

SVM

유튜브 주소 : <https://youtu.be/IbMd9e8TnoQ>

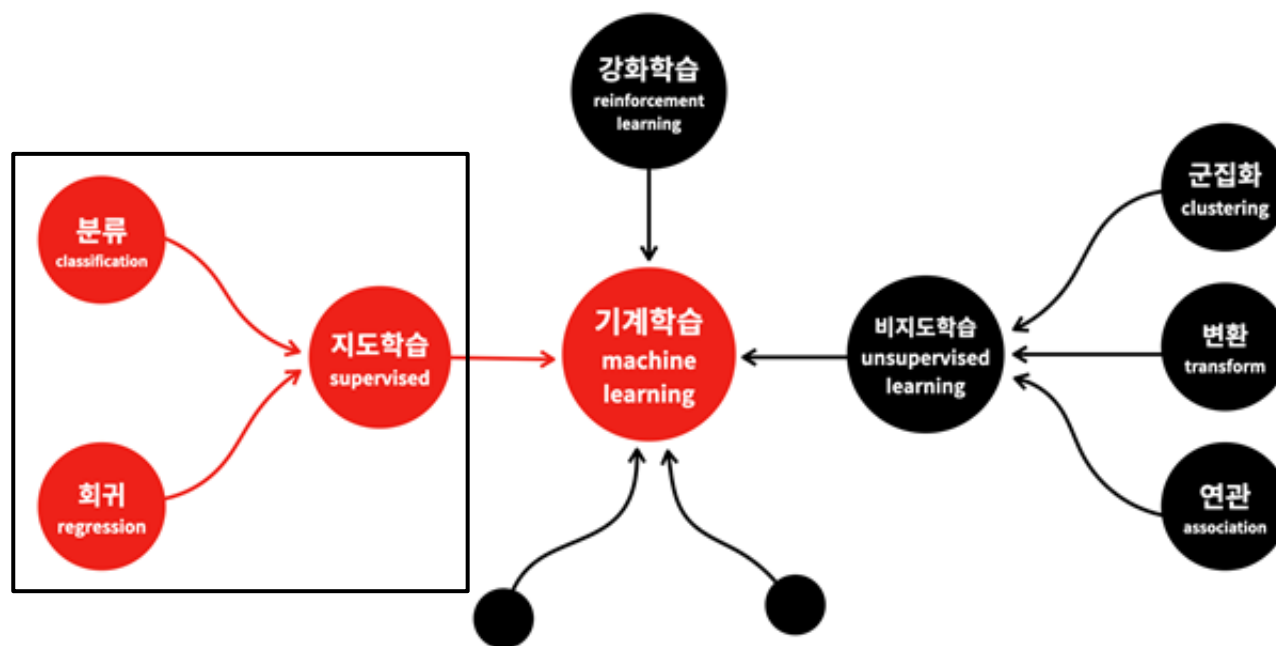
AI 학습 모델

SVM

SVM 실습

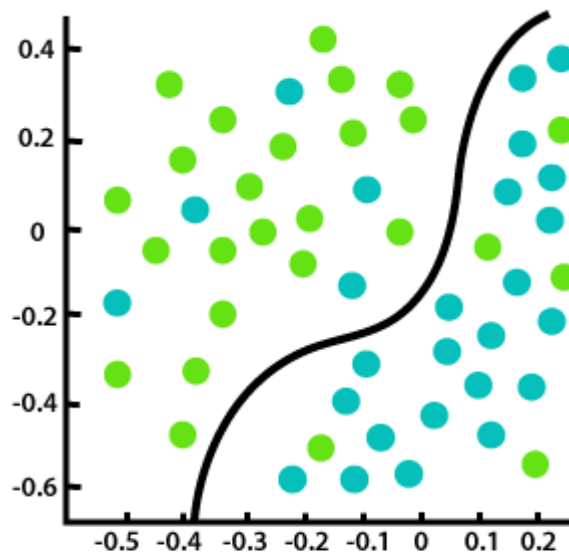
AI 학습 모델

- 분류 – 주어진 데이터를 정해진 카테고리에 따라 분류
- 회귀 – 어떤 데이터들의 특징을 기준으로 연속된 값(그래프)를 예측
- SVM – 분류, 회귀에 쓰이는 머신 러닝 방법

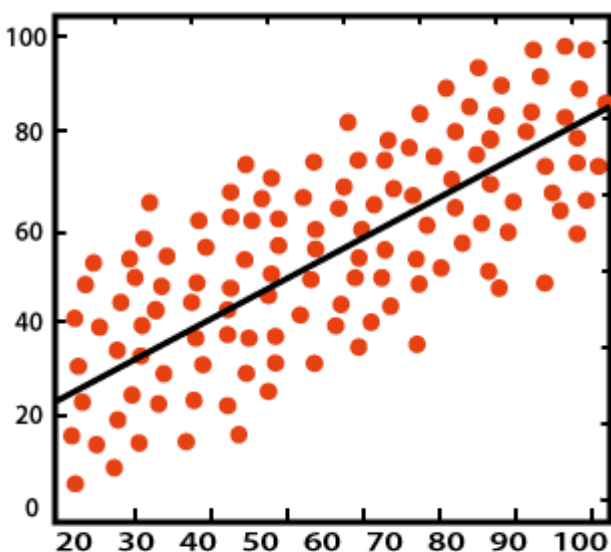


AI 학습 모델

- 분류(Classification) – 이진 분류, 다중 분류
 - 예측 변수가 이름 혹은 문자일 때 주로 사용
- 회귀(Regression) – 패턴, 트렌드, 경향 예측에 사용
 - 예측 변수가 숫자일 때 주로 사용



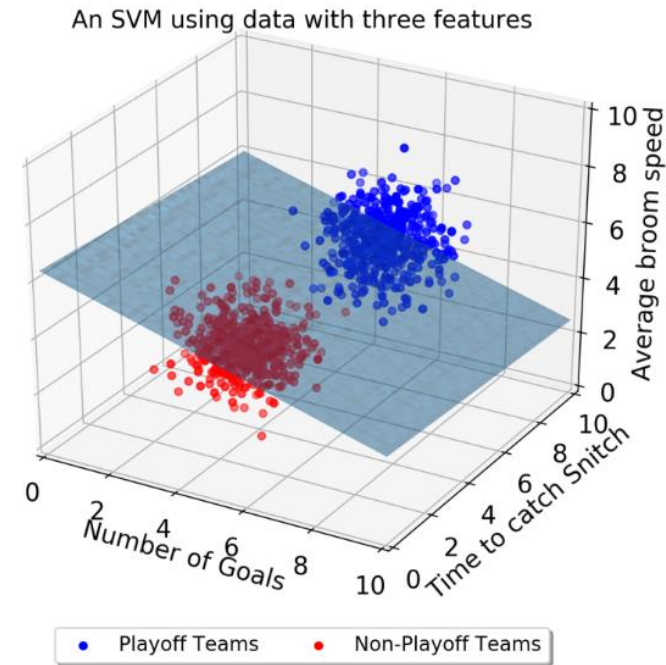
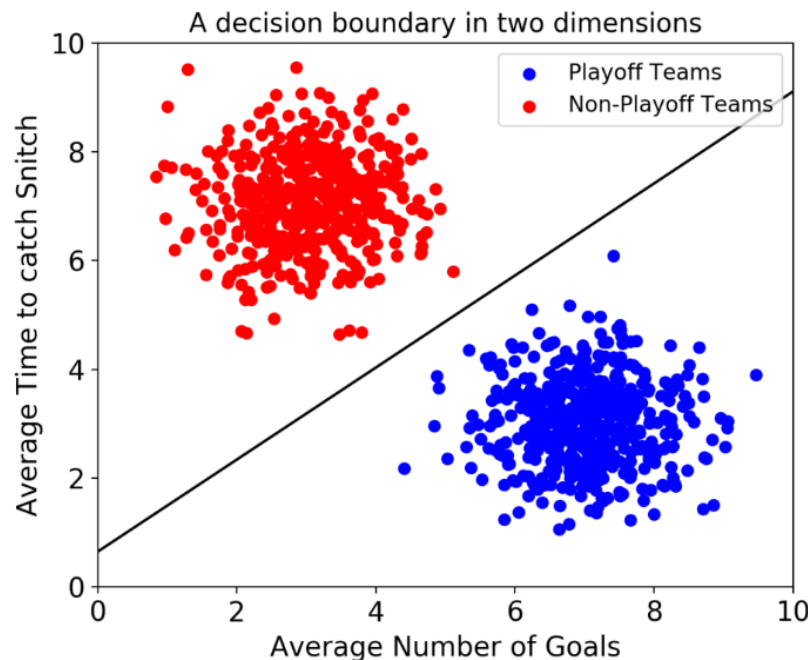
Classification



Regression

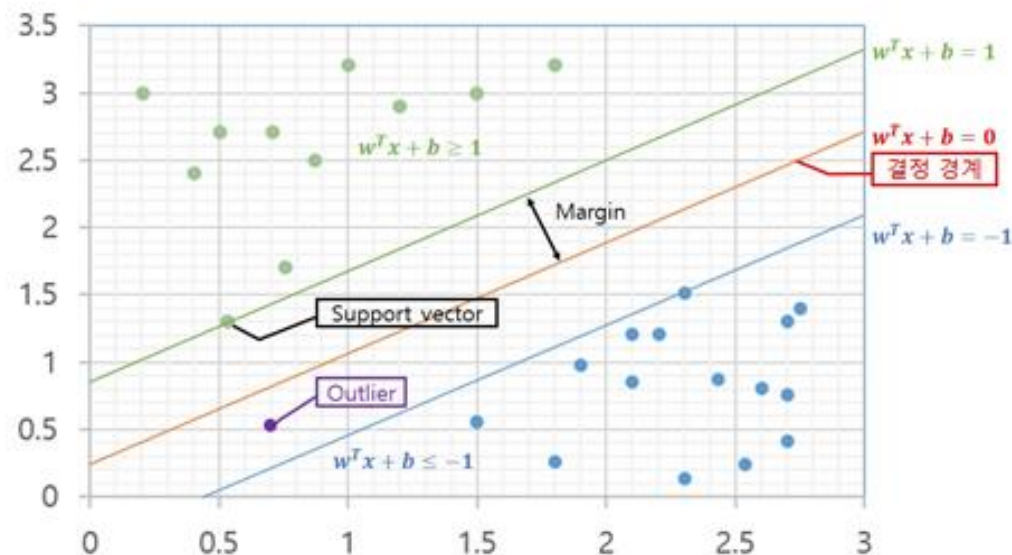
SVM

- SVM – Support Vector Machine
- 분류, 회귀에 쓰이는 지도 학습 머신 러닝 방법
- 결정 경계(Decision Boundary)를 정의하는 모델
 - 결정 경계 – 분류를 위한 기준 선
 - 초평면(Hyperplane) – 고 차원의 결정 경계



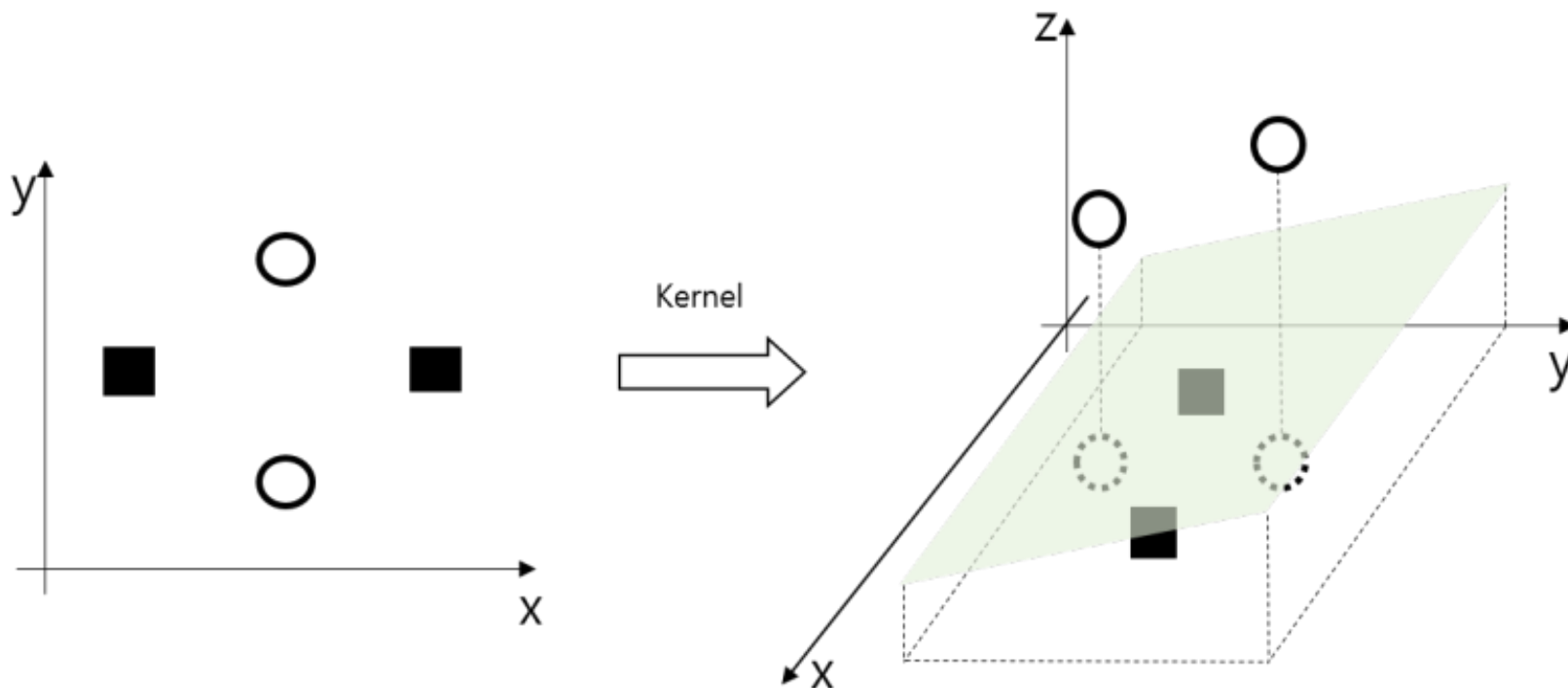
SVM 주요 용어

- 서포트 벡터 – 데이터군의 최전방 데이터 포인트
 - SVM의 결정 경계를 정의
 - N개의 속성을 가진 데이터들은 $n+1$ 개의 서포트 벡터가 존재
- 마진 – 결정 경계와 서포트 벡터 사이의 거리
 - 최적의 결정 경계는 마진을 최대화
- 이상치 – 데이터군과 떨어져 혼자 튀어나와 있는 데이터 포인트
 - 이상치의 허용은 fitting에 영향
 - Hard Margin / Soft Margin



SVM

- 커널 함수
 - 저차원의 데이터를 고차원으로 매핑하여 선형 분류를 하기 위해 사용
- 다항식 커널, 가우시안 커널, 시그모이드 커널
 - 가우시안 커널이 성능이 우수하여 가장 많이 쓰임



SVM

- 장점

- 노이즈 데이터와 이상치를 처리할 수 있음
- 선형 분리가 불가능한 데이터도 고차원 매핑을 통해 분리 가능
- 과적합 되는 경우가 적음
- 신경망보다 사용하기 쉬움
- 오류 데이터 영향이 적음

- 단점

- 여러 개의 조합 테스트 필요
- 학습 속도가 느림
- 해석이 어렵고 복잡한 블랙박스
- 모델 구축 시간이 오래 걸림

Q & A