Задача 1.

Покажем, что стратегия „отвечать средним значением таргета на объектах обучающей выборки, попавших в лист" является лучшей.

Матожидание MSE ошибки:

$$L(\hat{y}) = E[(Y-\hat{Y})^2] = \int_{-\infty}^{+\infty} (y-\hat{y}) f(y) dy =$$

$$= \int_{-\infty}^{+\infty} (y^2 - 2y\hat{y} + \hat{y}^2) f(y) dy =$$

$$= \int_{-\infty}^{+\infty} y^2 f(y) dy - 2\hat{y} \int_{-\infty}^{+\infty} y f(y) dy + \hat{y}^2,$$

т.к. $\int_{-\infty}^{+\infty} f(y) dy = 1$

Минимизируем эту функционалу:

$$\frac{\partial L(\hat{y})}{\partial \hat{y}} = 0 \iff 0 - 2\int_{-\infty}^{+\infty} y f(y) dy + 2\hat{y} = 0 \iff$$

$$\hat{y} = \int_{-\infty}^{+\infty} y f(y) dy = E[Y] \quad \blacksquare$$

Задача 2

$$quality = \frac{|L|}{|Q|} H(L) + \frac{|R|}{|Q|} H(R)$$

При mse: $H(X) = \sum_i (x_i - \bar{x})^2$

Предикат имеет вид feature < threshold

т.е алгоритм строит описываемые прямой $y = const$.

Чтобы добиться эффективности метода нужно вместо порога подбирать коэффициенты a, b нашей линейной модели $y = ax + b$ и в качестве функции ошибки $H(X)$ использовать ошибку линейной регрессии.

Задача 3.

Найдем энтропию многомерного нормального распределения.

$$P(x) = \frac{\exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)}{\sqrt{(2\pi)^n |\Sigma|}} -$$

- плотность распред-я, где

μ - вектор матожиданий,

Σ - матрица ковариаций

n - размерность вектор

$$H(X) = -\int_{-\infty}^{+\infty} P(x) \ln P(x)\, dx =$$

$$= -\int P(x) \left[-\frac{1}{2} \ln\left[(2\pi)^n |\Sigma|\right]\right] dx -$$

$$- \int P(x) \left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right] dx =$$

$$= \frac{1}{2} \ln\left[(2\pi)^n |\Sigma|\right] + \frac{1}{2} \int P(x) \left[(x-\mu)^T \Sigma^{-1} (x-\mu)\right] d$$

trace - след матрицы

$$\mathrm{trace}(AB) = \mathrm{trace}(BA)$$

$$\cancel{(x-\mu)^T \Sigma^{-1} (x-\mu) = \mathrm{trace}(x^T \Sigma^{-1} x) =}$$

$$\cancel{= \mathrm{trace}(\Sigma^{-1} x}$$

$$(x-\mu)^T \Sigma^{-1} (x-\mu) = \mathrm{trace}\left((x-\mu)^T \Sigma^{-1} (x-\mu) \right) =$$

$$= \mathrm{trace}\left(\Sigma^{-1} (x-\mu)^T (x-\mu) \right) = \mathrm{trace}(E) = n$$

Таким образом, продолжая вычисления,

$$H(x) = \frac{1}{2} \ln\left[(2\pi)^n |\Sigma| \right] + \frac{n}{2} =$$

$$= \frac{1}{2} \left[\ln\left[(2\pi e)^n |\Sigma| \right] \right]$$

∎