Chapter3. Model-based Anomaly Detection

고려대학교 산업경영공학과

DMQA Lab.

임 새 린

목차

- Auto-Encoder-based Anomaly Detection
- Support Vector-based Anomaly Detection
- Isolation Forest

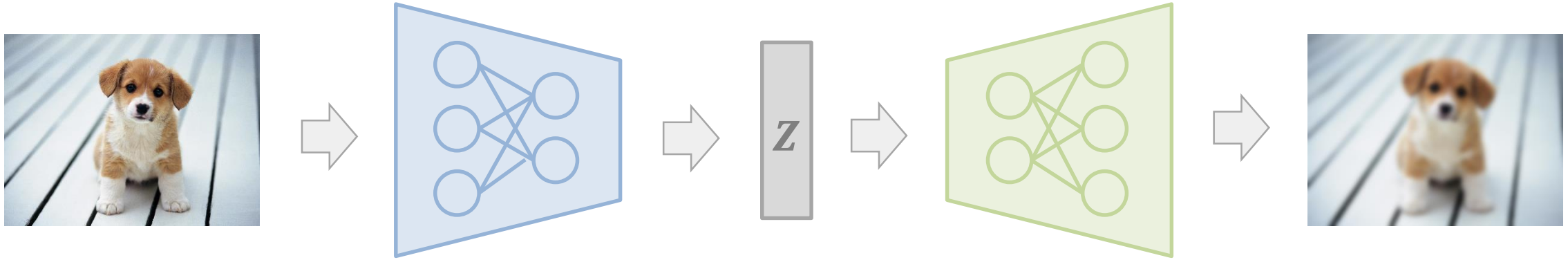
Auto-Encoder-based Anomaly Detection

Auto-Encoder-based Anomaly Detection

Definition

❖ Auto-Encoder

- Encoder : raw data를 저차원 벡터 z로 요약
- Decoder : 요약된 벡터 z를 다시 복원
- 주로 특징추출, 차원축소에 활용



$Raw\ data$
X

$Encoder$
$f(X)$

$Extracted\ feature$
Z

$Decoder$
$g(Z)$

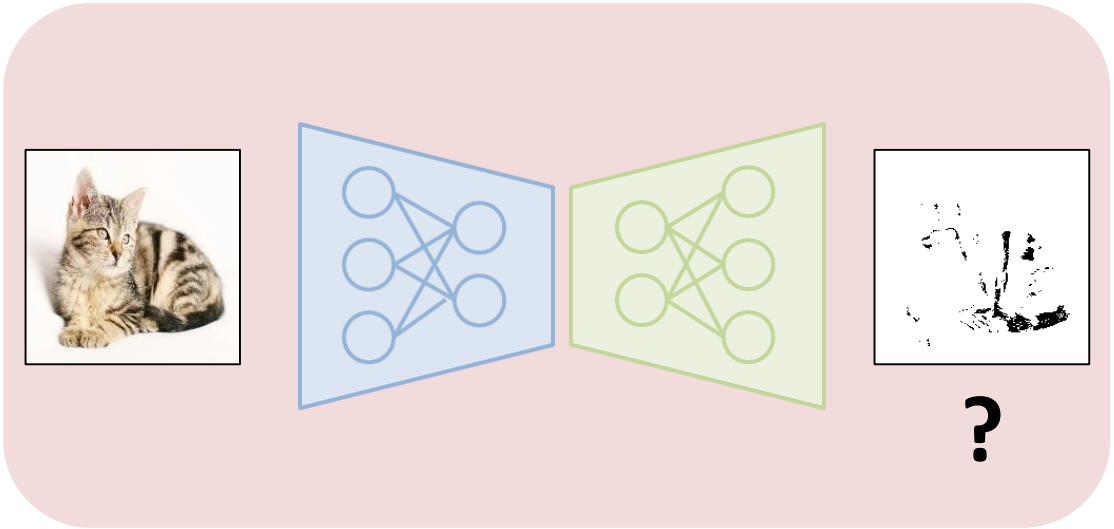
$Reconstruction$
X'

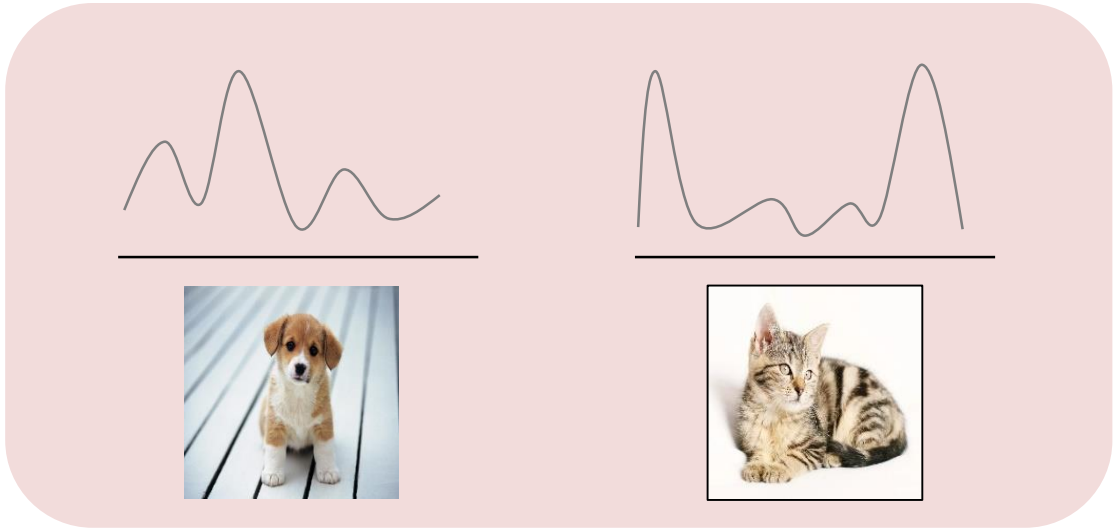
Auto-Encoder-based Anomaly Detection

Anomaly Detection

❖ How to use auto-encoder for anomaly detection?

- 가정1 : 학습 데이터와 다른 패턴을 가지는 데이터가 들어온다면 복원 성능이 떨어질 것이다.
- 가정2 : 이상 데이터는 정상 데이터와 패턴이 다를 것이다.
- 위 두 가정을 바탕으로 auto-encoder-based anomaly detection framework를 개발



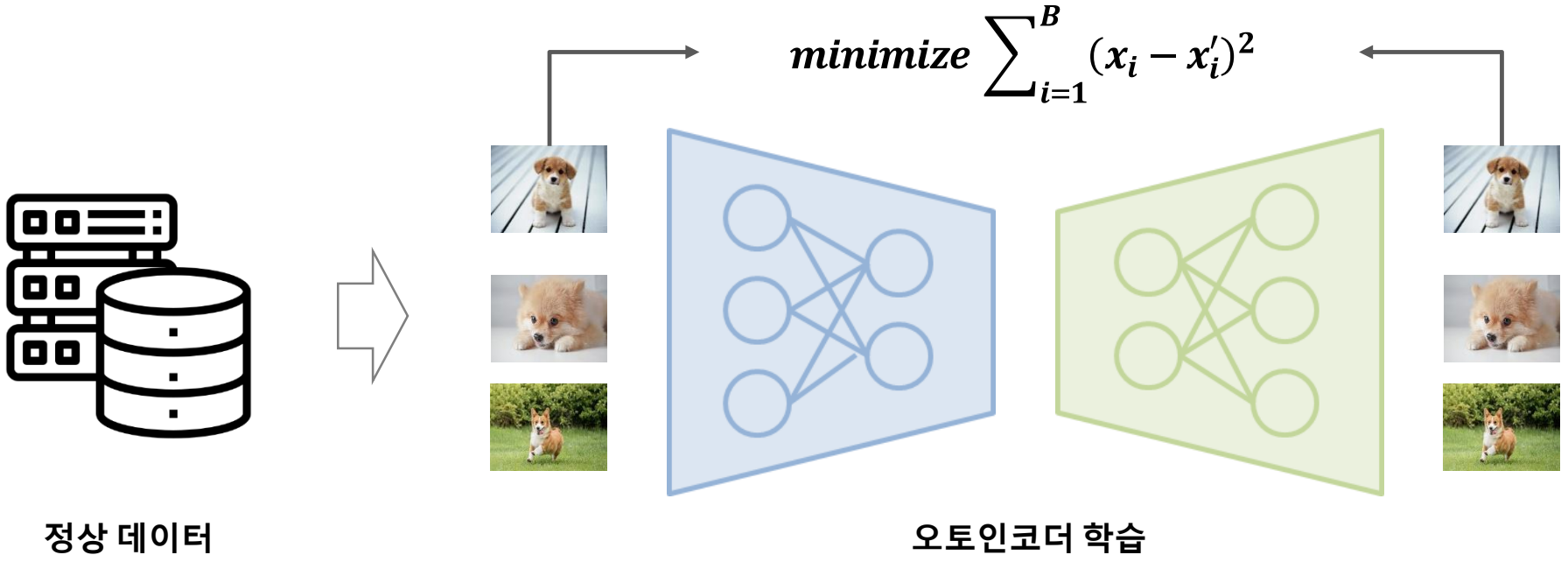


Auto-Encoder-based Anomaly Detection

Anomaly Detection

❖ Training Stage

- 정상 데이터만을 활용하여 오토인코더가 정상 패턴을 잘 복원할 수 있도록 학습
- 모델의 입력값과 복원값의 차이를 최소화 하는 방향으로 학습 진행

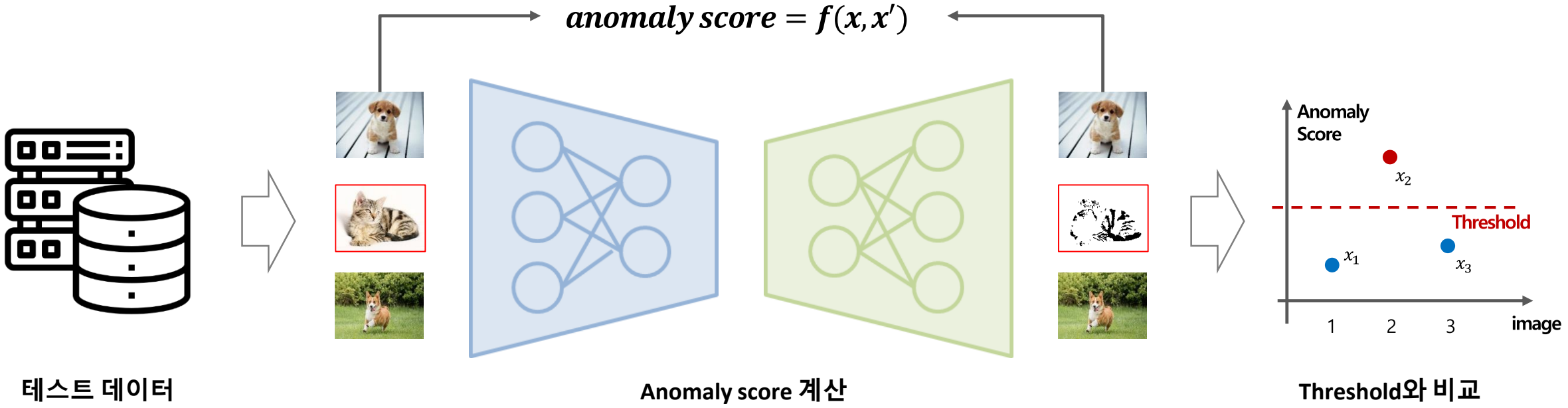


Auto-Encoder-based Anomaly Detection

Anomaly Detection

❖ Anomaly Detection Stage

- 학습이 완료된 오토인코더에 테스트 데이터를 입력하여 복원값 출력
- 입력값과 복원값을 통해 anomaly score를 계산하고 anomaly score가 threshold를 넘는다면 이상으로 판단

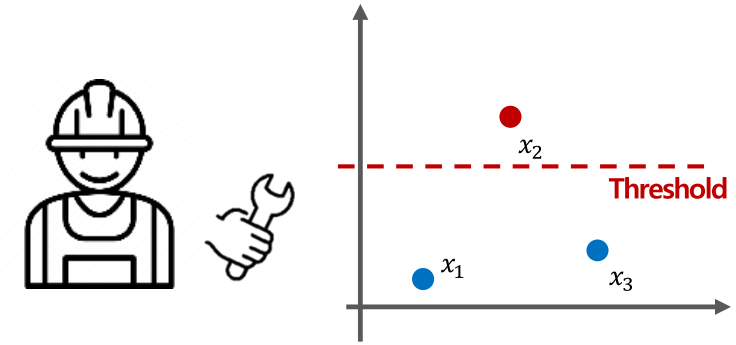


Auto-Encoder-based Anomaly Detection

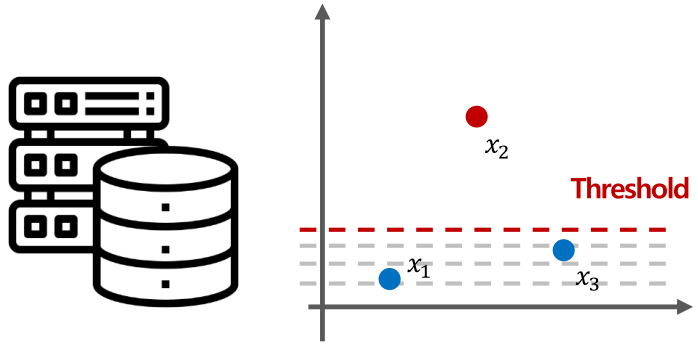
Anomaly Detection

❖ How to set the threshold?

- Domain knowledge
- Grid search with validation set



도메인 지식을 활용한 threshold 설정



검증 데이터를 활용한 threshold 설정

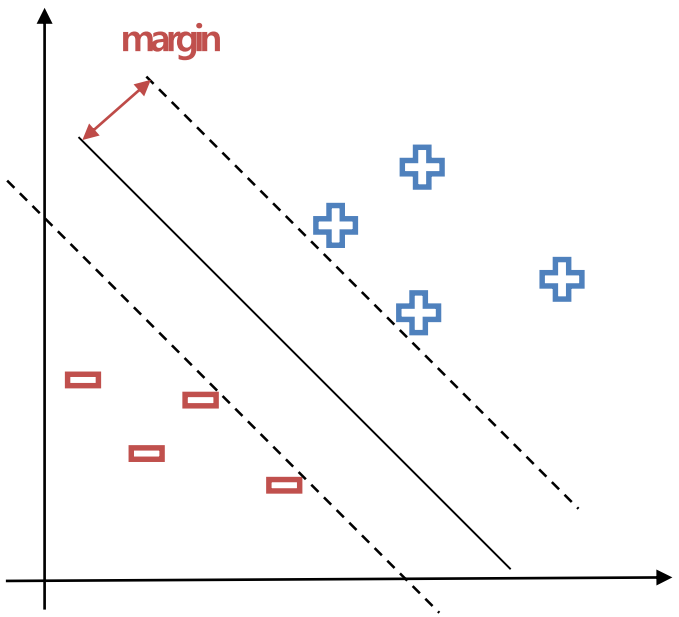
Support Vector-based Anomaly Detection

Support Vector-based Anomaly Detection

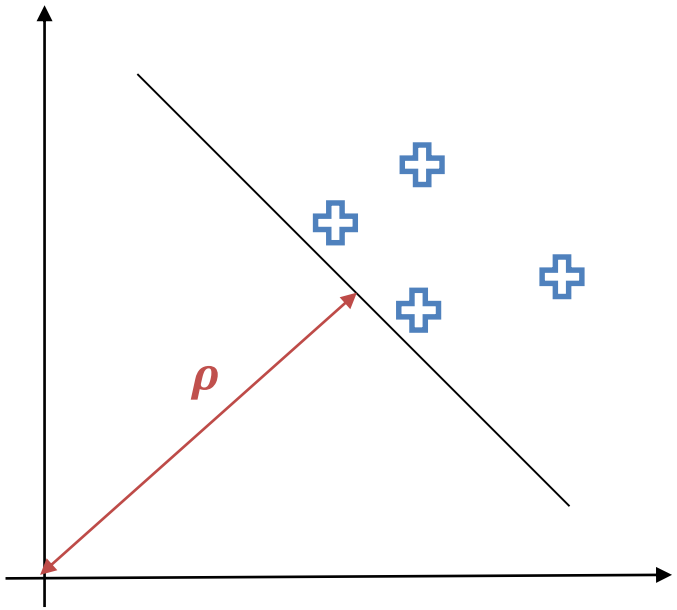
Definition

❖ One-class SVM & SVDD

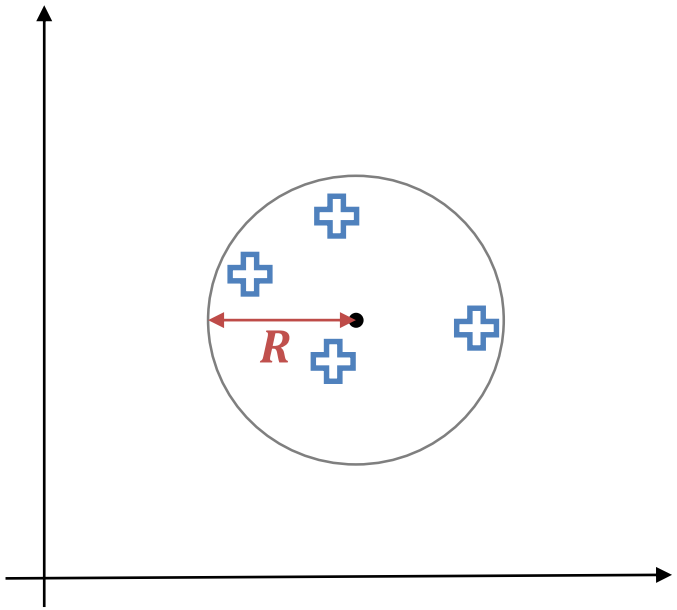
- One-class SVM : 원점으로부터 정상 데이터가 가장 멀어지도록 하는 hyperplane을 찾자
- SVDD : 정상 데이터를 모두 포함할 수 있는 가장 작은 hypersphere를 찾자



2-class SVM



1-class SVM



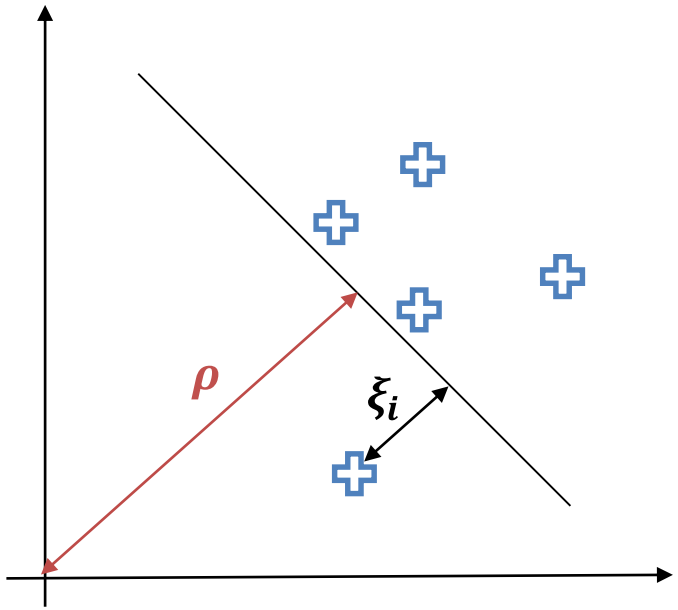
SVDD

Support Vector-based Anomaly Detection

One-class SVM

❖ One-class SVM

- One-class SVM : 원점으로부터 정상 데이터가 가장 멀어지도록 하는 hyperplane을 찾자



1-class SVM

$$\min_{w} \frac{1}{2}||w||^2 + \frac{1}{vl}\sum_{i=1}^{l}\xi_i - \rho$$

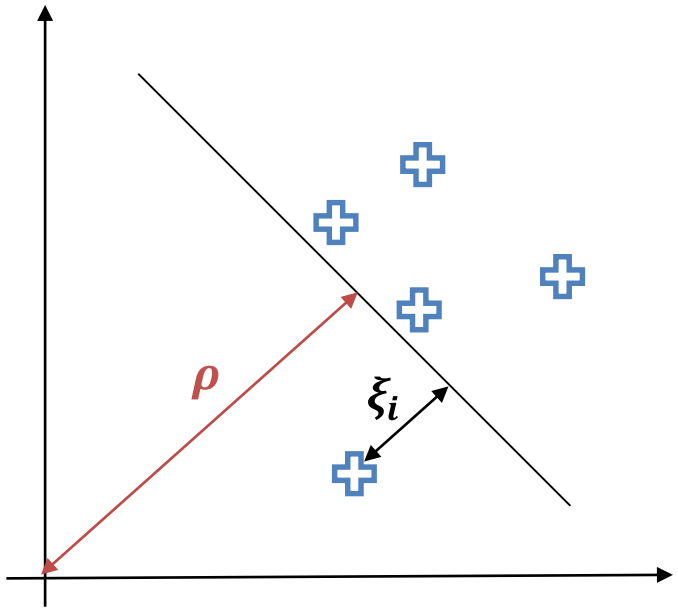
$$s.t.\ w \cdot \Phi(x_i) \geq \rho - \xi_i, \quad \xi_i, \geq 0 \quad for\ all\ i$$

Support Vector-based Anomaly Detection

One-class SVM

❖ One-class SVM

- One-class SVM : 원점으로부터 정상 데이터가 가장 멀어지도록 하는 hyperplane을 찾자



1-class SVM

원점과 거리 최대화

$$\min_{w} \frac{1}{2}||w||^2 + \frac{1}{vl}\sum_{i=1}^{l}\xi_i - \rho$$

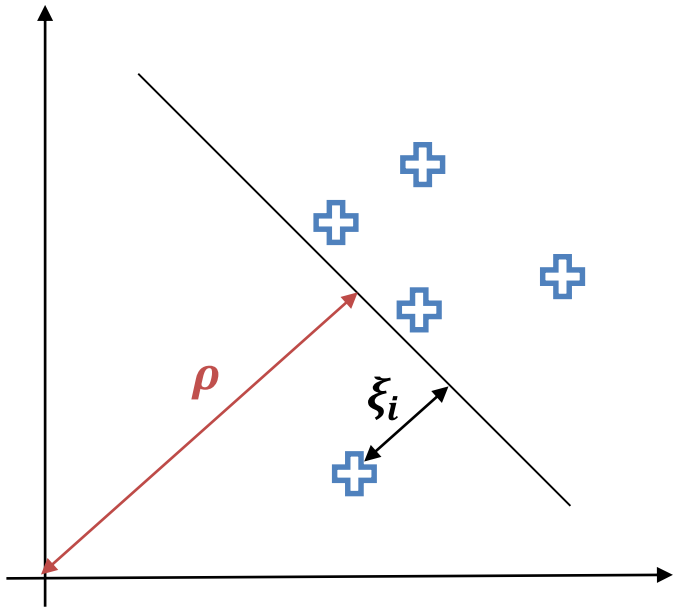
$$s.t.\ w \cdot \Phi(x_i) \geq \rho - \xi_i, \quad \xi_i, \geq 0 \quad for\ all\ i$$

Support Vector-based Anomaly Detection

One-class SVM

❖ One-class SVM

- One-class SVM : 원점으로부터 정상 데이터가 가장 멀어지도록 하는 hyperplane을 찾자



1-class SVM

Hyperplane 바깥에 있는 정상 데이터 패널티

$$\min_{w} \frac{1}{2}||w||^2 + \frac{1}{vl}\sum_{i=1}^{l}\xi_i - \rho$$

$$0 \le v \le 1$$

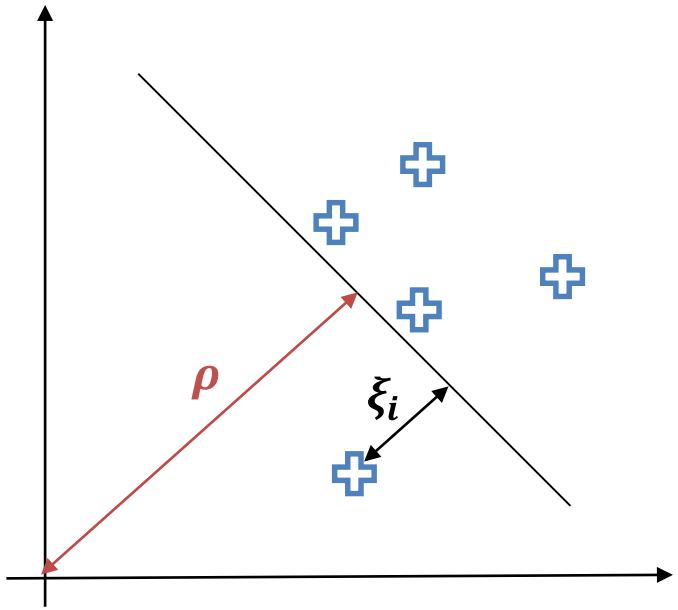
$$s.t.\ w \cdot \Phi(x_i) \ge \rho - \xi_i,\quad \xi_i, \ge 0 \quad for\ all\ i$$

Support Vector-based Anomaly Detection

One-class SVM

❖ One-class SVM

- One-class SVM : 원점으로부터 정상 데이터가 가장 멀어지도록 하는 hyperplane을 찾자



1-class SVM

모델 변동성 감소

$$\min_{w} \frac{1}{2}||w||^2 + \frac{1}{vl}\sum_{i=1}^{l}\xi_i - \rho$$

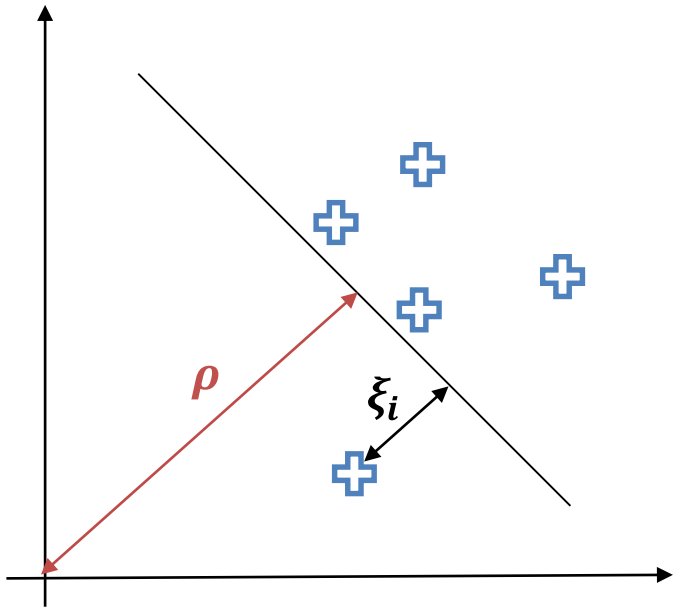
$$s.t.\ w \cdot \Phi(x_i) \geq \rho - \xi_i, \quad \xi_i, \geq 0 \quad for\ all\ i$$

Support Vector-based Anomaly Detection

One-class SVM

❖ One-class SVM

- One-class SVM : 원점으로부터 정상 데이터가 가장 멀어지도록 하는 hyperplane을 찾자



1-class SVM

$$\min_{w} \frac{1}{2}||w||^2 + \frac{1}{vl}\sum_{i=1}^{l}\xi_i - \rho$$

$$s.t.\ w \cdot \Phi(x_i) \geq \rho - \xi_i, \quad \xi_i, \geq 0 \quad for\ all\ i$$

모든 정상 데이터가
hyperplane 안쪽에 존재

Support Vector-based Anomaly Detection

One-class SVM

❖ One-class SVM

- One-class SVM : 원점으로부터 정상 데이터가 가장 멀어지도록 하는 hyperplane을 찾자

Primal Problem

$$L = \frac{1}{2}||w||^2 + \frac{1}{vl}\sum_{i=1}^{l}\xi_i - \rho$$

$$s.t.\ w \cdot \Phi(x_i) \geq \rho - \xi_i,\ \ \xi_i, \geq 0 \quad for\ all\ i$$

Lagrangian Problem

$$L = \frac{1}{2}||w||^2 + \frac{1}{vl}\sum_{i=1}^{l}\xi_i - \rho - \sum_{i=1}^{l}\alpha_i(w \cdot \Phi(x_i) - \rho + \xi_i) - \sum_{i=1}^{l}\beta_i\xi_i$$

$$s.t.\ \alpha_i \geq 0,\ \ \beta_i \geq 0\ for all\ i$$

KKT condition

- $\frac{\partial L_p}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{N}\alpha_i\Phi(x_i)$
- $\frac{\partial L_p}{\partial \rho} = 0 \rightarrow \sum_{i=1}^{N}\alpha_i = 1$
- $\frac{\partial L_p}{\partial \xi_i} = 0 \rightarrow \alpha_i = \frac{1}{vl} - \beta_i$

Support Vector-based Anomaly Detection

One-class SVM

❖ One-class SVM

- One-class SVM : 원점으로부터 정상 데이터가 가장 멀어지도록 하는 hyperplane을 찾자

Lagrangian Problem

$$L = \frac{1}{2}\left|\left|w\right|\right|^2 + \frac{1}{vl}\sum_{i=1}^{l}\xi_i - \rho - \sum_{i=1}^{l}\alpha_i(w \cdot \Phi(x_i) - \rho + \xi_i) - \sum_{i=1}^{l}\beta_i\xi_i$$

$$s.t.\ \alpha_i \geq 0,\ \ \beta_i \geq 0\ for all\ i$$

Dual Problem

$$\max\ L_D = \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j\Phi(x_i)\,\Phi(x_j) = \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j \mathrm{K}(x_i, x_j)$$

$$s.t.\ \sum_{i=1}^{N}\alpha_i = 1 and\ 0 \leq \alpha_i \leq \frac{1}{vl}$$

Support Vector-based Anomaly Detection

One-class SVM

❖ One-class SVM

- One-class SVM : 원점으로부터 정상 데이터가 가장 멀어지도록 하는 hyperplane을 찾자

KKT condition

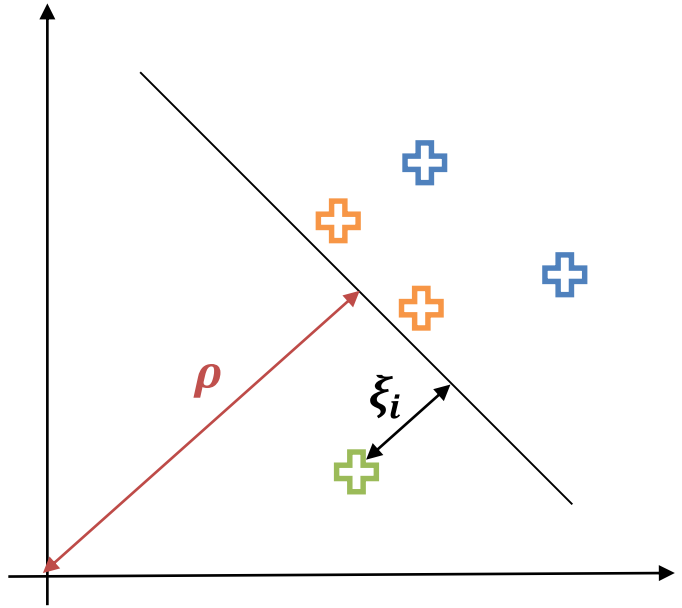
$$\alpha_i(w \cdot \Phi(x_i) - \rho + \xi_i) = 0$$

$$\alpha_i = \frac{1}{\nu l} - \beta_i \quad \& \quad \beta_i \xi_i = 0$$

Case 1 : $\alpha_i = 0 \rightarrow non - support\,vector$

Case 2 : $\alpha_i = \frac{1}{\nu l} \rightarrow \beta_i = 0 \rightarrow \xi_i > 0 \rightarrow support\,vector(outside)$

Case 3 : $0 < \alpha_i < \frac{1}{\nu l} \rightarrow \beta_i > 0 \rightarrow \xi_i = 0 \rightarrow support\,vector(on\,the\,hyperplane)$



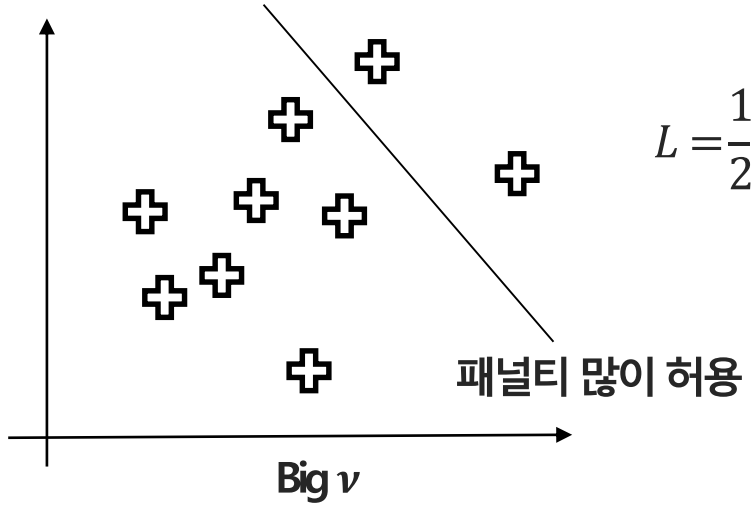
1-class SVM

Support Vector-based Anomaly Detection

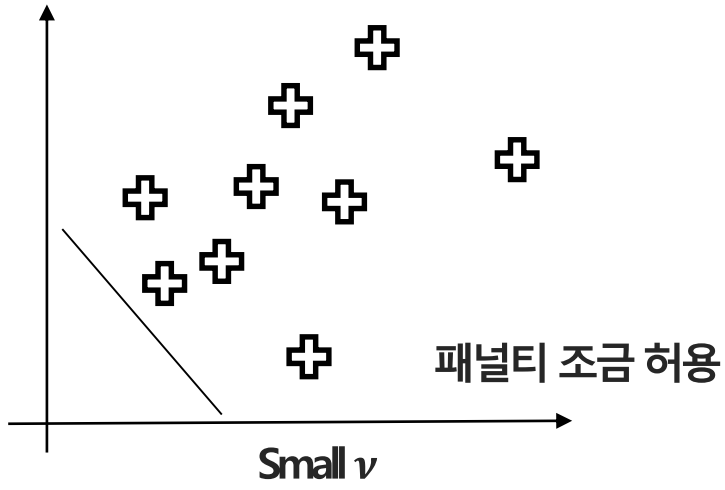
One-class SVM

❖ Role of ν

- α_i가 가질 수 있는 최대값 : $\frac{1}{\nu l}$ ($\because 0 \leq \alpha_i \leq \frac{1}{\nu l}$)
- 따라서 모든 α_i가 $\frac{1}{\nu l}$일 때, support vector의 개수 : νl → support vector 개수의 최소값($\because \sum_{i=1}^{l} \alpha_i = 1$)
- Case 2인 support vector의 최대값(outside) = νl
- 즉, ν는 support vector 개수의 최소값을 결정함과 동시에 에러를 허용한 support vector의 최대값을 결정
- ν의 값이 커질수록 패널티를 많이 줄 수 있기 때문에 hyperplane이 원점과 더 멀어지며 더 compact한 결정 경계 생성



$$L = \frac{1}{2}||w||^2 + \frac{1}{\nu l}\sum_{i=1}^{l} \xi_i - \rho$$

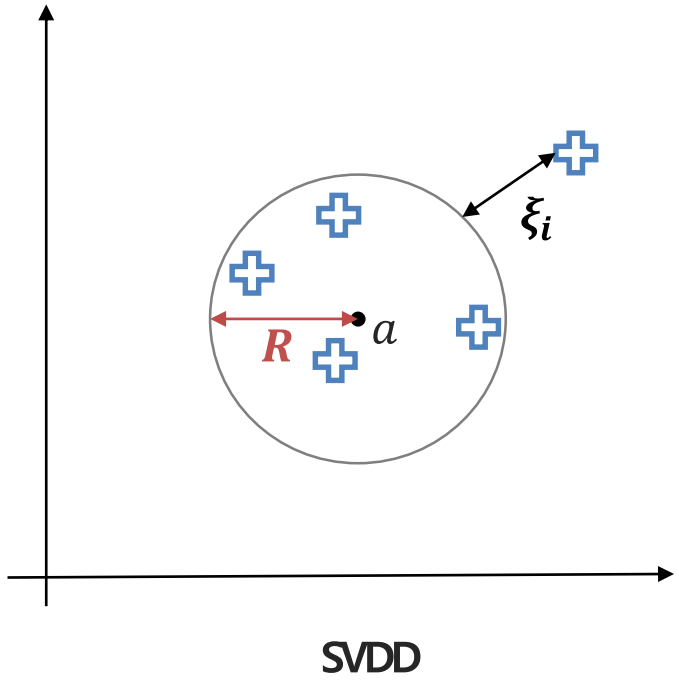


Support Vector-based Anomaly Detection

Support Vector Data Description

❖ SVDD

- SVDD : 정상 데이터를 모두 포함할 수 있는 가장 작은 hypersphere를 찾자



$$\min_{R,a,\xi_i} R^2 + C\sum_{i=1}^{l}\xi_i$$

$$s.t.\ \|\Phi(x_i) - a\|^2 \le R^2 + \xi_i,\quad \xi_i, \ge 0 \quad for\ all\ i$$

Support Vector-based Anomaly Detection

Support Vector Data Description

❖ SVDD

- SVDD : 정상 데이터를 모두 포함할 수 있는 가장 작은 hypersphere를 찾자

Primal Problem

$$L = R^2 + C\sum_{i=1}^{l}\xi_i$$

$$s.t.\ \|\Phi(x_i) - a\|^2 \le R^2 + \xi_i,\ \ \xi_i, \ge 0 \quad for\ all\ i$$

Lagrangian Problem

$$L = R^2 + C\sum_{i=1}^{l}\xi_i - \sum_{i=1}^{l}\alpha_i\left(R^2 + \xi_i - \left(\Phi(x_i)\,\Phi(x_j)\right) - 2a\Phi(x_i) + a^2\right) - \sum_{i=1}^{l}\beta_i\xi_i$$

$$s.t.\ \alpha_i \ge 0,\ \ \beta_i \ge 0\ for\ all\ i$$

KKT condition

- $\frac{\partial L_p}{\partial R} = 0 \rightarrow 2R - 2R\sum_{i=1}^{N}\alpha_i = 0 \rightarrow \sum_{i=1}^{N}\alpha_i = 1$
- $\frac{\partial L_p}{\partial a} = 0 \rightarrow a = \sum_{i=1}^{N}\alpha_i\,\Phi(x_i)$
- $\frac{\partial L_p}{\partial \xi_i} = 0 \rightarrow C - \alpha_i - \beta_i = 0$

Support Vector-based Anomaly Detection

Support Vector Data Description

❖ SVDD

- SVDD : 정상 데이터를 모두 포함할 수 있는 가장 작은 hypersphere를 찾자

Primal Problem

$$L = R^2 + C\sum_{i=1}^{l}\xi_i \quad -\sum_{i=1}^{l}\beta_i\xi_i$$
$$-\sum_{i=1}^{l}\alpha_i\left(R^2+\xi_i-\left(\Phi(x_i)\,\Phi(x_j)\right)-2a\Phi(x_i)+a^2\right)$$

$$s.t.\ \alpha_i \geq 0,\ \ \beta_i \geq 0\ for all\ i$$

Lagrangian Problem

$$L_D = \sum_{i=1}^{l}\alpha_i\,\Phi(x_i)\,\Phi(x_j) - \sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j\,\Phi(x_i)\,\Phi(x_j)$$
$$= \sum_{i=1}^{l}\alpha_i\,K(x_i,x_j) - \sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j\,K(x_i,x_j)$$

$$s.t.\ 0 \leq \alpha_i \leq C$$

Support Vector-based Anomaly Detection

Support Vector Data Description

❖ SVDD

- SVDD : 정상 데이터를 모두 포함할 수 있는 가장 작은 hypersphere를 찾자

KKT condition

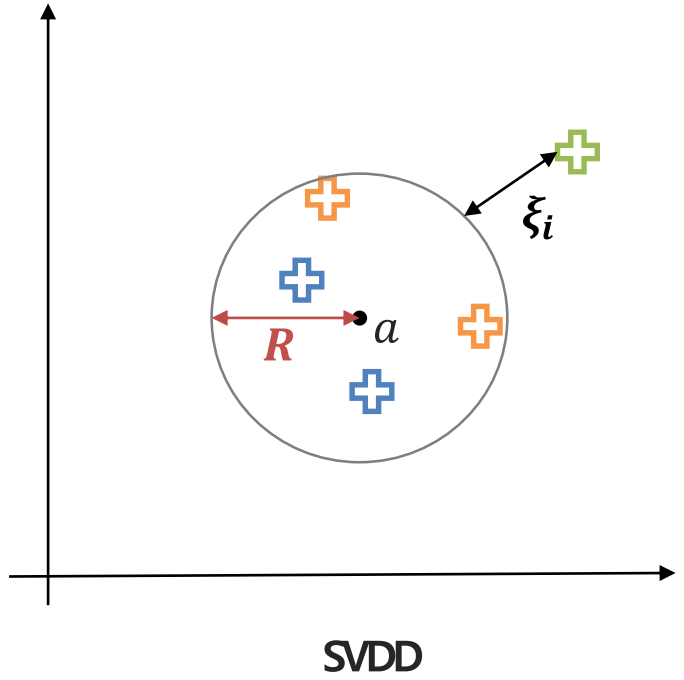
$$\alpha_i \left(R^2 + \xi_i - \left(\Phi(x_i)\, \Phi(x_j) \right) - 2a\Phi(x_i) + a^2 \right) = 0$$

$$C - \alpha_i - \beta_i = 0 \quad \& \quad \beta_i \xi_i = 0$$

Case 1 : $\alpha_i = 0 \rightarrow non - support\,vector$

Case 2 : $\alpha_i = C \rightarrow \beta_i = 0 \rightarrow \xi_i > 0 \rightarrow support\,vector(outside)$

Case 3 : $0 < \alpha_i < C \rightarrow \beta_i > 0 \rightarrow \xi_i = 0 \rightarrow support\,vector(on\,the\,hyperplane)$

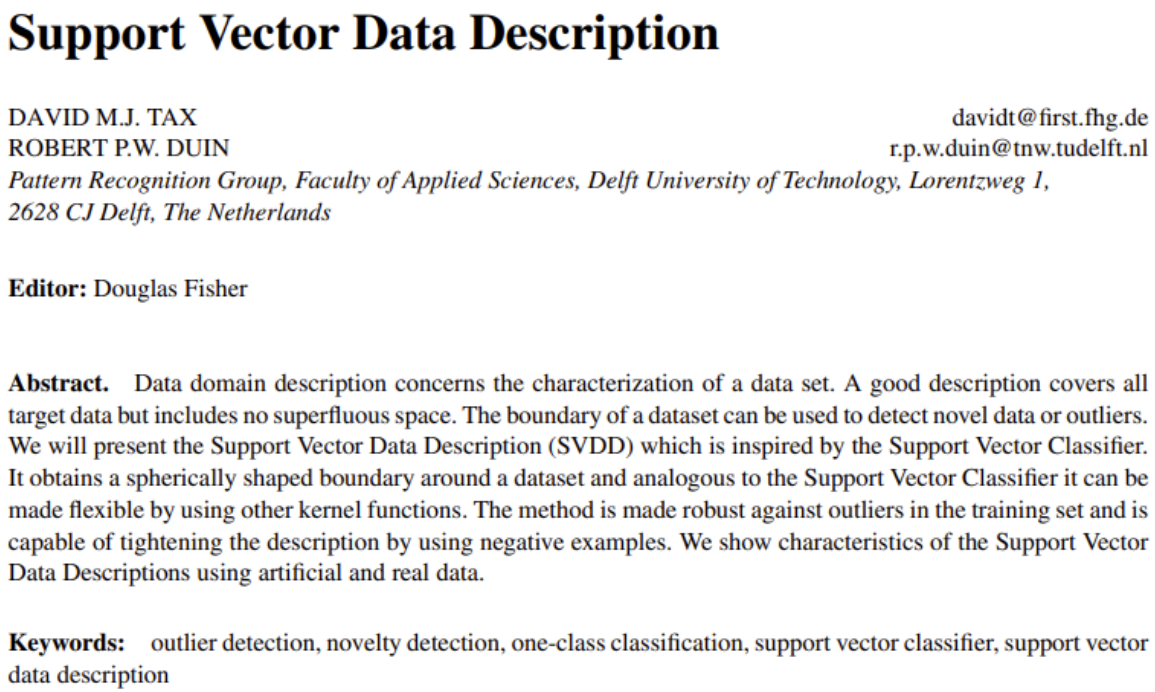


Support Vector-based Anomaly Detection

Support Vector Data Description

❖ One class SVM = SVDD?

- 모든 데이터가 unit norm vector로 정규화된다면 SVDD와 one class SVM은 같아진다.



Support Vector Data Description

DAVID M.J. TAX davidt@first.fhg.de
ROBERT P.W. DUIN r.p.w.duin@tnw.tudelft.nl
Pattern Recognition Group, Faculty of Applied Sciences, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

Editor: Douglas Fisher

Abstract. Data domain description concerns the characterization of a data set. A good description covers all target data but includes no superfluous space. The boundary of a dataset can be used to detect novel data or outliers. We will present the Support Vector Data Description (SVDD) which is inspired by the Support Vector Classifier. It obtains a spherically shaped boundary around a dataset and analogous to the Support Vector Classifier it can be made flexible by using other kernel functions. The method is made robust against outliers in the training set and is capable of tightening the description by using negative examples. We show characteristics of the Support Vector Data Descriptions using artificial and real data.

Keywords: outlier detection, novelty detection, one-class classification, support vector classifier, support vector data description

Tax, David MJ, and Robert PW Duin. "Support vector data description." *Machine learning* 54.1 (2004): 45-66.

Isolation Forest

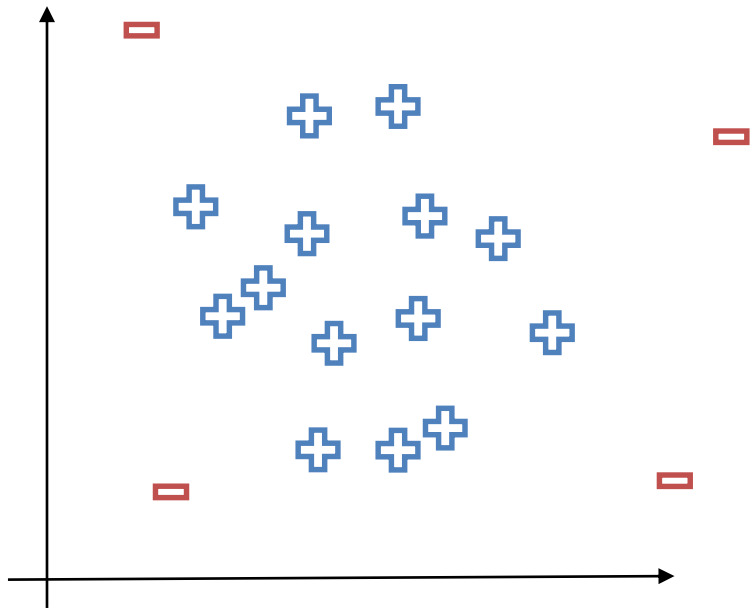
Isolation Forest

Motivation

❖ Motivation of isolation forest

- 일반적으로 이상치의 개수는 적음
- 또한, 정상 데이터와는 특정 변수의 값이 많이 다를 가능성이 높음

관측치	변수A	변수B	범주
1	3	0.3	정상
2	2	0.7	정상
3	6	0.5	정상
4	1	0.6	정상
5	7	0.7	정상
6	35	0.5	이상
7	1	0.6	정상

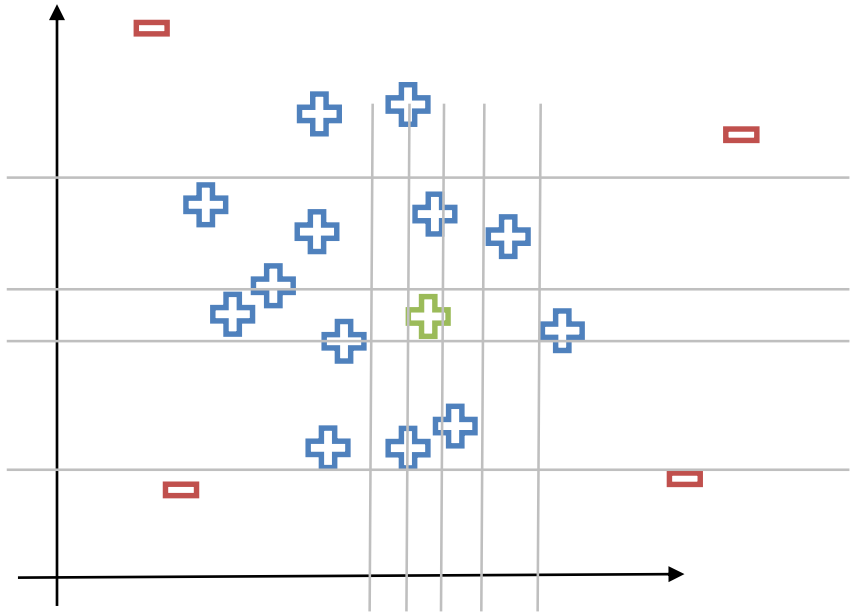


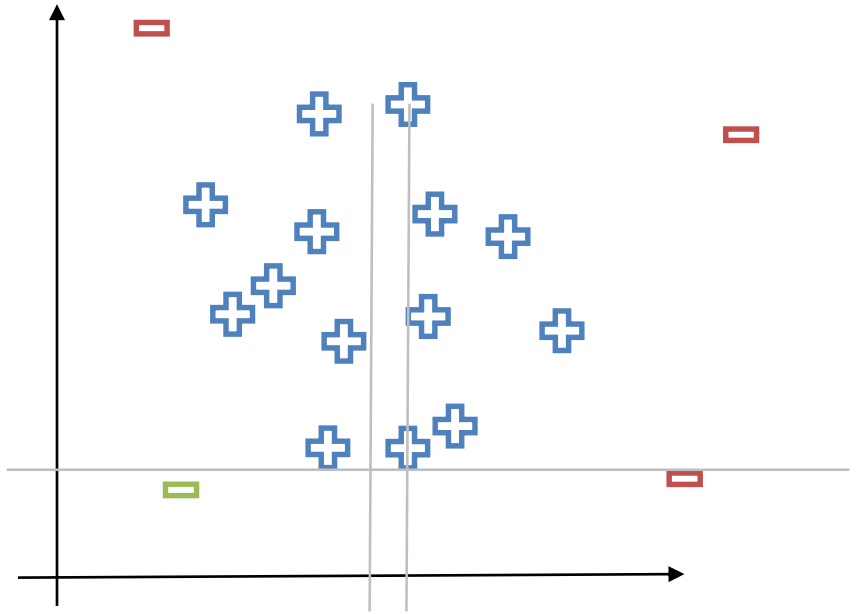
Isolation Forest

Idea

❖ Idea of isolation forest

- 일반적으로 이상치의 개수는 적음
- 또한, 정상 데이터와는 특정 변수의 값이 많이 다를 가능성이 높음



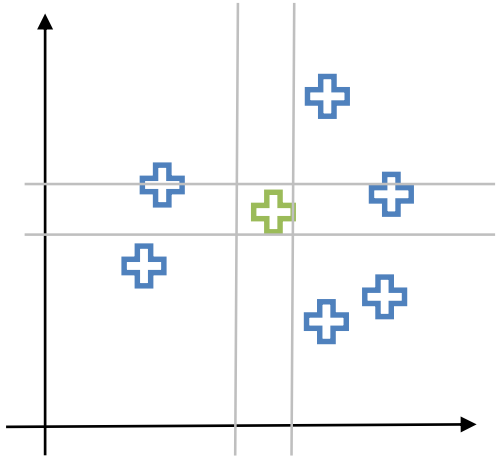


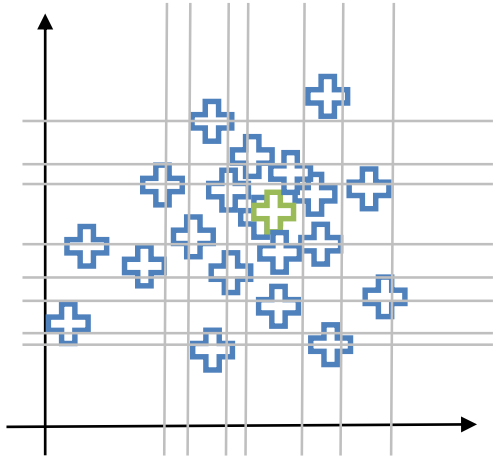
Isolation Forest

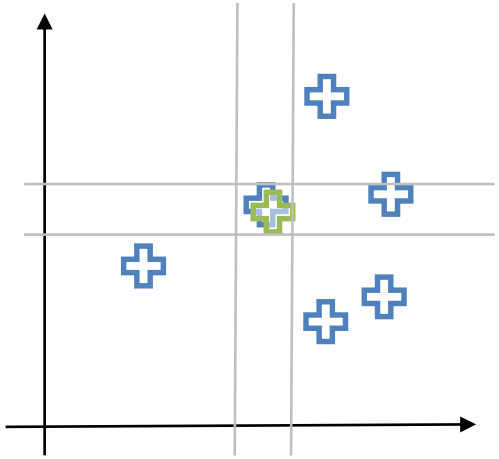
algorithm

❖ split method and terminal conditions

- Split variable과 value는 랜덤하게 선택
- 알고리즘 종료 조건으로 세 가지 제시
  - 입력 관측치가 고립
  - Tree가 사전에 설정한 최대 깊이에 도달
  - 고립된 영역에서 입력 관측치와 값이 같은 관측치들만 존재







Isolation Forest

algorithm

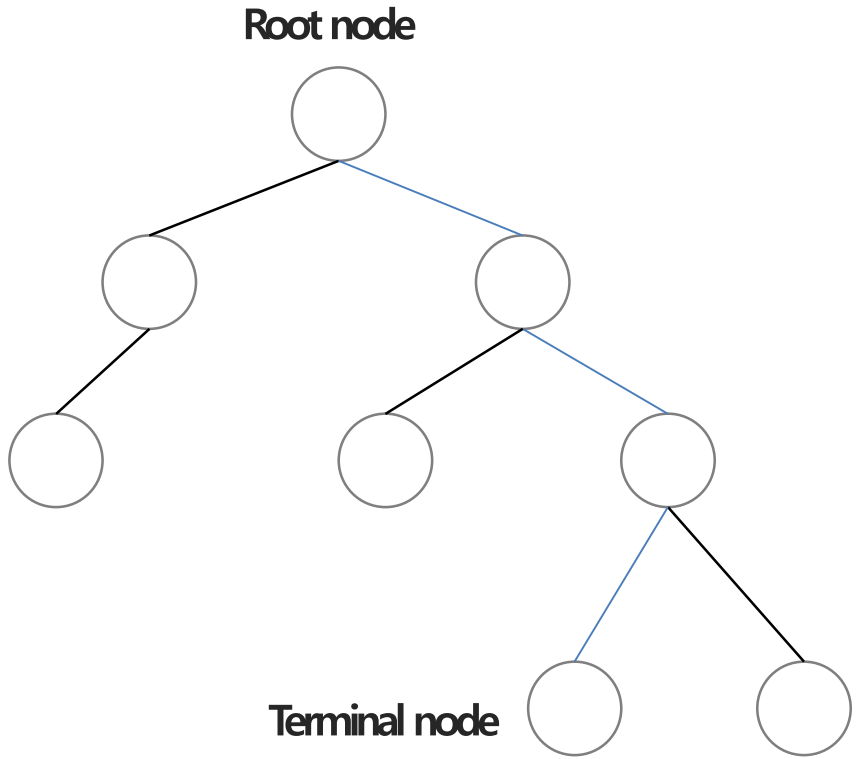
❖ Anomaly score for isolation forest

- 이론적인 평균 path length c(n)을 통해 관측치 x에서 실제로 고립까지 걸린 평균 path length h(x)를 정규화
- Path length : root node에서 terminal 노드까지 거쳐간 node의 수 = split 횟수
- Tree에서 path length가 짧을수록 이상치 스코어는 1에 가까워짐

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n}$$

$$H(n) = \ln(n) + 0.57721(Euler's\ constant)$$

$$anomaly\ score\ s(x,n) = 2^{-\frac{E(h(x))}{c(n)}} \begin{cases} E(h(x)) \to 0, s \to 1 \\ E(h(x)) \to c(n), s \to 0.5 \\ E(h(x)) \to n-1, s \to 0 \end{cases}$$



감사합니다