

Лабораторная работа №1.

Мета-признаки

Задание:

1. Скачайте архив с наборами данных для задачи классификации или изучите API сайта OpenML.
2. Реализуйте несколько мета-признаков из каждой группы: базовые, статистические и структурные.
3. Выберите один набор данных. Измените в нём порядок строк, столбцов и категорий. Убедитесь, что мета-признаки от этого не изменились.
4. Выберите не менее трёх алгоритмов обучения с учителем и меру оценки алгоритма обучения (валидацию).
5. Постройте мета-набор данных. Для каждого набора данных определите наилучший алгоритм для него.
6. Визуализируйте мета-набор данных проекцией на двумерном графике. Используйте разные цвета для разных мета-классов.
7. Протестируйте несколько алгоритмов обучения с учителем (включая наивный алгоритм) на полученном наборе данных.

Примечание:

- Используйте небольшие наборы данных, чтобы 5-й шаг не занял слишком много времени. Но в итоге должно быть использовано не менее 300 наборов данных.
- Если на 5-м шаге набор данных получился сильно несбалансированным, попробуйте использовать другие алгоритмы или гиперпараметры алгоритмов обучения. Но не стоит пытаться настраивать гиперпараметры под каждый набор данных.
- На 7-м шаге алгоритмы обучения с учителем могут отличаться от используемых на 5-м шаге алгоритмов. Также можно использовать другие гиперпараметры для них.