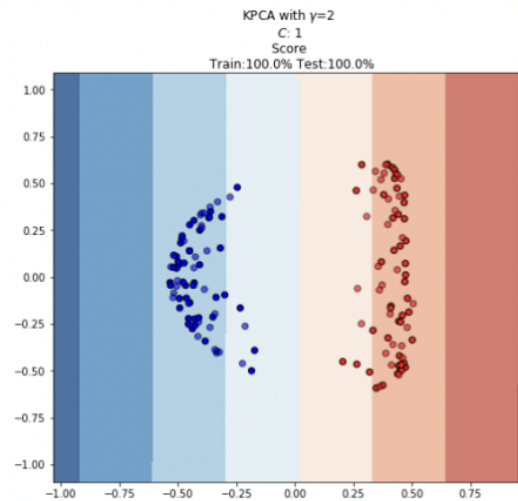
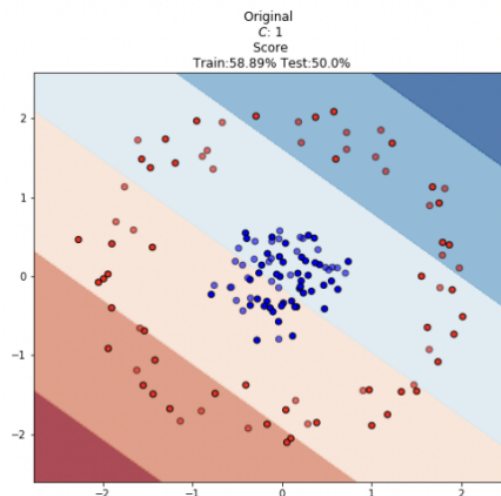




# 차원축소

## 도입

- 변수의 수가 많아 차원이 높은경우 발생하는 문제들
  - 시각화가 어렵다. (3차원 이상은 어떻게 표현해야 할까??)
  - 신경망 또는 알고리즘의 연산시 계산이 복잡해 진다.
- 해결방법
  - 꼭 필요한 변수들만 추출한다.
  - 기존의 변수들을 이용해 새로운 변수를 만들어 낸다.



## 차원축소의 방법

- Feature Selection

- 기존의 변수들중 상관관계가 존재하는 변수를 추출한다.
  - Goal
  - $P_{original} > P_{new}$
  - $x_1, x_2, x_3, \dots, x_n \rightarrow x_1, x_2, \dots, x_m$  단,  $m < n$ 
    - 장점 : 기존의 특성이 가진 성질을 유지하면서 데이터를 줄일수 있다.
    - 단점 : Feature를 압축하지 않고 줄여준다.
- Feature Extraction
    - 기존의 변수들을 변환해 새로운 변수를 추출한다.
    - Goal
    - $P_{original} > P_{new}$
    - $x_1, x_2, x_3, \dots, x_n \rightarrow y_1, y_2, \dots, y_m$  단,  $m < n$ 
      - 장점 : 각각의 변수가 내재한 정보가 많아 진다.
      - 단점 : 생성된 특성은 사람이 이해할수 없는 데이터이다.

## AutoEncoder 의 결과는 무엇을 나타내는 걸까?

