

1. Расскажите, как работает регуляризация в решающих деревьях, какие параметры мы штрафует в данных алгоритмах?

Регуляризация помогает решить проблему переобученности модели. Переобучение в большинстве случаев проявляется в том, что в получающихся многочленах слишком большие коэффициенты. Регуляризация в решающих деревьях штрафует признаки за очень малые и большие веса в модели путем добавления нового члена к ошибке. Параметр  $\lambda$  позволяет управлять строгостью штрафа. Смысл регуляризации заключается в минимизации функционала ошибки с ограничением весов. Различают L1-регуляризацию (lasso, регуляризация через манхэттенское расстояние - прибавление к лосс-функции суммы весов в модуле) и L2-регуляризацию (ridge, регуляризация Тихонова - прибавление суммы весов в квадрате с множителем  $\lambda$ ).

2. По какому принципу рассчитывается "важность признака (feature\_importance)" в ансамблях деревьев?

На каждом шаге построения дерева выбирается признак и его значение, которые дают наилучшее разбиение данных, снижают степень неопределённости. В разных алгоритмах применяются разные способы подсчета прироста информации, чтобы вычислить наилучшее разбиение (критерий Джини, энтропия при каждом разбиении). Важность признака в дереве определяется степенью снижения меры неопределенности. Чем выше признак в дереве, тем он важнее. В ансамблях деревьев важность признака считается как среднее арифметическое по всем важностям данного признака для каждого дерева.