

Тема «Иерархические методы»
Самостоятельная работа
Реализация деревьев решений

Выполните указанные ниже задания в виде отдельного файла Jupyter notebook, сохраните его и прикрепите к данному заданию в системе (в имени файла укажите вашу фамилию и инициалы, а также номер модуля и название данной темы).

1. Вы работаете с файлом `dataset_Facebook.csv`, выполните загрузку данных из него. Целевая переменная – `Paid`.
2. Выполните предобработку данных при необходимости.
3. Выполните машинное обучение по модели деревьев решений.
4. Исследуйте влияние глубины дерева на модель, используя 5-fold кросс-валидацию в диапазоне от 1 до 15. Обратите внимание на подбор метрики.

Примеры метрик приведены ниже: Valid options are ['accuracy', 'adjusted_mutual_info_score', 'adjusted_rand_score', 'average_precision', 'completeness_score', 'explained_variance', 'f1', 'f1_macro', 'f1_micro', 'f1_samples', 'f1_weighted', 'fowlkes_mallows_score', 'homogeneity_score', 'mutual_info_score', 'neg_log_loss', 'neg_mean_absolute_error', 'neg_mean_squared_error', 'neg_mean_squared_log_error', 'neg_median_absolute_error', 'normalized_mutual_info_score', 'precision', 'precision_macro', 'precision_micro', 'precision_samples', 'precision_weighted', 'r2', 'recall', 'recall_macro', 'recall_micro', 'recall_samples', 'recall_weighted', 'roc_auc', 'v_measure_score']

5. Постройте графики сравнения результатов на обучающем и тестовом множестве с указанием метрики.
6. Выполните подбор гиперпараметров по методу случайного перебора. В пространство поиска включите несколько критериев.
7. Выведите параметры лучшей модели, наилучшее значение метрики и важность признаков.