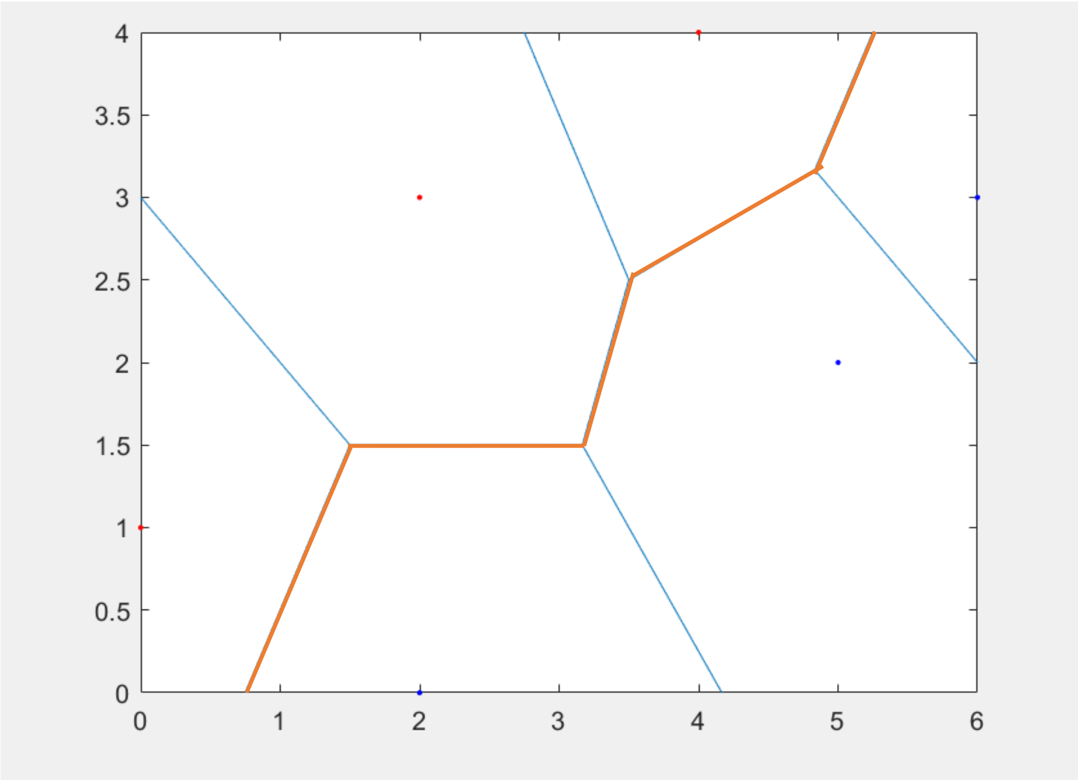
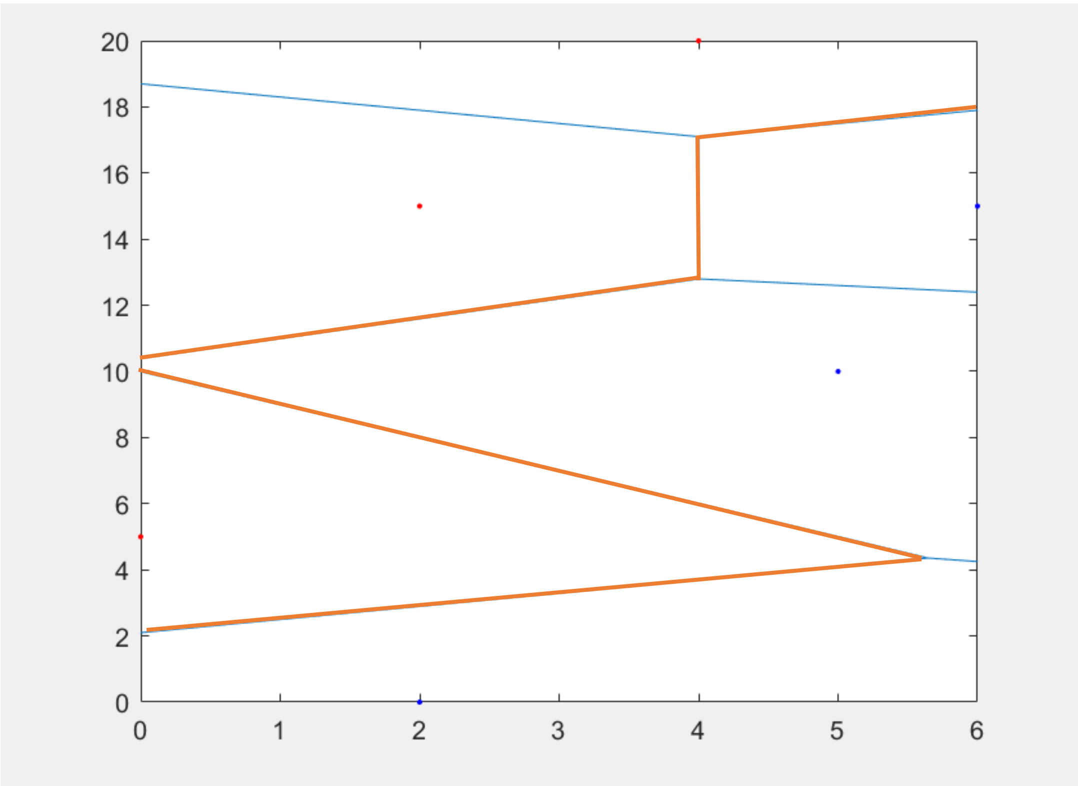
2.

(a).The orange line ids the decision boundary.



(b).

The orang line is the decision boundary.



Since the y-coordinate of each point was multiplied by 5, the importance of y-coordinate has been addressed.

(c). It will be classified as red.
At least two points should be added. And the coordinates of these two points should meet $(x-1)^2 + (y-2)^2 = 2$.

(d). The time complexity of knn is O(nd).
Using data condense, the worst case is O(nd).
Using the KD-tree, the worst case is O(nd), and the average case is O(logn).

1. (a) $G_{ij} = \vec{x}_i^T \vec{z}_j$

$X = [\vec{x}_1, \cdots, \vec{x}_n] \in \mathbb{R}^{d\times n}$ $Z = [\vec{z}_1, \cdots, \vec{z}_n] \in \mathbb{R}^{d\times m}$

$G = [G_{11}, \cdots, G_{nm}]$

Thus, $G = X^T Z$

(b) Since $D_{ij}^2 = (\vec{x}_i - \vec{z}_j)^2$, $S_{ij} = \vec{x}_i^T \vec{x}_i$, $R_{ij} = \vec{z}_j^T \vec{z}_j$

$D_{ij}^2 = S_{ij} - 2G_{ij} + R_{ij}$

$S = \begin{pmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_n \end{pmatrix} \cdot (x_1, \cdots, x_n)$ $R = \begin{pmatrix} \hat{z}_1 \\ \vdots \\ \hat{z}_m \end{pmatrix} \cdot (z_1, \cdots, z_m)$

$\therefore D^2 = S - 2G + R$

① $D_{ij}^2 = (\vec{x}_i - \vec{z}_j)^2 \geq 0.$

② It should be D

③ $D_{ij} = \sqrt{D_{ij}^2} = \sqrt{(\vec{x}_i - \vec{z}_j)^2}$

$D = \sqrt{S - 2G + R}$

So, it need to compute the root square.