
인공지능보안

CICIDS 데이터 셋

곽상열(M2021520)

문성현(M2021522)

Kookmin University

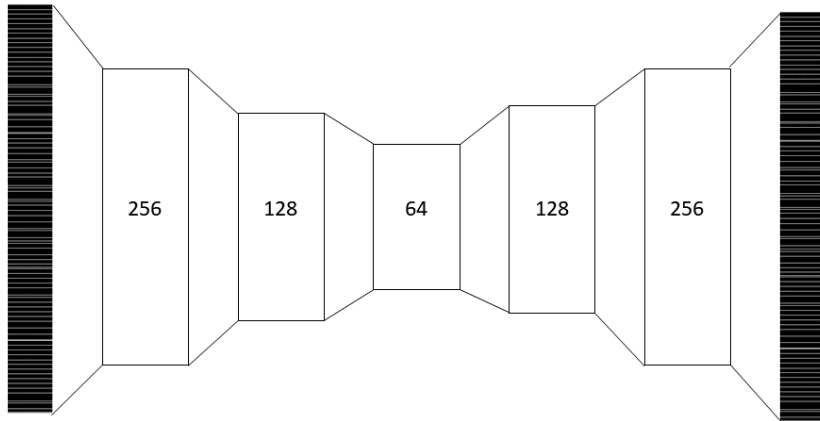
Anomaly Detection

- **Anomaly Detection**

- 오토인코더를 이용한 netflow 이상탐지
- 오토인코더를 사용해서 netflow를 분석해 오토인코더를 사용한 Anomaly Detection model이 어떤 공격을 잘 감지하는지 혹은 어떤 공격에 취약한지 분석

Overview

- Overview



- AutoEncoder

- 정상 데이터로 학습된 오토인코더가 input으로 비정상 데이터를 받았을 때 복원력이 떨어진다는 점을 이용

Data Preprocessing

• Data Preprocessing

	ts	te	td	sa	da	sp	dp	pr	ipkt	ibyt	opkt	obyte	Label
0	2017-07-03 1:00:01	2017-07-03 1:00:01	83823	192.168.10.25	23.194.182.12	59385	443	6	5	479.0	10	713.0	BENIGN
1	2017-07-03 1:00:01	2017-07-03 1:00:01	3	23.194.182.12	192.168.10.25	443	59385	6	0	0.0	2	37.0	BENIGN
2	2017-07-03 1:00:01	2017-07-03 1:01:59	118699862	8.6.0.1	8.0.6.4	0	0	0	0	0.0	76	0.0	BENIGN
3	2017-07-03 1:00:01	2017-07-03 1:00:01	86739	192.168.10.3	192.168.10.1	61098	53	17	1	67.0	1	51.0	BENIGN
4	2017-07-03 1:00:01	2017-07-03 1:01:41	100934792	192.168.10.25	192.168.10.3	51147	53	17	4	684.0	4	142.0	BENIGN

Data Preprocessing

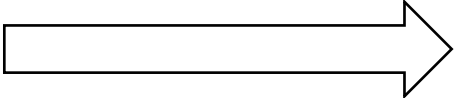
· 세션 시작시간/ 종료시간

2017-07-03 1:00:01
2017-07-03 1:01:22
2017-07-03 1:01:53
⋮
⋮
2017-07-03 1:02:01

- 날짜 부분은 중요하지 않다고 생각하여 제외
- 각 자리별로 0~9 사이로 원-핫 인코딩

Data Preprocessing

· 출발지/도착지 Ip

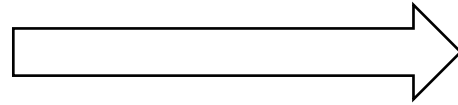
192.168.10.25		192 168 010 025
23.194.182.12		023 194 182 012
192.168.10.3		192 168 010 003
8.6.0.1		008 006 000 001
:		:

- 고정된 크기의 input을 만들기 위해 0을 패딩
- 패딩 된 결과값에 대해서 원-핫 인코딩

Data Preprocessing

· 출발지/도착지 port

59385
443
0
61098
⋮

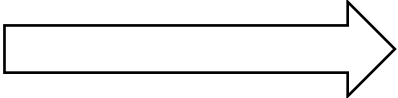


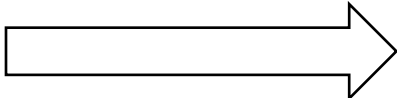
59385
00443
00000
61098
⋮

- 가장 큰 수의 자리 수에 맞춰 패딩
- 패딩된 값에 대해 원-핫 인코딩

Data Preprocessing

· 출발지 도착지 패킷과 바이트 수

5		000005
0		000000
17		000017
629		000629
⋮		⋮

479.0		000000479
0.0		000000000
67.0		000000067
684.0		000000684
⋮		⋮

- 가장 큰 수의 자리 수에 맞춰 패딩
- 패딩된 결과값에 대해서 원-핫 인코딩

Data Preprocessing

- **Protocol**

- Protocol 값의 종류가 3 종류이므로 3개의 클래스로 원-핫 인코딩
- TCP: 6 UDP: 17 ICMP: 1

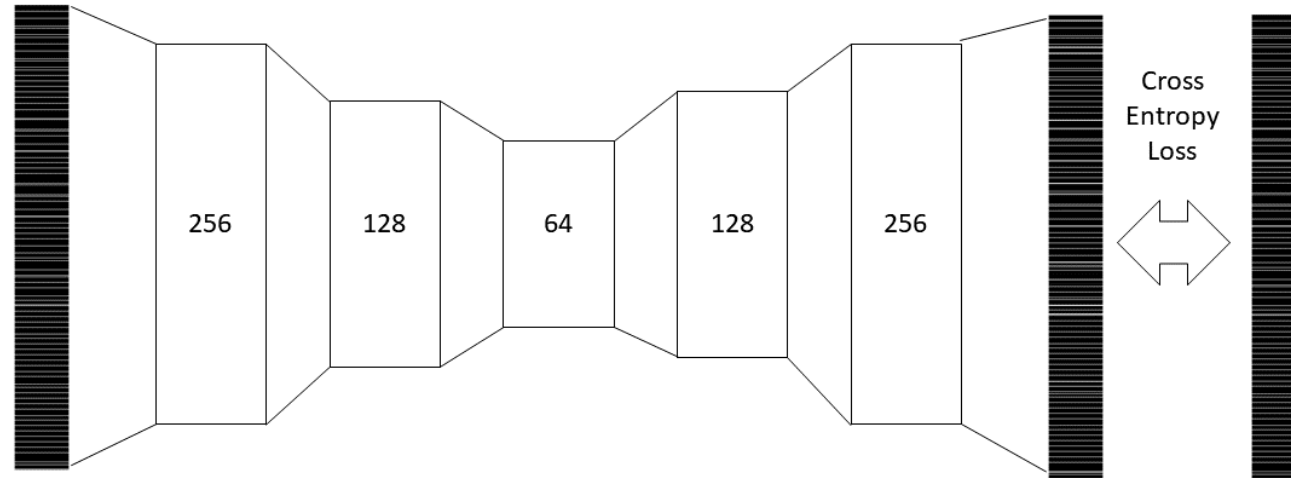
Data Preprocessing

- **Label**

- "BENIGN"인지 아닌지를 판별하여, 정상/비정상 2개의 클래스로 원-핫 인코딩

Network

- **AutoEncoder**



Training

- **Training**
 - Adam Optimizer
 - 10 epochs
 - 79.8% Accuracy

Testing

- **Testing**

- Accuracy 75.07%
- Accuracy 81.24% (DoS Hulk, PortScan 제거)

BENIGN	7617
DoS Hulk	1036
PortScan	664
DDoS	549
DoS GoldenEye	40
FTP-Patator	24
DoS slowloris	22
SSH-Patator	19
DoS Slowhttptest	18
Bot	6
Web Attack-Brute Force	3
Web Attack-XSS	2

테스트 데이터의 포함된 label

BENIGN	1452
PortScan	631
DoS Hulk	311
DDoS	65
FTP-Patator	12
SSH-Patator	9
DoS slowloris	6
Web Attack-Brute Force	3
Bot	2
DoS GoldenEye	1
Web Attack-XSS	1

예측 실패한 label