(b.) Define the sets $A, B \subset \mathbb{R}$ by

$$A = [0,1] \cup [4,5] \cup [8,9] \cup \dots \cup [4(m-1), 4(m-1)+1]$$

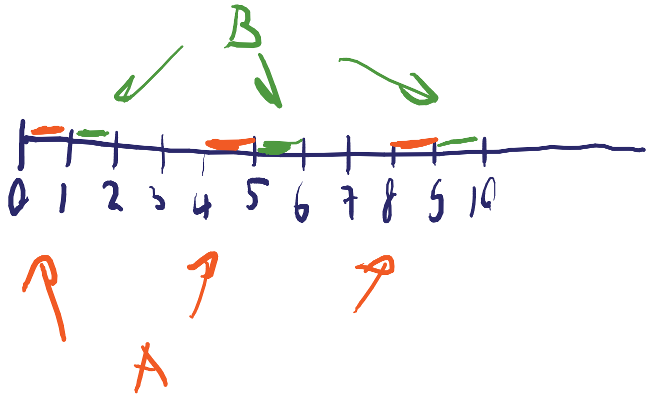
$$B = [1,2] \cup [5,6] \cup \dots \cup [4(m-1)+1, 4(m-1)+2]$$

Suppose that

$$P(Y=0) = P(Y=1) = \frac{1}{2}$$

and, given $Y=0$, X is uniformly distributed on A

given $Y=1$, X is uniform on B.



Clearly, $R^* = 0$.

The key observation is that if for each $j = 0, 1, \dots, m-1$, there is at most one data point X_i in $[j, j+2]$, then the risk of the nearest neighbor rule equals

$R(g_n) = \mathbb{P}(g_n(X) \neq Y \mid D_n) = \frac{1}{2}$ and therefore

$$E\,R(g_n) \geq \frac{1}{2} \cdot \mathbb{P}(\text{at most one } X_i \text{ in every "bin" } [j, j+2])$$

$$= \frac{1}{2}\left[1 - \mathbb{P}(\exists j \in \{0, \dots, m-1\} \text{ with at least two } X_i\text{'s in } [j, j+2])\right]$$

$$\geq \frac{1}{2}\left[1 - m\,\underbrace{\mathbb{P}(\text{there are at least } 2\ X_i\text{'s in } [0,2])}\right]$$

union bound

$$= \mathbb{P}\left(\text{Binom}\left(n, \tfrac{1}{m}\right) \geq 2\right)$$

$$= \sum_{k=2}^{n} \binom{n}{k} \left(\frac{1}{m}\right)^k \left(1 - \frac{1}{m}\right)^{n-k} \leq \frac{n \binom{n}{2}}{m^2} \leq \frac{n^3}{m^2}$$

$$\geq \frac{1}{2}\left[1 - \frac{n^3}{m}\right] > \frac{1}{4} \text{ if } m > 2n^3$$