### Варианты заданий для Лабораторной работы №4

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **№** | **Набор данных (Dataset)** | **Ссылка на датасет** | **Алгоритм дерева (Часть 1)** | **Алгоритм бэггинга** | **Алгоритм бустинга** |
| 1 | Прогнозирование оттока клиентов банка | [Ссылка](https://www.kaggle.com/datasets/shubhammeshram579/bank-customer-churn-prediction) | ID3 | RandomForestClassifier | AdaBoostClassifier |
| 2 | Прогнозирование инсульта | [Ссылка](https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset) | C4.5 | BaggingClassifier | GradientBoostingClassifier |
| 3 | Качество красного вина | [Ссылка](https://www.kaggle.com/datasets/uciml/red-wine-quality-cortez-et-al-2009) | ID3 | RandomForestClassifier | XGBoost |
| 4 | Прогнозирование сердечной недостаточности | [Ссылка](https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction) | C4.5 | ExtraTreesClassifier | CatBoost |
| 5 | Набор данных о курении | [Ссылка](https://www.kaggle.com/datasets/kukuroo3/body-signal-of-smoking) | ID3 | BaggingClassifier | AdaBoostClassifier |
| 6 | Удержание клиентов телеком-оператора | [Ссылка](https://www.kaggle.com/datasets/blastchar/telco-customer-churn) | C4.5 | RandomForestClassifier | GradientBoostingClassifier |
| 7 | Покупка в социальных сетях | [Ссылка](https://www.google.com/search?q=https://www.kaggle.com/datasets/rakeshpanigrahi/social-network-ads) | ID3 | ExtraTreesClassifier | XGBoost |
| 8 | Оценка риска по кредиту | [Ссылка](https://www.kaggle.com/datasets/uciml/german-credit) | C4.5 | BaggingClassifier | CatBoost |
| 9 | Прогнозирование диабета | [Ссылка](https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database) | ID3 | RandomForestClassifier | AdaBoostClassifier |
| 10 | Обнаружение мошенничества с онлайн-платежами | [Ссылка](https://www.kaggle.com/datasets/rupakroy/online-payments-fraud-detection-dataset) | C4.5 | BaggingClassifier | GradientBoostingClassifier |

**Расшифровка алгоритмов для Часть 2:**

* **Бэггинг:**
  + RandomForestClassifier: Случайный лес, классический и мощный алгоритм бэггинга.
  + BaggingClassifier: "Чистый" бэггинг, который можно использовать с любым базовым классификатором (по умолчанию DecisionTreeClassifier).
  + ExtraTreesClassifier: Ансамбль экстремально рандомизированных деревьев.
* **Бустинг:**
  + AdaBoostClassifier: Адаптивный бустинг, один из первых и наиболее известных алгоритмов бустинга.
  + GradientBoostingClassifier: Градиентный бустинг, основа для многих современных мощных алгоритмов.
  + XGBoost: Экстремальный градиентный бустинг, высокопроизводительная реализация.
  + CatBoost: Градиентный бустинг от Яндекса, отлично работающий с категориальными признаками.