$$\therefore \quad N_{R_{m'}} \cdot w^*_{m'} - \sum_{x_i \in R_{m'}} y_i = 0$$

$$\boxed{\therefore \quad w^*_{m'} = \sum_{x_i \in R_{m'}} \frac{y_i}{N_{R_{m'}}}}$$

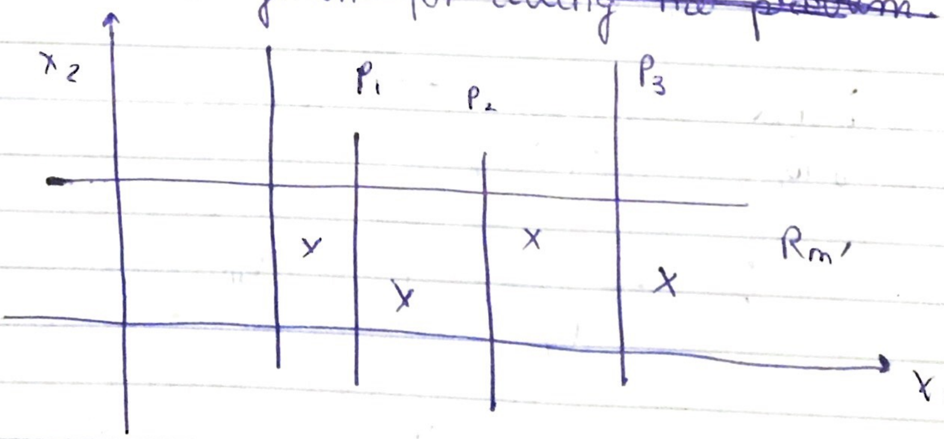
(b) For a given region $R_{m'}$, containing $N_{R_{m'}}$ data points, and a given feature threshold x_j, if we want to find an optimal value threshold t_k by trying different values, then we need to try

atmost $N_{R_{m'}} - 1$ values of t_k.

This is because, given a feature to threshold in $R_{m'}$, we can atmost get $N_{R_{m'}}$ different outputs splitting a region if the points are not aligned.

Q3.

~~(a) I used Python for coding the problem.~~



($N_{R_{m'}} - 1$) possible decision boundaries within $R_{m'}$.

PROBLEM ON READING

Q1. (a) It is a case of semi supervised learning problem. Since we use only given training sample, its transductive.

(b) As we are out of original training data, and use new data, it is inductive.

(c) As all the data available is labeled, it is a type of supervised learning.

(d) It is a case of semi supervised learning problem. It's of inductive type

---

Q2. CART applied to regression.

(a) $$\text{Cost}\,\{(\underline{x}_i, y_i) \in R_{m'}\} = \sum_{\underline{x}_i \in R_{m'}} (y_i - w_{m'})^2$$

When we differentiate wrt $w_{m'}$,

$$\frac{\partial\, \text{Cost}}{\partial\, w_{m'}} = \sum_{\underline{x}_i \in R_{m'}} -2\,(y_i - w_{m'}) = 0$$

$$\therefore \sum_{\underline{x}_i \in R_{m'}} -2y_i + 2w_{m'} = 0$$

$$\therefore \sum_{\underline{x}_i \in R_{m'}} w_{m'} - y_i = 0$$