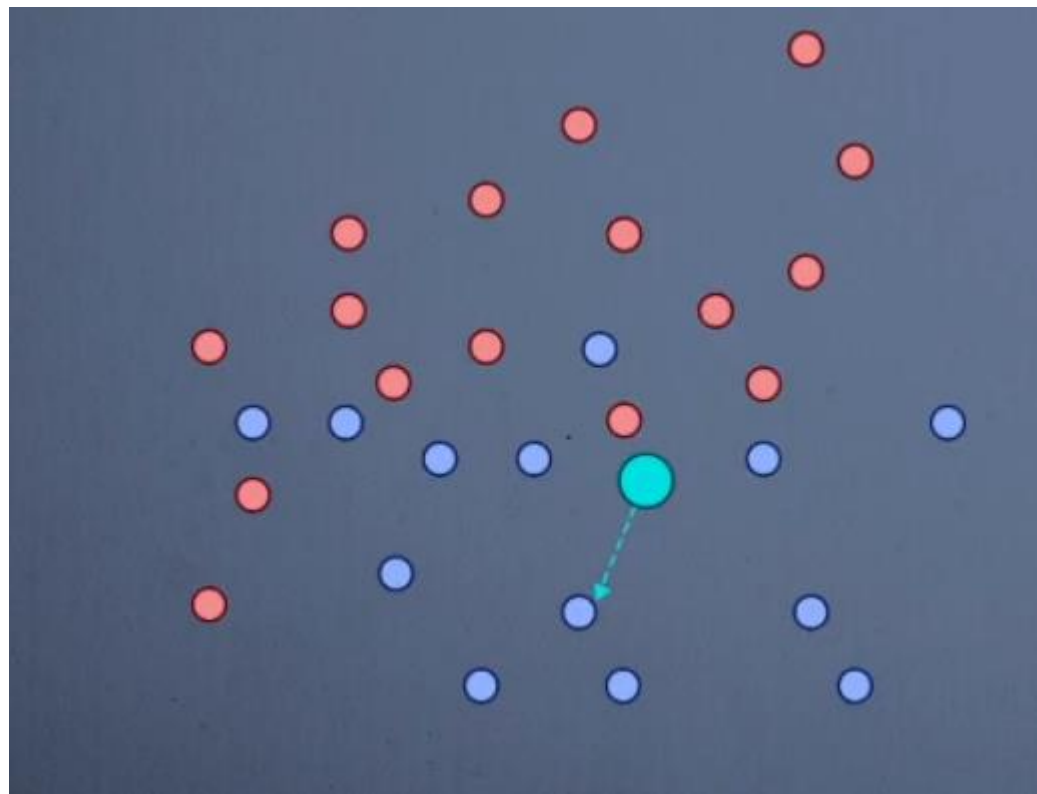


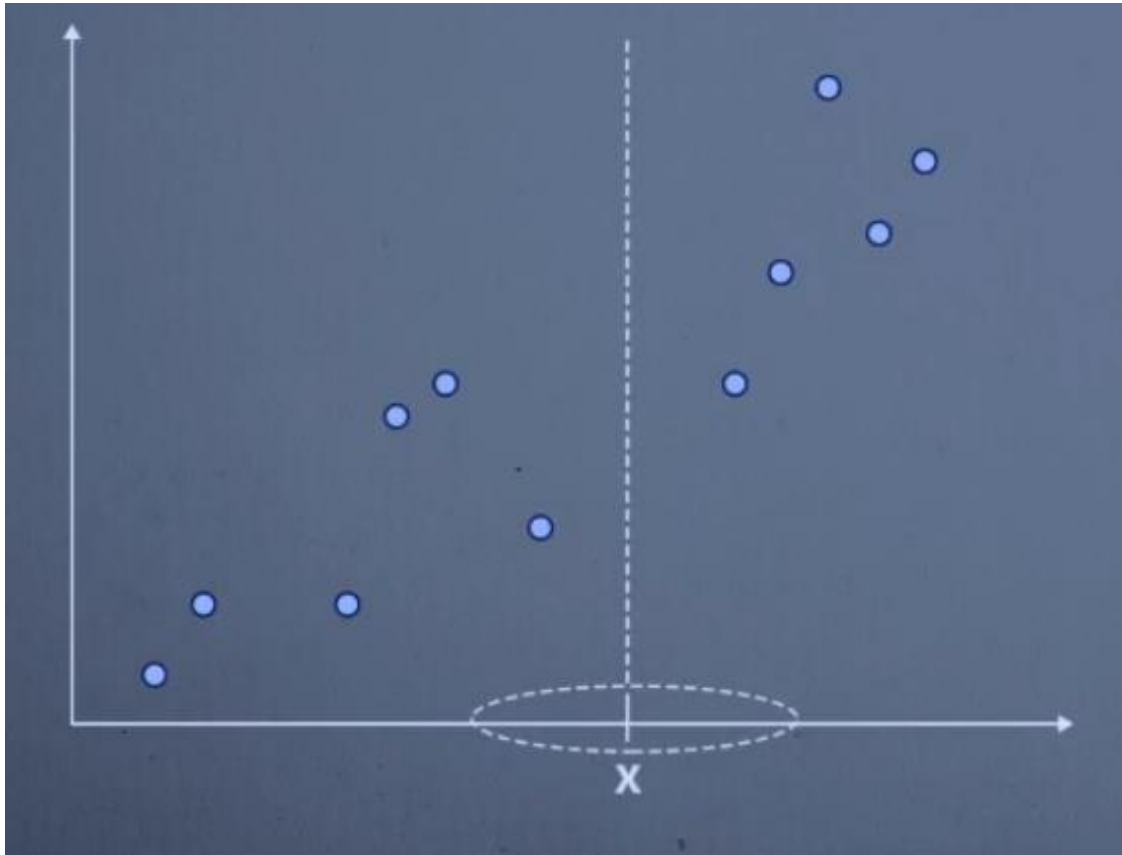
K-Nearest Neighbors(KNN)

KNN - Classification

- 제일 가까운 것을 찾는다.
- $K=1, 3, 5$.. 몇 개를 볼 것인가?
- 점점 경계가 단순해짐



KNN - Regression



입력(x)에 대해 가까운 값을 통해 y값을 예측
K=1, 3, 5
점점 경계가 단순해짐

KNN - 변형

Variation

» Classification (k=5)

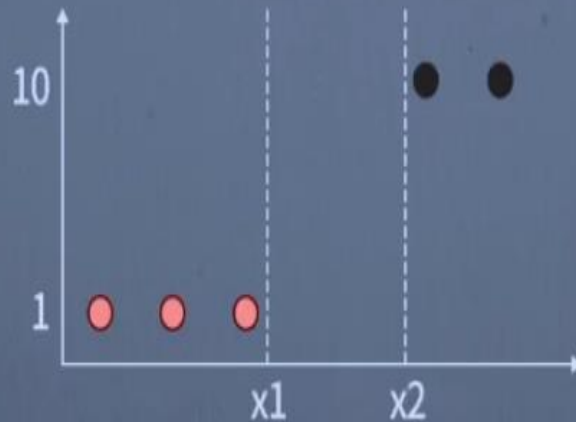
- X1: Red or Black?
- X2: Red or Black?



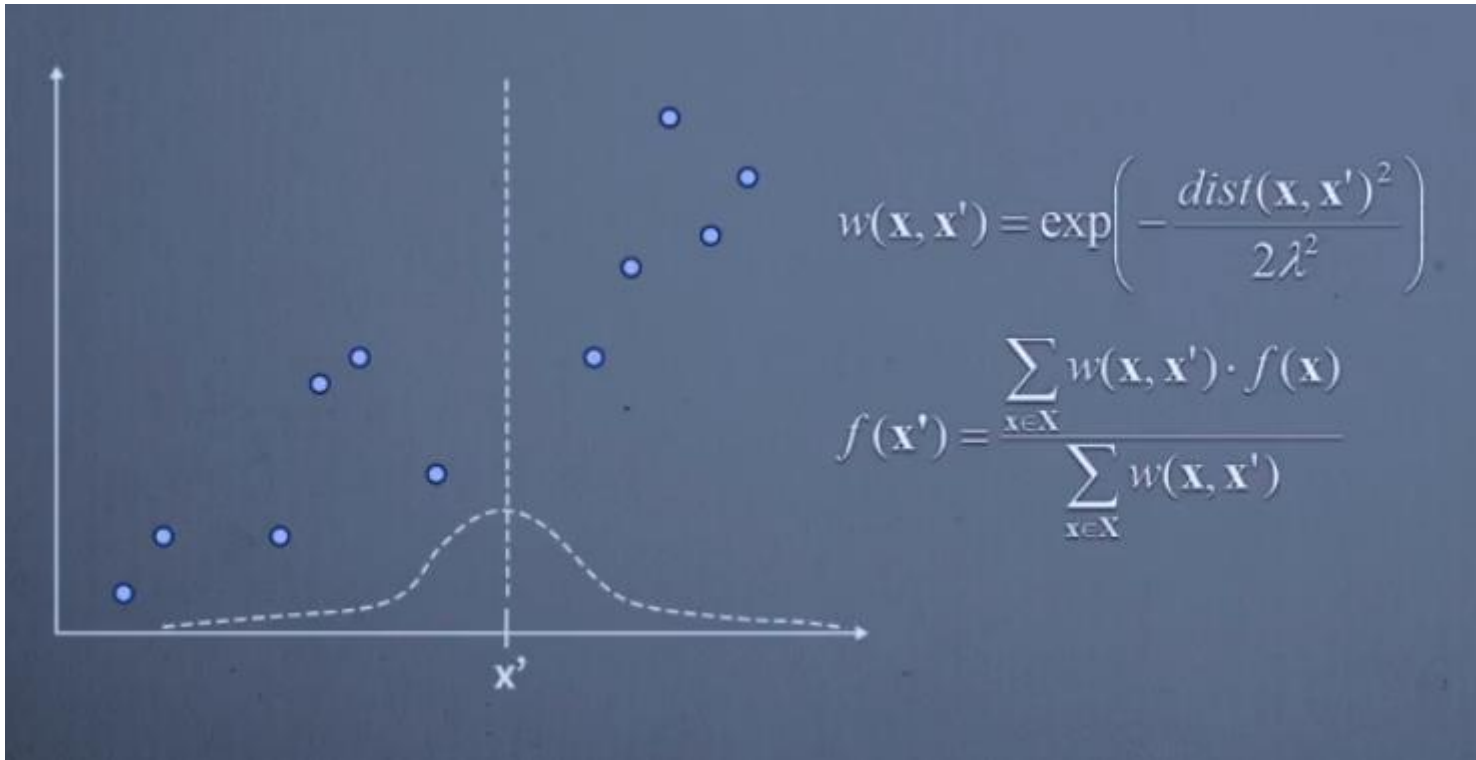
개수와 거리를 같이 고려하자
(x1, x2를 같은 것이라고 판단하기 애매함)
거리를 weigh로 변환하기
=> Weight 함수를 어떻게 지정할 것인가?

» Regression (k=5)

- X1: close to 10 or 1
- X2: close to 10 or 1

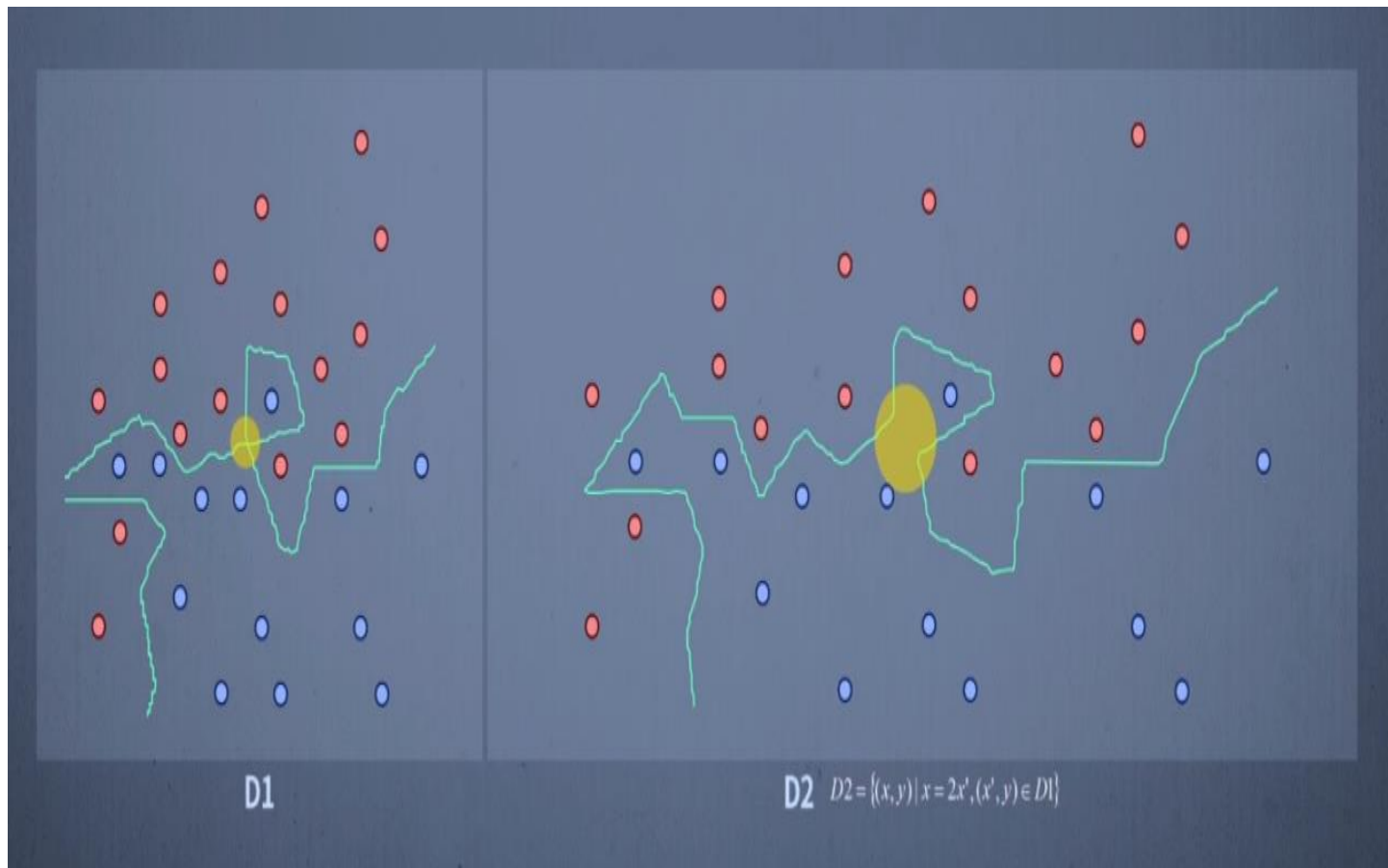


KNN - 변형2



모든 데이터를 사용하자
거리를 weigh로 변환하기
=> Weight 함수를 어떻게 지정할 것인가?
Gaussian weight

고려할 점



분포는 똑같으나 x축으로 2배 된 경우?
경계가 다르게 나옴
X, Y축 변경한 경우 예측 결과 변함

장점과 단점

- 장점
 - No Training
 - No Information loss
- 단점
 - 노이즈에 민감
 - 시간이 굉장히 많이 걸림(모든 데이터 저장해야 함)