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

L1 регуляризация берет веса по модулю и оценивает данные по медиане. L2 регуляризация берет квадраты весов и оценивает данные по средним значениям.

Так же L1 регуляризация помогает избавиться от не нужных признаков.

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

Для задач классификации насчитывается по критерию Gini или по Information Gain.

Для задач регрессии по параметру уменьшения дисперсии.