$$\text{RSS} = \sum_{i=1}^{n} \left[y_i - (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \right]^2$$

$$\text{minimize} \quad \underbrace{\sum_{i=1}^{n} \left(y_i - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p) \right)^2}_{\text{residual sum of squares}} + \underbrace{\sum_{j=1}^{p} |\beta_j|}_{\text{penalty}}$$

$$\text{minimize} \quad \underbrace{\sum_{i=1}^{n} \left(y_i - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p) \right)}_{\text{residual sum of squares}} \quad \text{where} \quad \underbrace{\sum_{j=1}^{p} |\beta_j| \leq t}_{\text{penalty}}$$

$$\text{minimize} \quad \underbrace{\sum_{i=1}^{n} \left[y_i - (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \right]^2}_{\text{residual sum of squares}} + \lambda \underbrace{\sum_{j=1}^{p} P(\beta_j)}_{\text{penalty}}$$

$$\underbrace{E(y_0 - \hat{y}_0)}_{\text{expected residual}} = \underbrace{\text{Var}(\hat{y}_0)}_{\text{variance of model}} + \underbrace{(\text{Bias}(\hat{y}_0))^2}_{\text{squared bias of model}} + \underbrace{\text{Var}(\epsilon)}_{\text{variance of error}}$$

$$\hat{\beta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon$$

$$\hat{f}(X_1, X_2, \ldots, X_p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$$

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon$$

1

Let n be an integer. Prove that $4 \nmid n^2 - 3$.

We will consider three cases. First, assume that n is even. Then $n = 2k$ for some integer k. Hence,

$$n^2 - 3 = (2k)^2 - 3 \tag{1}$$
$$= 4k^2 - 3 \tag{2}$$

which is not a multiple of 4.

Now, assume that n is odd. Then $n = 2k + 1$ for some integer k. Then

$$n^2 - 3 = (4k + 1)^2 - 3 \tag{3}$$
$$= (16k^2 + 8k + 1) - 3 \tag{4}$$
$$= 16k^2 + 8k - 2 \tag{5}$$
$$= 4(4k^2 + 2k) - 2 \tag{6}$$

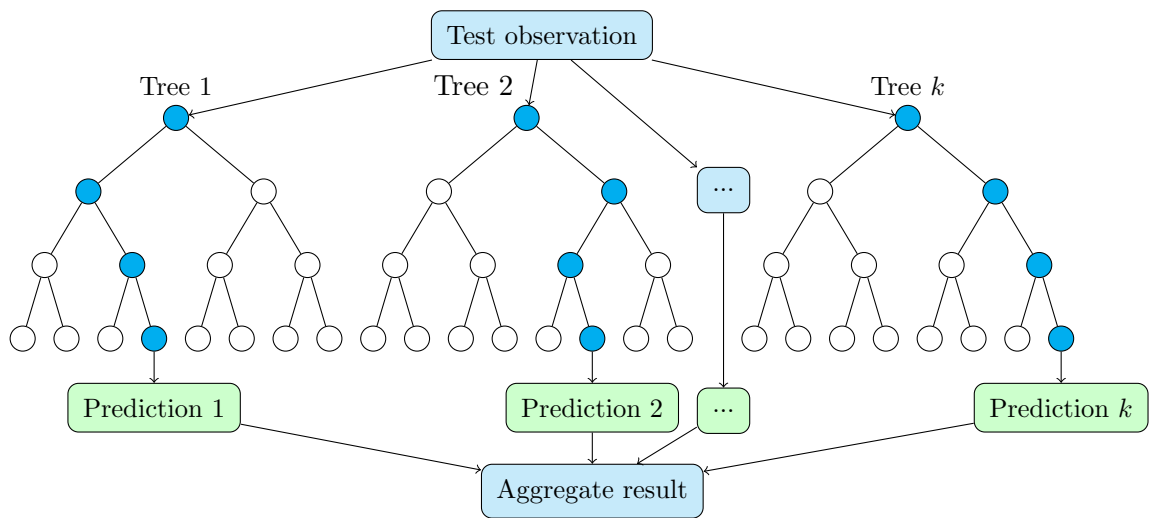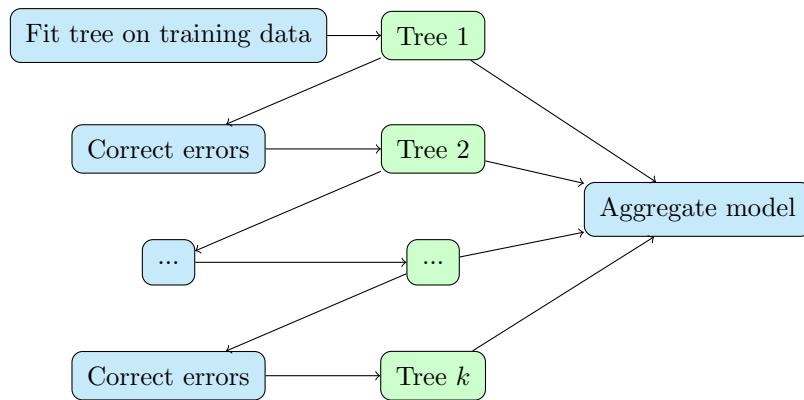which is also not a multiple of 4. Thus, for all integers n, $n^2 - 3$ is not a multiple of four. Therefore, $4 \nmid n^2 - 3$ for all integers n.

Figure 1: Random forest

Figure 2: Gradient boosting