Problem 3.

linear reg.

coef stands for θ_i,

p-value should be small (near 0)

⇒ reject $H_0: \theta_i = 0$ (para. does not matter)

80% traing & 20% testing (test-size =0.2)

Q-Q plot should be close to 45° line,

R^2 should be close to 1

we can find out that some parameters (features) are not related to prize.

then modify the code based on canvas file.

GBM,

still needs to calculate R^2.

n-estimator : trees / iteration steps.

max-depth : 4 / 16 leaves

learning rate : 0.01

alpha : 1 = λ in pdf.

honestly, I don't write the python code, the above is a draft.