

Наименование работы:

Композиции алгоритмов

Цель работы:

Ознакомиться с методами построения композиций базовых алгоритмов

Задачи:

- a) В соответствии с вариантом выбрать набор данных;
- b) Обучить по одной модели для каждого типа ансамблирования алгоритмов:
 - 1. boosting;
 - 2. bagging;
 - 3. случайный лес;
 - 4. голосование;
- c) С помощью обученных моделей сделать прогноз для тестовых данных;
- d) Оценить прогнозы по метрикам из лаб. №6 или лаб. №7;
- e) Объединить обученные ранее модели в ансамбли с помощью:
 - 1. блендинга;
 - 2. стекинга;
- f) Сравнить методы ансамблирования алгоритмов и сделать выводы. Уметь

объяснить базовое устройство каждого из них.

Для реализации рекомендуется использовать язык программирования Python 3.x и библиотеки Pandas, Numpy, Sklearn

Оформление результатов:

Результаты лабораторной работы оформляются в виде отчета в формате PDF.

Структура отчета:

- 1. Титульный лист;
- 2. Основная часть;
- 3. Заключение.

В основной части приводится описание выполнения каждой из поставленных задач в виде текста и скриншотов программного кода

В заключении приводятся практически значимые выводы по проделанной работе

Таблица 1 - наборы данных

| № | Датасет | Описание |
|---|---|--|
| 1 | Walmart Recruiting: Trip Type Classification https://www.kaggle.com/c/walmart-recruiting-trip-type-classification/data | Данные о посещениях магазина Walmart покупателями. Целевой признак: тип посещения магазина. |
| 2 | IEEE-CIS Fraud Detection https://www.kaggle.com/c/ieee-fraud-detection/data | Данные об онлайн-транзакциях. Целевой признак: является ли транзакция мошеннической |
| 3 | Home Credit Default Risk https://www.kaggle.com/c/home-credit-default-risk/data | Данные о заемщиках банка. Целевой признак: способен ли заёмщик выплатить кредит |
| 4 | House Prices: Advanced Regression Techniques https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data | Данные о продаваемых домах и их характеристиках. Целевой признак: цена дома |
| 5 | Restaurant Revenue Prediction https://www.kaggle.com/c/restaurant-revenue-prediction/data | Данные о ресторанах, их местоположении, типе и т.д. Целевой признак: выручка ресторана за год |
| 6 | Sberbank Russian Housing Market https://www.kaggle.com/c/sberbank-russian-housing-market/data | Данные о недвижимости и макроэкономике России. Целевой признак: стоимость недвижимости |