работы классификации по ближайшим соседям.
5. Дайте определения показателей качества классификации: аккуратность (accuracy), точность (precision), полнота (recall), F-мера.
6. Объясните идею ансамблевой классификации и работу следующих ансамблевых методов: бэггинг, бустинг, случайный лес. Назовите преимущества и недостатки каждого метода.