

Лабораторная работа №2 – Обучение под наблюдением

Техники контролируемой классификации

Изучите следующие 5 техник контролируемой классификации:

- **Основанные на критерии Байеса:** вероятностный классификатор, как правило, бинарный (две категории), основанный на гипотезе независимости входных признаков; каждой категории присваивается вероятность, а выход классификатора – это наиболее вероятная категория (оптимизация осуществляется на основе обучающего набора данных).
- **Основанные на функциях:** проблема классификации моделируется с помощью функционального представления входных данных, которое оптимизируется с использованием обучающего набора. Один из методов, основанных на функциях, – Support Vector Machines (SVM), который оптимизирует гиперплоскость для разделения данных в пространстве признаков.
- **Основанные на правилах решения:** базируются на генерации и оптимизации набора правил типа "если – то", используя обучающие данные; правила не обязательно являются взаимоисключающими.
- **Основанные на деревьях:** представляют информацию в виде древовидной структуры, где категории представлены листьями дерева, узлы соответствуют атрибутам данных, а ветви представляют собой отношения значений атрибутов, ведущие к конкретному узлу.
- **Мета-методы:** объединяют несколько существующих классификаторов; обычно обучаются итеративно, добавляя набор "слабых" классификаторов к глобальному классификатору; по мере добавления данных, они переоцениваются: ошибочно классифицированные данные получают больший вес, а корректно классифицированные – меньший, чтобы следующий слабый классификатор фокусировался на ошибочно классифицированных данных.

Выберите для реализации (на языке Python) один метод из каждой вышеуказанной категории (или как минимум четыре). Объясните концепцию и основные шаги каждого алгоритма. Выберите или создайте набор данных в соответствии с требованиями каждого разработанного алгоритма.

Примечание! Оцените полученные результаты (precision, recall, f1-score, confusion_matrix).

Проанализируйте и прокомментируйте полученные результаты.

```
from sklearn.metrics import classification_report, confusion_matrix,  
accuracy_score
```