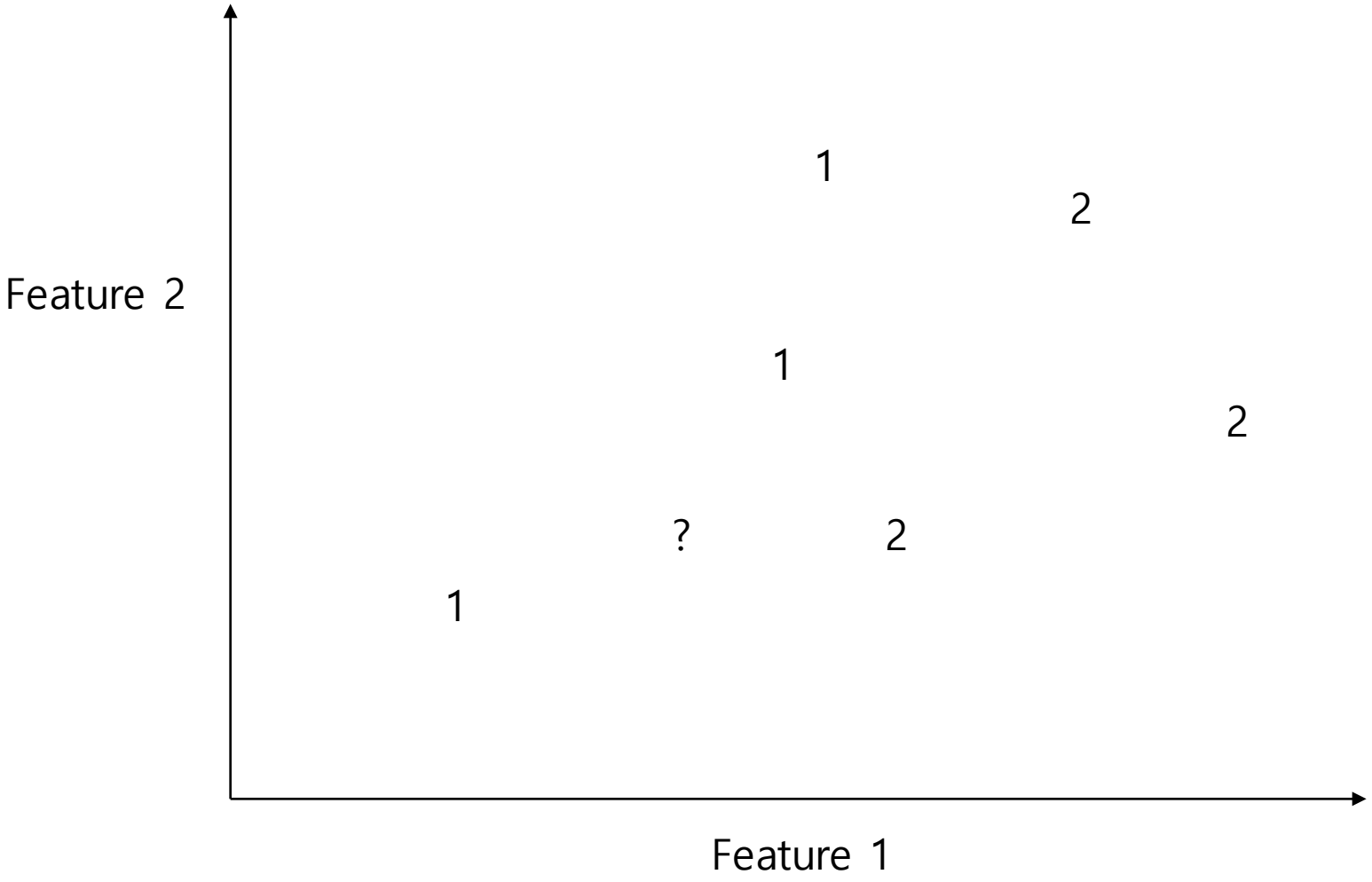
kNN Decision Boundary at k=1 and Voronoi Tessellation

Sept. 2015
Youn-Hee Han
http://link.koreatech.ac.kr

kNN Algorithm at k=1

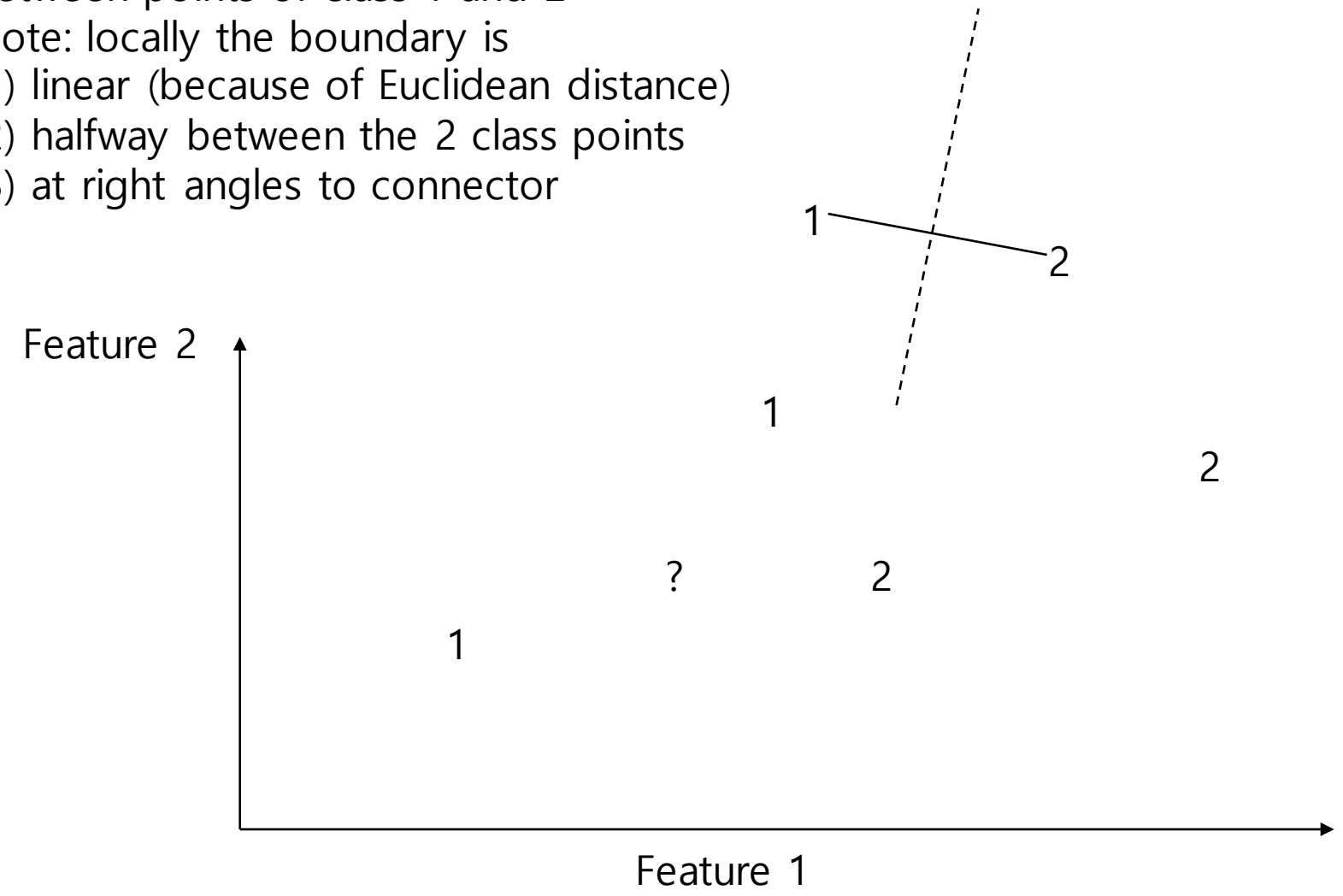
- Algorithm
  - y is a new feature vector whose class label is unknown
  - Search the training set for the closest feature vector to y
    - let this "closest feature vector" be x(j)
    - How are "closest x" vectors determined?
      - typically use minimum Euclidean distance
  - Classify y with the same label as x(j),
    - i.e. y is assigned label c(j)
  - Side note: this produces a "Voronoi tesselation" of the d-space
    - each point "claims" a cell surrounding it
    - cell boundaries are polygons

How should the new point be classified?

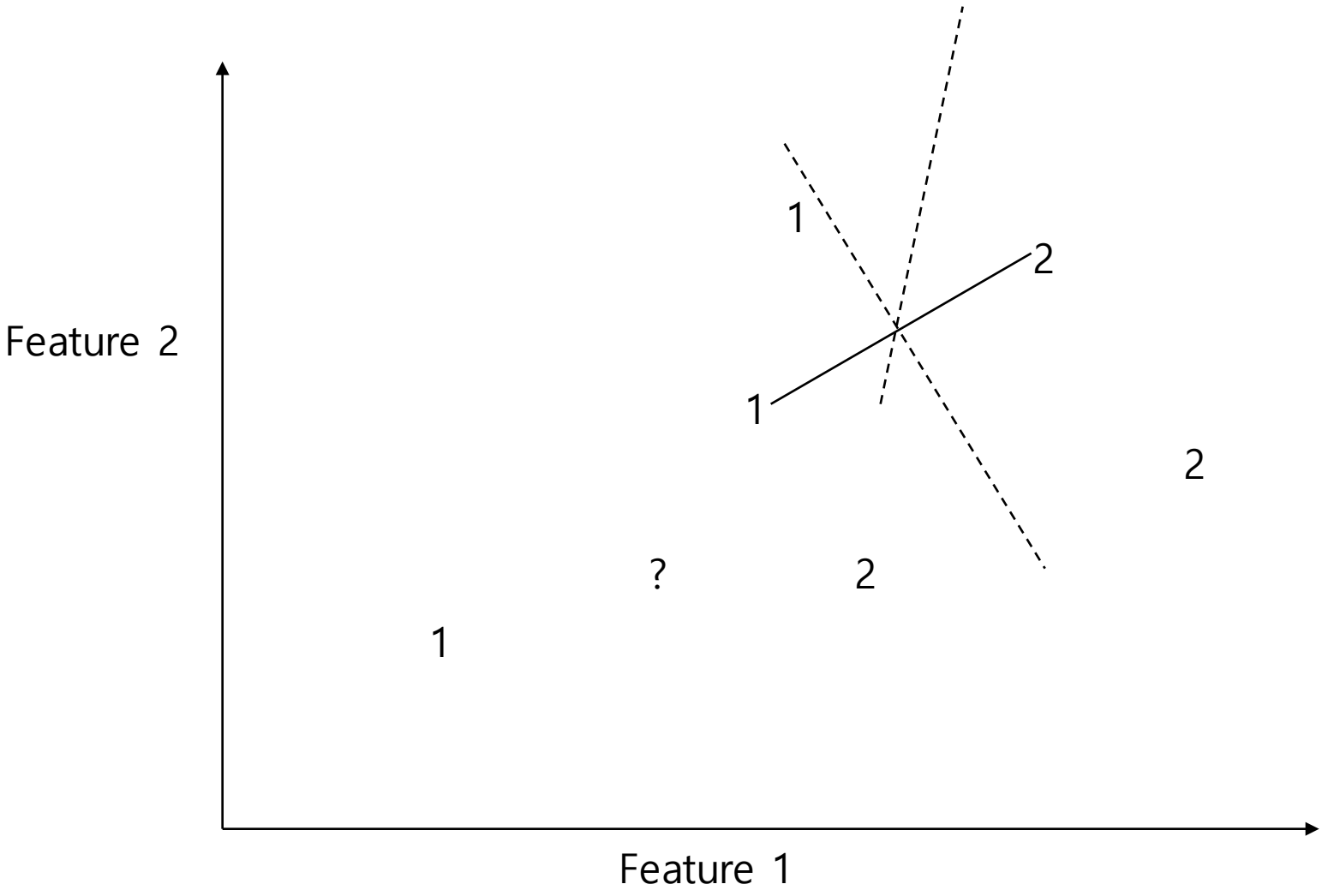


Local Decision Boundaries

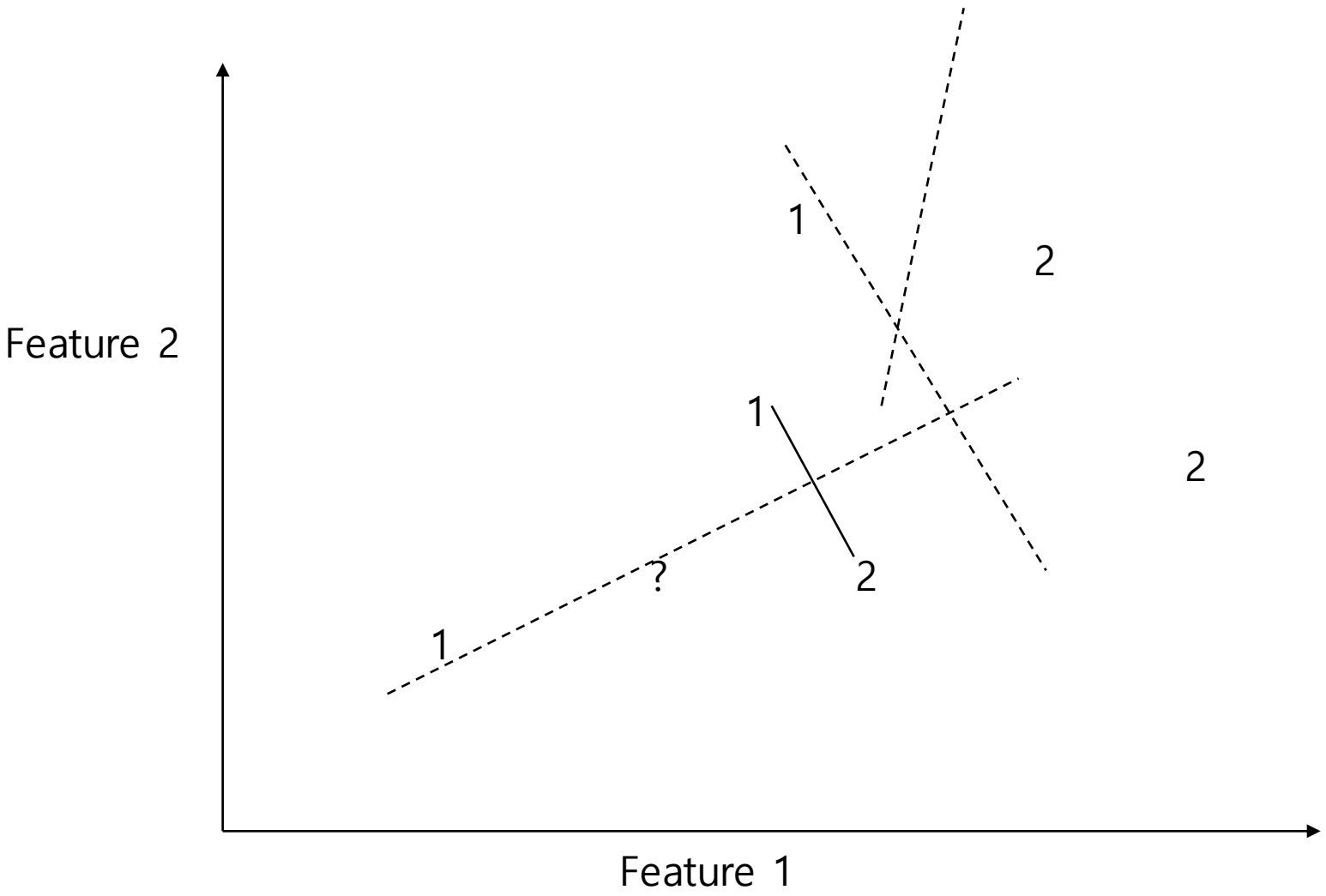
Boundary? Points that are equidistant between points of class 1 and 2
Note: locally the boundary is
(1) linear (because of Euclidean distance)
(2) halfway between the 2 class points
(3) at right angles to connector



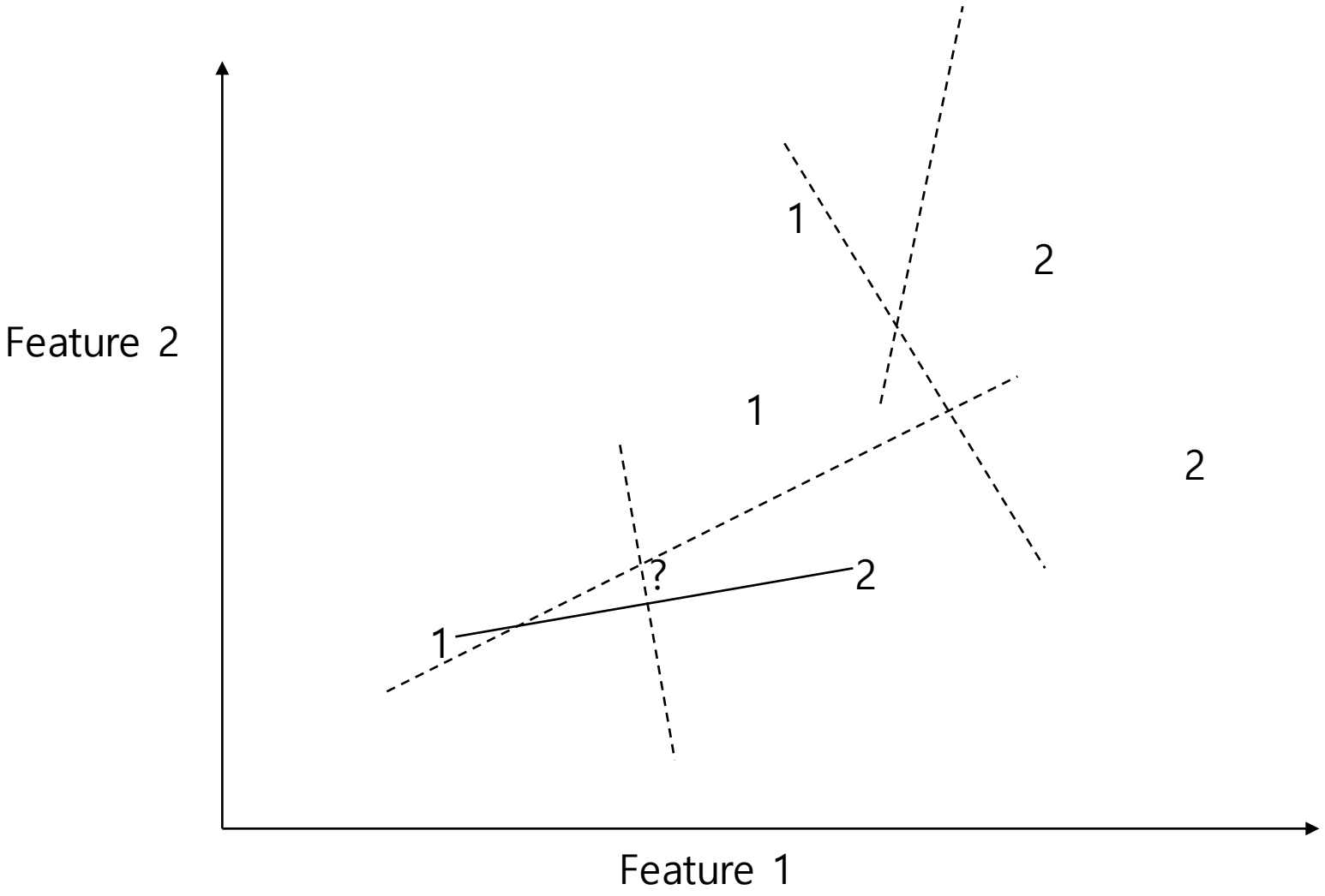
Finding the Decision Boundaries



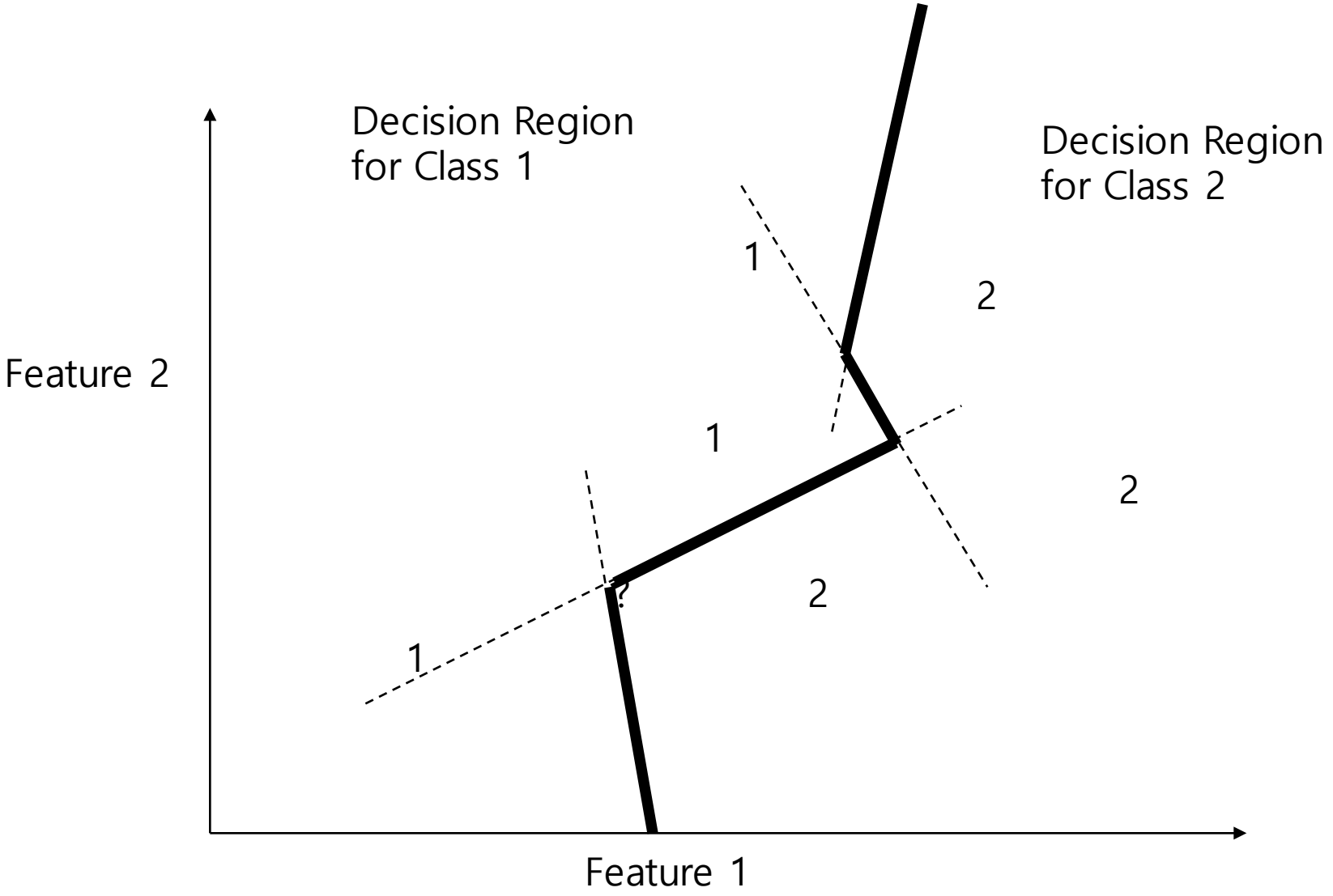
Finding the Decision Boundaries



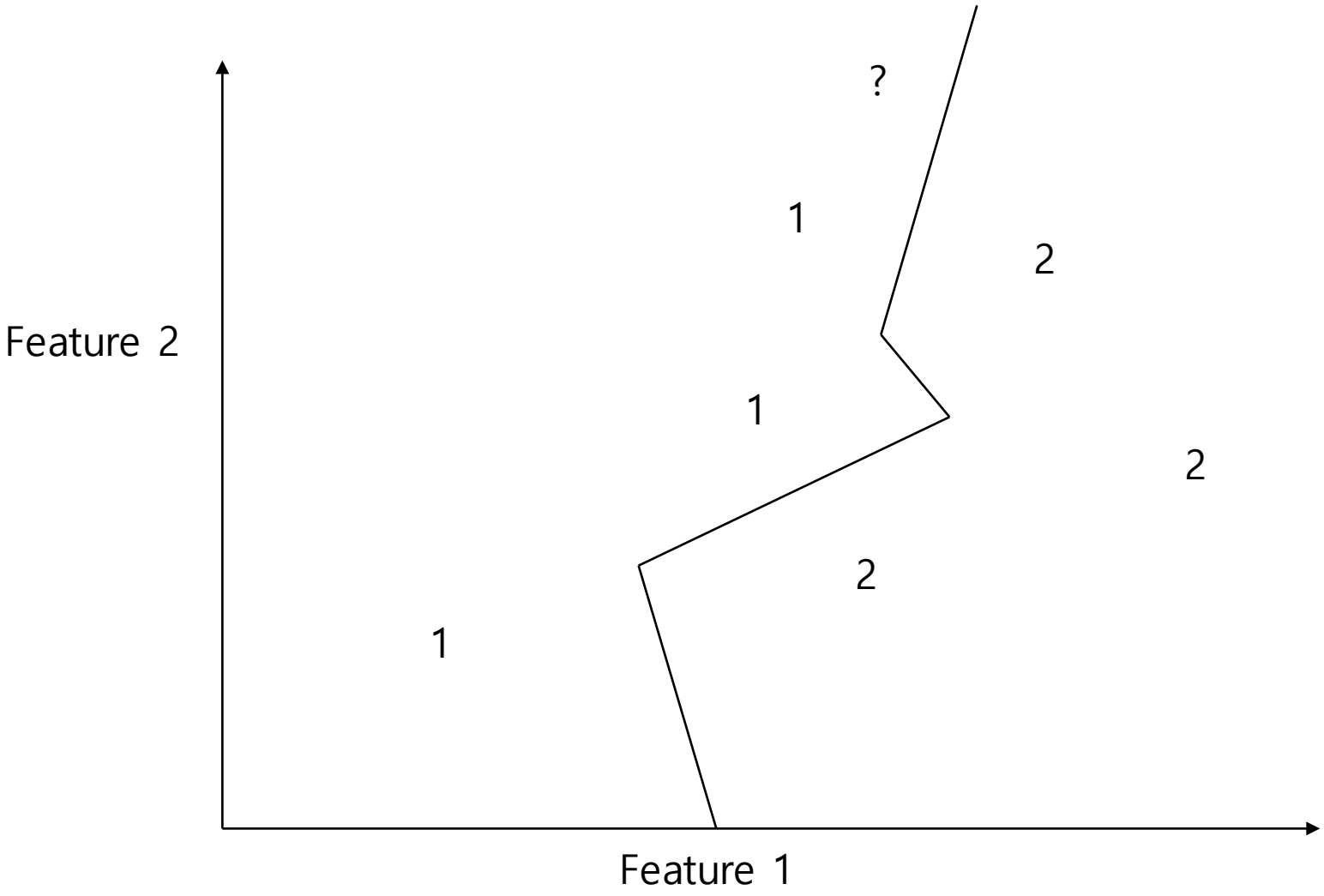
Finding the Decision Boundaries



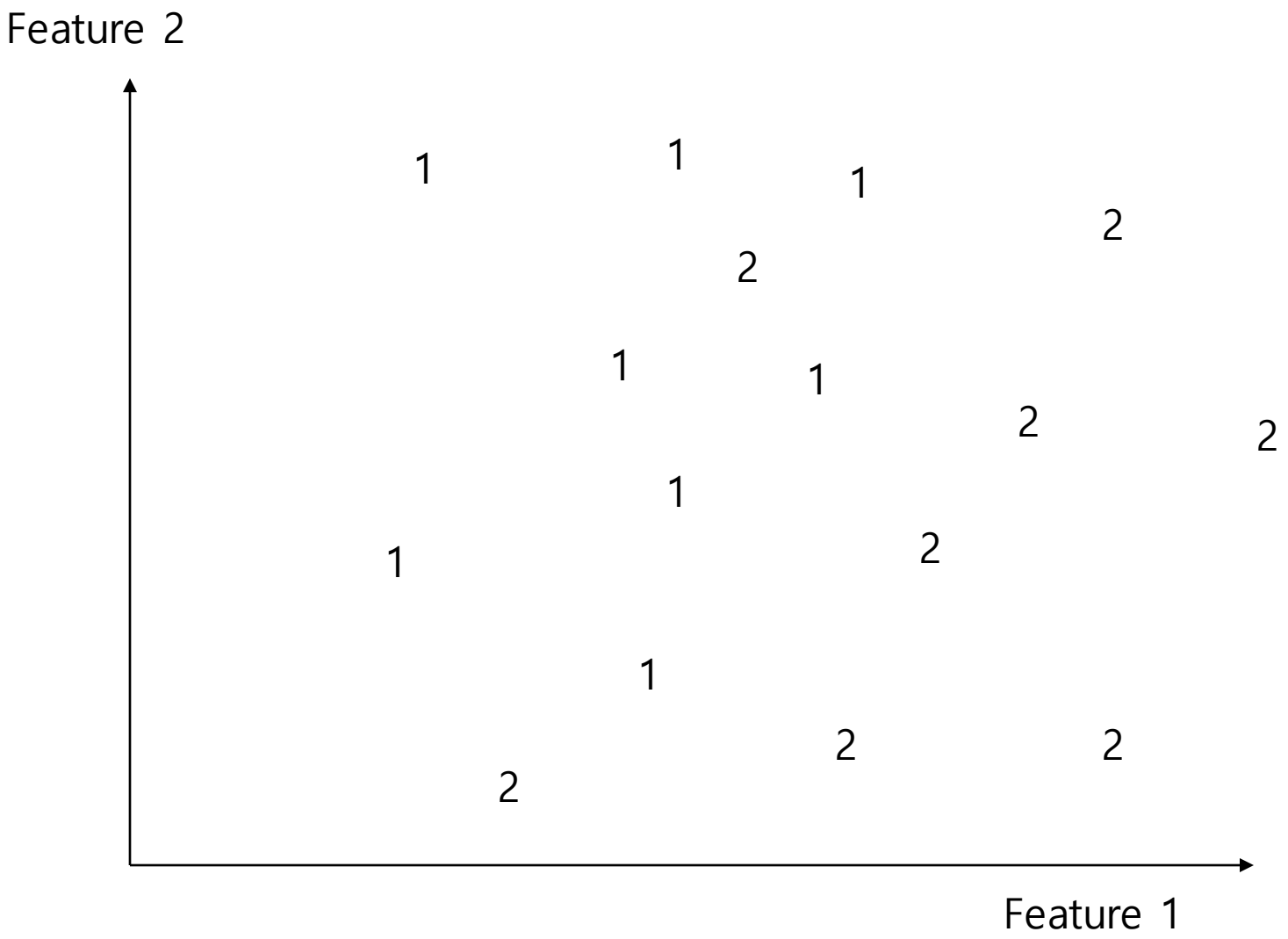
Overall Boundary = Piecewise Linear



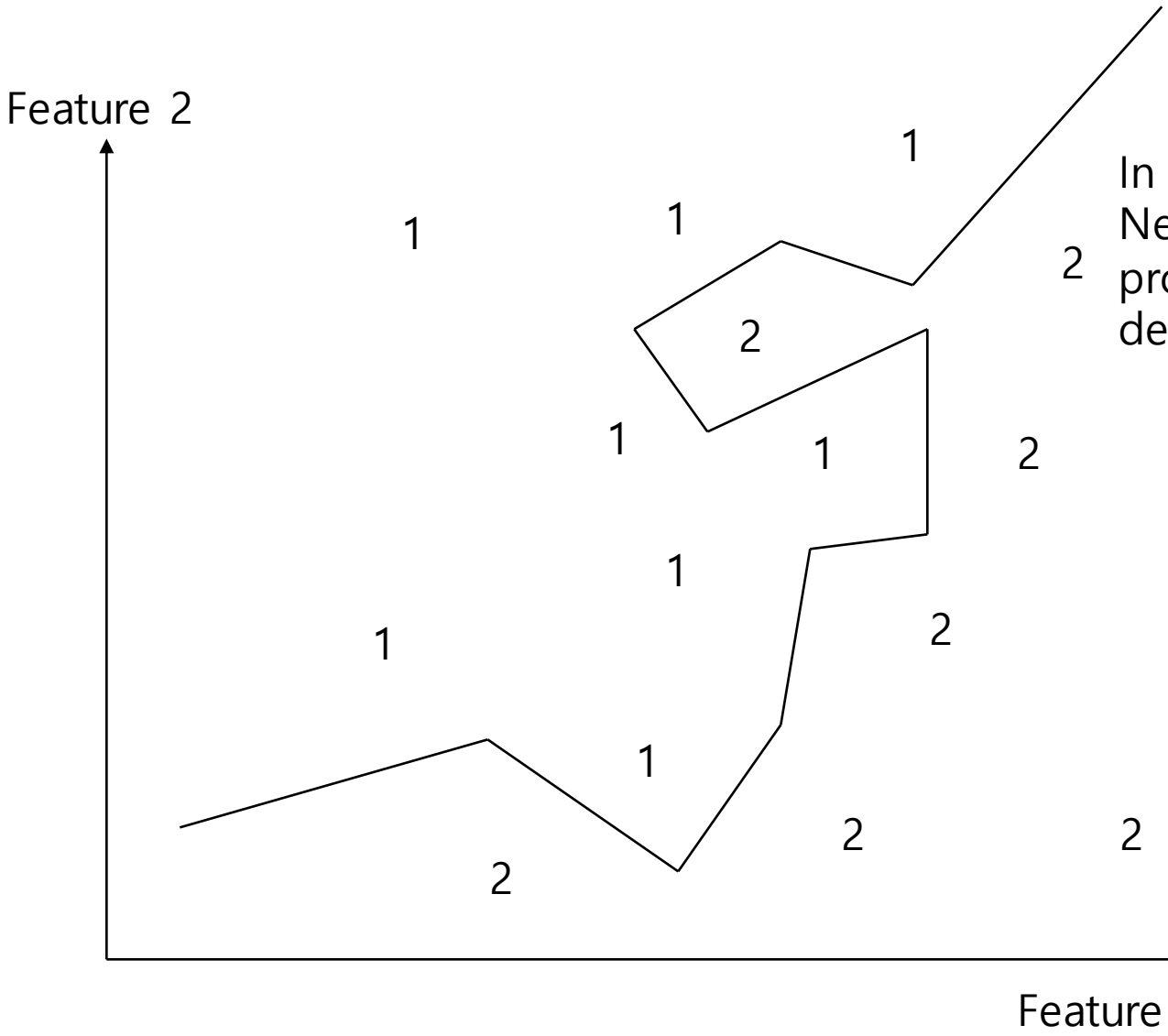
Geometric Interpretation of kNN (k=1)



More Data Points



More Complex Decision Boundary



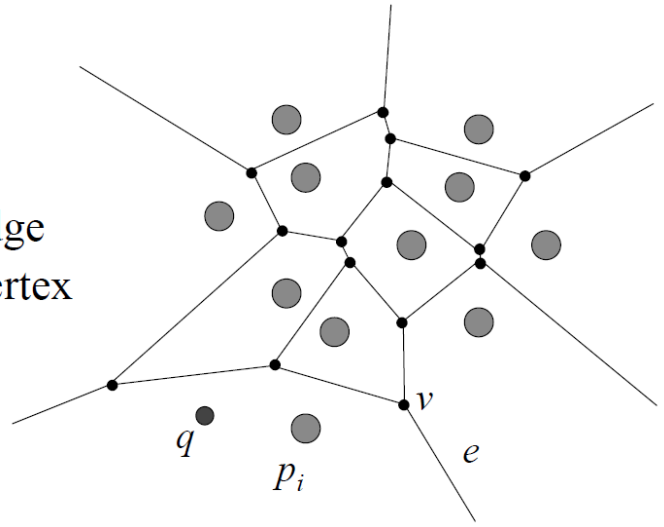
In general:
Nearest-neighbor classifier produces piecewise linear decision boundaries

Voronoi diagram (tesellation)

◈Voronoi diagram

- The Voronoi diagram is formed from lines that bisect and are perpendicular to the lines that connect two neighboring vertices.
- Each point *s* has a Voronoi cell *V(s)* consisting of all points closer to *s* than to any other points

p_i : site points
q : free point
e : Voronoi edge
v : Voronoi vertex

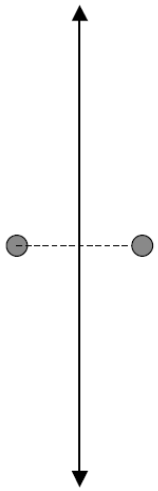


A point q lies in the cell corresponding to a site $p_i \in P$
iff
Euclidean_Distance(q, p_i) < Euclidean_distance(q, p_j),
for each $p_i \in P, j \neq i$.

Voronoi diagram

- Voronoi diagram examples
  - 1 point

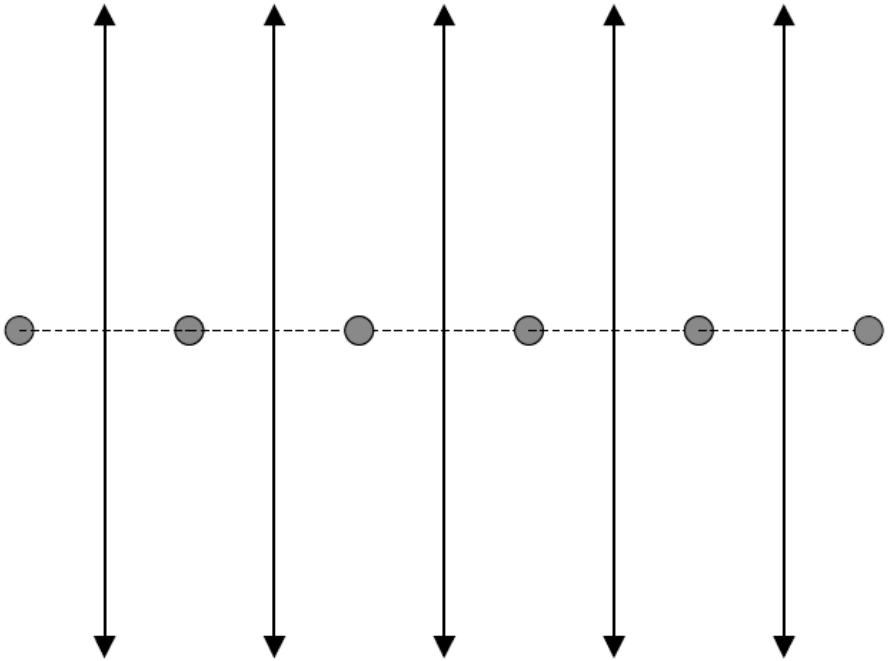
  - 2 points form "a perpendicular bisector"



Voronoi Diagram is a line that extends infinitely in both directions, and the two half planes on either side.

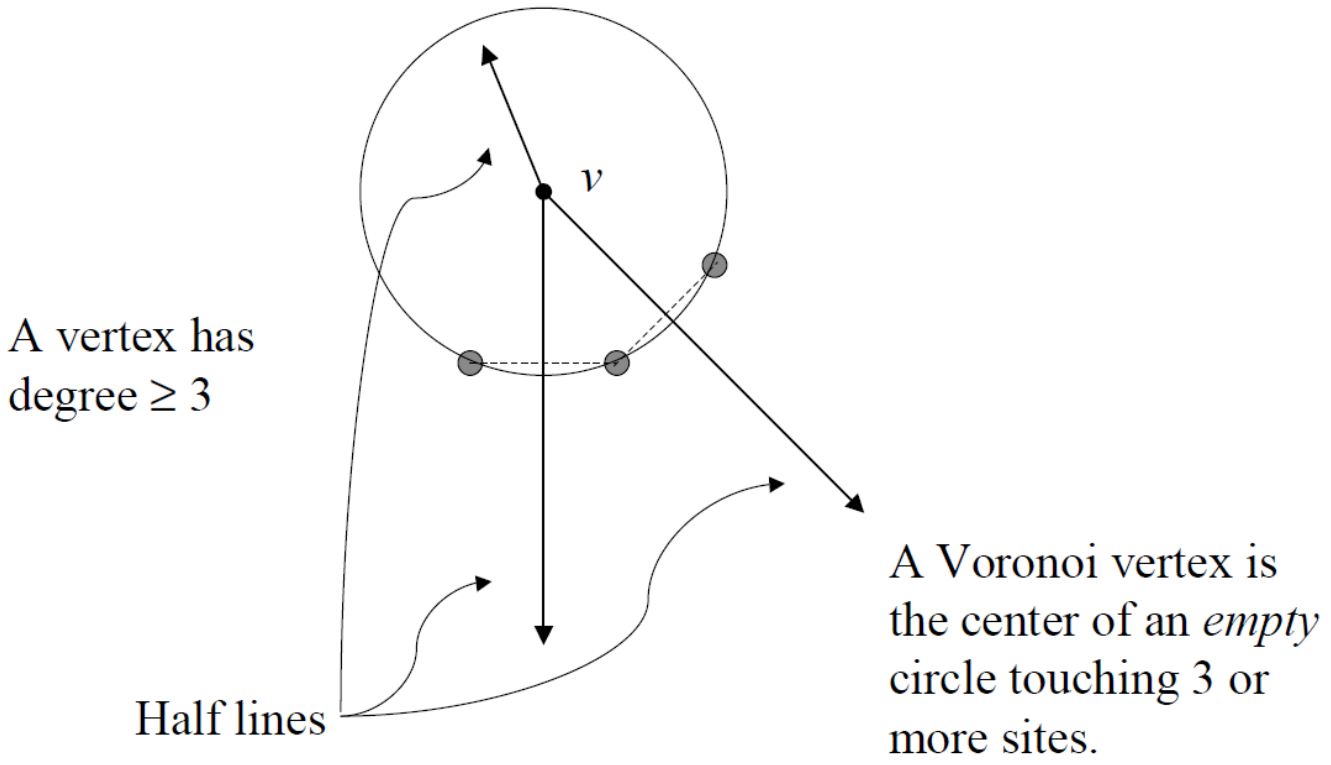
Voronoi diagram

- Voronoi diagram examples
  - Collinear points form “a series of parallel lines”



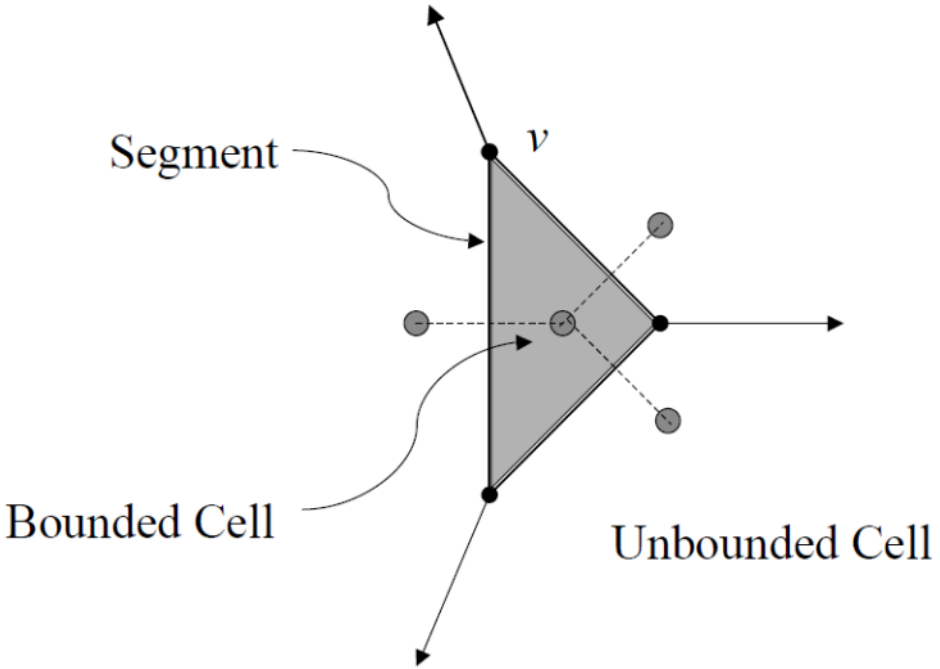
Voronoi diagram

- Voronoi diagram examples
  - Non-collinear points form "Voronoi half lines" that meet at a vertex



Voronoi diagram

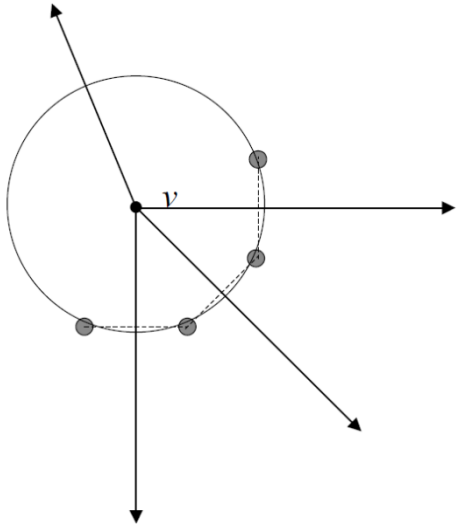
◈Voronoi cells and segments



Voronoi diagram

- Which of the following is true for 2-D Voronoi diagrams?
  - Four or more non-collinear points are…
    1) sufficient to create a bounded cell
    2) necessary to create a bounded cell
    3) 1 and 2
    4) none of above

- Four points' degeneration case of bounded cell:

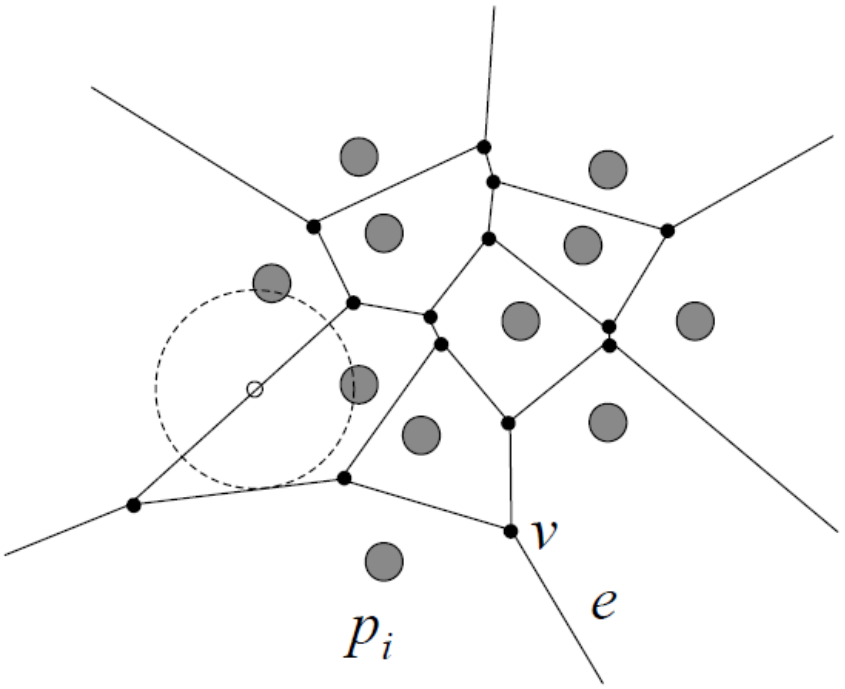


Property I of Voronoi diagram

A point q lies on a Voronoi edge between sites p_i and p_j *iff* the largest empty circle centered at q touches only p_i and p_j

- A Voronoi edge is a subset of locus of points equidistant from p_i and p_j

p_i : site points
e : Voronoi edge
v : Voronoi vertex

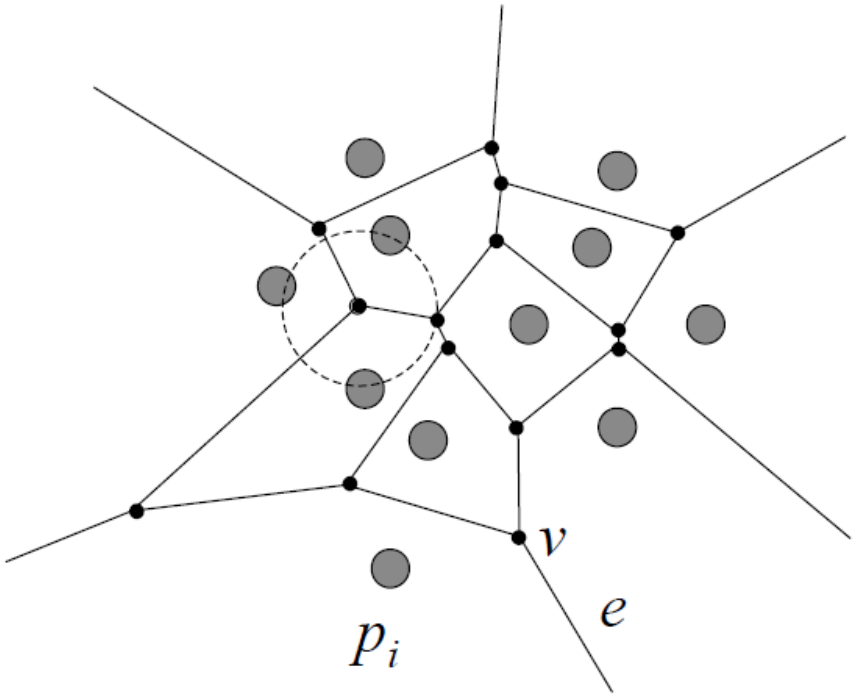


Property II of Voronoi diagram

A point q is a vertex *iff* the largest empty circle centered at q touches at least 3 sites

- A Voronoi vertex is an intersection of 3 more segments, each equidistant from a pair of sites

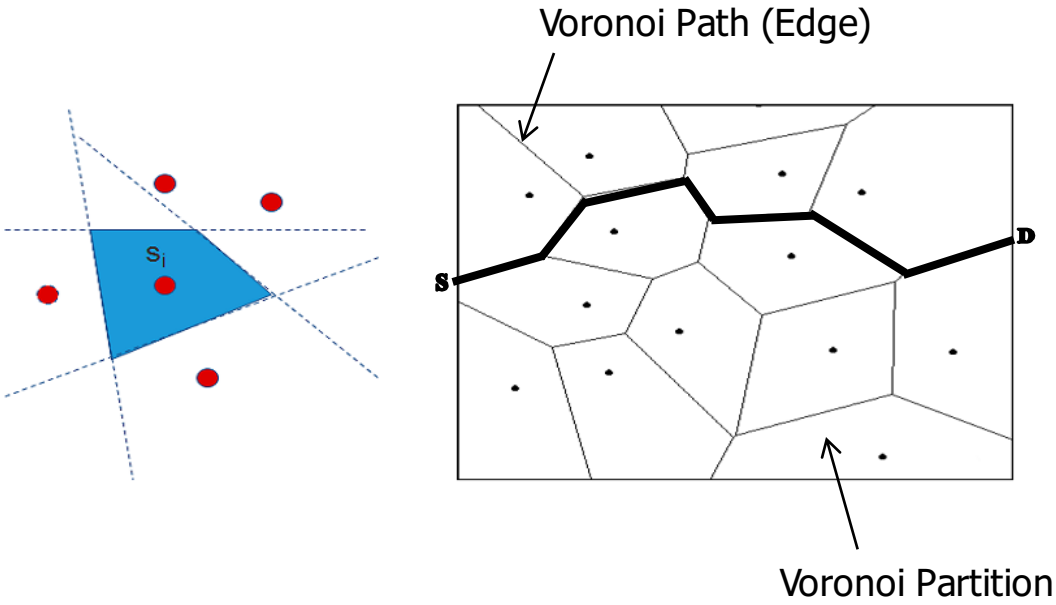
p_i : site points
e : Voronoi edge
v : Voronoi vertex



Surveillance Path

◈Maximal Breach Path

Voronoi diagram



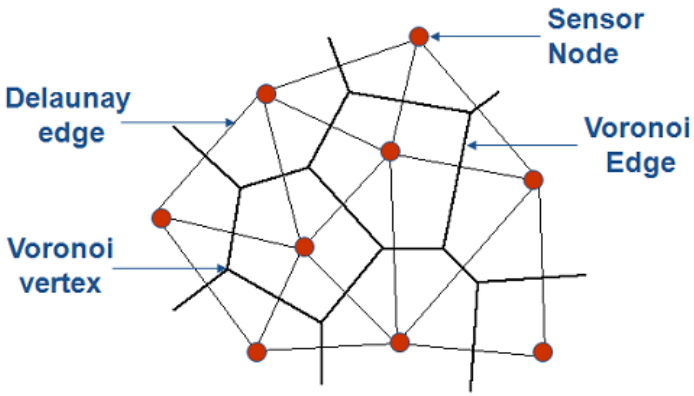
– Voronoi Path (= Maximal Breach Path)

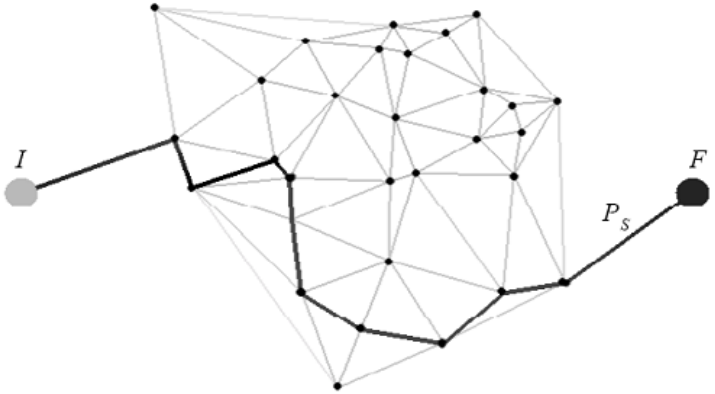
- The path where the surveillance level is the lowest
- The path where its closest distance to any sensor is as large as possible.

Surveillance Path

◈Maximal Support Path

Delaunay triangulation





– Delaunay Triangulation Path (= Maximal Support Path)

- The path where the surveillance level is the greatest
- The path where its closest distance to any sensor is as short as possible.