**머신러닝 12주차 과제**

**20145128**

**박인근**

서포트 벡터 머신(SVM) 이란?

**서포트 벡터 머신**(support vector machine, **SVM**)은 [기계 학습](https://ko.wikipedia.org/wiki/%EA%B8%B0%EA%B3%84_%ED%95%99%EC%8A%B5)의 분야 중 하나로 패턴 인식, 자료 분석을 위한 [지도 학습](https://ko.wikipedia.org/wiki/%EC%A7%80%EB%8F%84_%ED%95%99%EC%8A%B5) 모델이며, 주로 [분류](https://ko.wikipedia.org/wiki/%EB%B6%84%EB%A5%98)와 [회귀 분석](https://ko.wikipedia.org/wiki/%ED%9A%8C%EA%B7%80_%EB%B6%84%EC%84%9D)을 위해 사용한다. 두 카테고리 중 어느 하나에 속한 데이터의 집합이 주어졌을 때, SVM 알고리즘은 주어진 데이터 집합을 바탕으로 하여 새로운 데이터가 어느 카테고리에 속할지 판단하는 비[확률적](https://ko.wikipedia.org/wiki/%ED%99%95%EB%A5%A0) 이진 [선형 분류](https://ko.wikipedia.org/wiki/%EC%84%A0%ED%98%95_%EB%B6%84%EB%A5%98) 모델을 만든다. 만들어진 분류 모델은 데이터가 사상된 공간에서 경계로 표현되는데 SVM 알고리즘은 그 중 가장 큰 폭을 가진 경계를 찾는 알고리즘이다. SVM은 선형 분류와 더불어 비선형 분류에서도 사용될 수 있다. 비선형 분류를 하기 위해서 주어진 데이터를 고차원 특징 공간으로 사상하는 작업이 필요한데, 이를 효율적으로 하기 위해 [커널 트릭](https://ko.wikipedia.org/w/index.php?title=%EC%BB%A4%EB%84%90_%ED%8A%B8%EB%A6%AD&action=edit&redlink=1)을 사용하기도 한다.

정의

일반적으로, 서포트 벡터 머신은 분류 또는 회귀 분석에 사용 가능한 [초평면](https://ko.wikipedia.org/wiki/%EC%B4%88%ED%8F%89%EB%A9%B4)(hyperplane) 또는 초평면들의 집합으로 구성되어 있다. 직관적으로, 초평면이 가장 가까운 학습 데이터 점과 큰 차이를 가지고 있으면 분류 오차(classifier error)가 작기 때문에 좋은 분류를 위해서는 어떤 분류된 점에 대해서 가장 가까운 학습 데이터와 가장 먼 거리를 가지는 초평면을 찾아야 한다. 일반적으로 초기의 문제가 유한 차원 공간에서 다루어지는데, 종종 데이터가 [선형 구분](https://ko.wikipedia.org/wiki/%EC%84%A0%ED%98%95_%EA%B5%AC%EB%B6%84_%EA%B0%80%EB%8A%A5)이 되지 않는 문제가 발생한다. 이러한 문제를 해결하기 위해 초기 문제의 유한 차원에서 더 높은 차원으로 대응시켜 분리를 쉽게 하는 방법이 제안되었다. 그 과정에서 계산량이 늘어나는 것을 막기 위해서, 각 문제에 적절한 [커널 함수](https://ko.wikipedia.org/w/index.php?title=%EC%BB%A4%EB%84%90_%ED%95%A8%EC%88%98&action=edit&redlink=1) {\displaystyle k(x,y)}를 정의한 SVM 구조를 설계하여 [내적 연산](https://ko.wikipedia.org/wiki/%EC%A0%90%EA%B3%B1)을 초기 문제의 변수들을 사용해서 효과적으로 계산할 수 있도록 한다.[[3]](https://ko.wikipedia.org/wiki/%EC%84%9C%ED%8F%AC%ED%8A%B8_%EB%B2%A1%ED%84%B0_%EB%A8%B8%EC%8B%A0#cite_note-3) 높은 차원 공간의 초평면은 점들의 집합과 상수 벡터의 내적 연산으로 정의된다. 초평면에 정의된 벡터들은 데이터 베이스 안에 나타나는 이미지 벡터 매개 변수들과의 선형적 결합이 되도록 선택된다. 이 선택된 초평면에서, 초평면에 대응된 점 x{\displaystyle x}는 다음과 같은 관계가 성립한다. 

{\displaystyle \textstyle \sum \_{i}\alpha \_{i}k(x\_{i},x)=\mathrm {constant} .}만약 {\displaystyle k(x,y)}가x {\displaystyle x} 와 y{\displaystyle y}가 점점 멀어질 수록 작아진다면, 각각의 합은 테스트 점 {\displaystyle x}x와 그와 대응되는 데이터 점{\displaystyle x\_{i}}xixi의 근접성의 정도를 나타내게 된다. 이러한 방식으로, 위 커널식의 합은 구별하고 싶은 집합안에 있는 데이터 점과 테스트 점간의 상대적인 근접성을 측정하는데 사용될 수 있다. 초기 공간에서 볼록하지 않는 집합안의 점 x{\displaystyle x}x가 높은 차원으로 대응되었을 때 오히려 더 복잡하고 어려워질 수도 있는데 이런 부분을 주의해야 한다.