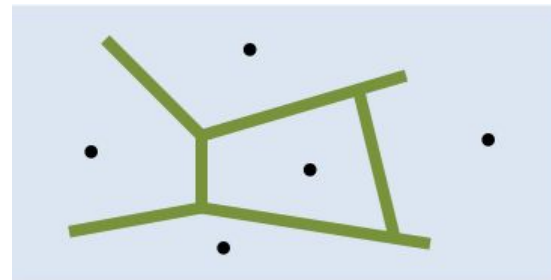
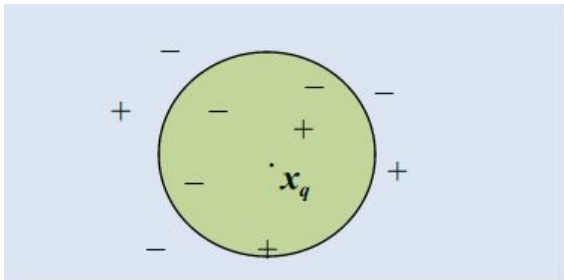


## Algorithme des $k$ plus proches voisins

- Chaque objet est représenté par un point dans un espace à  $n$  dimensions
- Les plus proches voisins sont définis en terme de distance (euclidienne, manhattan, ...) ou dissimilarité
- La fonction de prédiction peut être discrète/nominale ou continue
  - . discrète : valeur la plus fréquente
  - . continue : moyenne
- Diagramme de Voronoï : surface de décision induite par le plus proche voisin



- Pondération en fonction de la distance
  - . pondère la contribution de chacun des  $k$  voisins en fonction de sa distance à l'objet à classer
  - . plus on est proche, plus on a de poids
- Robuste dans le cas de données bruitées
- Problème de dimensionnalité : la distance peut être dominée par des attributs non pertinents
  - . solution : normalisation des dimensions ou élimination des attributs non pertinents