JKIICE Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지 Vol. 26, No. 12: 1872~1879, Dec. 2022

가중치 VAE 오버샘플링(W-VAE)을 이용한 보안데이터셋 샘플링 기법 연구

강한바다[1] · 이재우[2*]

A Data Sampling Technique for Secure Dataset Using Weight VAE Oversampling(W-VAE)

Hanbada Kang[1] · Jaewoo Lee[2*]

[1]Graduate Student, Department of Convergence Security, Chung-Ang University, Seoul, 06974 Korea
[2*]Assistant Professor, Department of Industrial Security, Chung-Ang University, Seoul, 06974 Korea

요 약

최근 인공지능 기술이 발전하면서 해킹 공격을 탐지하기 위해 인공지능을 이용하려는 연구가 활발히 진행되고 있다. 하지만, 인공지능 모델 개발에 핵심인 학습데이터를 구성하는데 있어서 보안데이터가 대표적인 불균형 데이터라는 점이 큰 장애물로 인식되고 있다. 이에 본 눈문에서는 오버샘플링을 위한 데이터 추출에 딥러닝 생성 모델인 VAE를 적용하고 K-NN을 이용한 가중치 계산을 통해 클래스별 오버샘플링 개수를 설정하여 샘플링을 하는 W-VAE 오버샘플링 기법을 제안한다. 본 논문에서는 공개 네트워크 보안 데이터셋인 NSL-KDD를 통해 ROS, SMOTE, ADASYN 등 총 5가지 오버샘플링 기법을 적용하였으며 본 논문에서 제안한 오버샘플링 기법이 F1-Score 평가지표를 통해 기존 오버샘플링 기법과 비교하여 가장 효과적인 샘플링 기법임을 증명하였다.

ABSTRACT

Recently, with the development of artificial intelligence technology, research to use artificial intelligence to detect hacking attacks is being actively conducted. However, the fact that security data is a representative imbalanced data is recognized as a major obstacle in composing the learning data, which is the key to the development of artificial intelligence models. Therefore, in this paper, we propose a W-VAE oversampling technique that applies VAE, a deep learning generation model, to data extraction for oversampling, and sets the number of oversampling for each class through weight calculation using K-NN for sampling. In this paper, a total of five oversampling techniques such as ROS, SMOTE, and ADASYN were applied through NSL-KDD, an open network security dataset. The oversampling method proposed in this paper proved to be the most effective sampling method compared to the existing oversampling method through the F1-Score evaluation index.

키워드 : 인공지능, 정보보안, 오버샘플링, VAE

Keywords : AI, Information Security, Over Sampling, VAE

Received 20 October 2022, **Revised** 5 November 2022, **Accepted** 9 November 2022

*** Corresponding Author** Jaewoo Lee(E-mail:jaewoolee@cau.ac.kr, Tel:+82-2-820-5935)
Assistant Professor, Department of Industrial Security, Chung-Ang University, Seoul, 06974 Korea

Open Access http://doi.org/10.6109/jkiice.2022.26.12.1872 print ISSN: 2234-4772 online ISSN: 2288-4165

Ⅰ. 서 론

최근 인공지능 기술의 확산으로 사이버 공격자들 역시 인공지능 기술 활용을 통한 공격 효율성 증대나 새로운 형태의 공격 방식을 시도하고 있다. 그에 따라 보안업계 또한 인공지능 기술을 이용하여 해커의 공격을 탐지하고 차단하려는 시도들이 이어지고 있다. 하지만 보안데이터는 데이터가 특정 클래스에 편향되어있는 대표적인 불균형 데이터로 이러한 편향된 보안데이터의 특성은 인공지능 활용에 있어서 큰 장애물로 인식되고 있다.

불균형 데이터는 다른 클래스와 비교하여 상대적으로 많은 데이터가 존재하는 다수 클래스(majority class)와 적은 데이터가 존재하는 소수 클래스(minority class)로 구분할 수 있다. 불균형 데이터로 인공지능 모델을 학습시키면 데이터양이 많은 클래스로 편향되어 학습되기 때문에 학습양이 적은 클래스에 대해서 낮은 분류 성능을 보이는 문제가 발생한다[1].

이러한 문제를 해결하기 위해서는 데이터 클래스 간 균형을 맞춰주는 과정이 필요한데 이러한 작업을 데이터 샘플링(Data Sampling)이라 한다. 데이터 샘플링은 다수 클래스와 소수 클래스 중 어느 클래스의 데이터 샘플 수를 조정하느냐에 따라 언더 샘플링(Under-sampling) 기법과 오버 샘플링(Over-sampling) 기법으로 나누어진다[2].

언더 샘플링은 다수의 클래스에서 일부 데이터를 추출하고 소수의 클래스의 데이터는 모두 추출하는 샘플링 방식이다. 언더 샘플링은 대량의 데이터를 다루어야 할 때 전체 데이터 크기를 줄일 수 있기 때문에 자원 활용에 효율적이라는 장점이 있다. 그러나 데이터를 추출하는 과정에서 데이터 유실이 발생할 수 있다는 점에서 단점이 있다[3]. 언더 샘플링 기법으로는 RUS(Random Under-sampling)과 ENN(Edited Nearest Neighbor), Tomek Links 등이 있다[4].

오버 샘플링은 소수의 클래스에 데이터를 추가하여 소수 클래스와 다수 클래스간의 데이터 불균형을 줄이는 샘플링 방식이다. 오버 샘플링은 언더 샘플링과 반대로 샘플링 과정에서 데이터의 손실이 없다는 장점이 있다. 오버 샘플링의 단점으로는 랜덤 오버샘플링 같은 데이터 복제 방식의 경우 같은 데이터를 반복 학습하기 때문에 분류 모델의 학습 시 과적합(Overfitting)이 발생할 수 있다. 또한 SMOTE 같은 샘플링 방법의 경우 K-NN (k-nearest neighbors) 알고리즘을 사용하여 소수클래스와 다수클래스 사이의 샘플을 생성하기 때문에 다수 클래스 주변에 생성된 데이터가 인공지능 학습 시 노이즈로 작용해 모델 성능을 떨어뜨린다[2]. 이러한 오버 샘플링 기법으로는 ROS(Random Over-sampling), SMOTE (Synthetic Minority Over-sampling Technique)[5], ADASYN (Adaptive synthetic sampling)[6] 등이 있다[4].

본 눈문에서는 보안데이터의 불균형 문제를 해결하기 위해 ADASYN에서 사용한 기법인 데이터 클래스의 밀도를 가중치로 두어 샘플링 개수를 산출하고 VAE (Variational Auto Encoding)를 기반으로 데이터의 특성을 학습하여 추가데이터를 생성하는 W-VAE 오버샘플링 기법을 제시하였다.

본 논문의 공헌은 다음과 같다.

- 이미지 데이터 분야에 사용되는 VAE 알고리즘이 보안데이터 오버샘플링에 효과적임을 증명하였다.
- 클래스마다 가중치를 주어 단순 VAE 오버샘플링보다 인공지능 분류 성능을 향상시켰다.

본 논문의 구성은 다음과 같다. 2장에서는 오버샘플링 및 VAE 관련 연구를 기술하고, 3장에서는 본 눈문에서 제안하는 오버 샘플링 기법을 설명한다. 4장에서는 제안한 모델을 기존 샘플링 기법들과 비교하기 위한 실험 방법을 설명하고 5장에서 실험 결과를 기술한다. 6장에서 결론에 대해 요약하고 향후 연구 방향을 제시한다.

Ⅱ. 관련 연구 및 배경

2.1. 오버샘플링

오버샘플링 기법은 다양한 방식으로 연구되어 왔다. Chawla 등은[5] K-NN 알고리즘을 이용하여 소수 클래스에서 가장 가까운 K개의 이웃을 선택한 후 필요한 오버샘플링 양에 따라 각각의 방향으로 하나의 샘플을 생성하는 SMOTE 기법을 제안하였다. Haibo 등은[6] 기존의 SMOTE 기법에서 소수 클래스의 밀도를 기준으로 소수 클래스 중 학습하기 어려운 클래스를 구분하여 해당 클래스에 초점을 맞추어 샘플링하는 방식인 ADASYN 기법을 제안하였다.

이강혁은[7] 가우시안 혼합 모델을 이용하여 소수 클래스를 군집화한 후 SMOTE, ADASYN 등의 오버샘플

링 기법을 이용하여 소수 클래스의 데이터를 증가시키는 방법을 제시하였고 자동차, 전복 등 그림 데이터와 의료 데이터인 저밀도 지단백-콜레스테롤(LDL-Cholesterol) 데이터셋에 적용하여 효과를 검증하였다. 최윤희 등은[8] 생성적 적대 신경망(GAN)을 기반으로한 CTGAN 모델을 이용한 오버샘플링 기법을 제안하였고 보안데이터인 CICIDS2017 데이터셋에 해당 기법을 적용하여 효과를 검증하였다. 유승태 등은[9] CNN과 GAN을 결합한 DCGAN을 이용한 오버샘플링 기법을 제시하였고 보안데이터인 NGIDS-DS를 이용하여 효과를 검증하였다.

2.2. VAE(Variational Auto Encoding)

VAE는 GAN과 같이 대표적인 딥러닝 생성모델(Generativ Model) 중 하나이다. Kingma 등은[10] encoder와 decoder 구조를 가진 딥러닝 모델인 오토인코더(Auto Encoder)에 베이지안(Bayesian) 추론 개념을 더하여 VAE 모델을 제시하였으며 손글씨 숫자(MNIST)와 사람 얼굴(Frey Face)의 그림 데이터셋을 학습시켜 해당 모델이 입력 값의 정보를 잘 학습하였음을 검증하였다.

VAE는 Encoder를 통해서 입력 값(X)으로부터 잠재변수 확률 분포의 모수(μ, σ)를 추정하여 평균(μ), 표준편차(σ)로부터 임의의 값을 샘플링하고 이를 decoder에 투입하여 입력값을 복원하는 딥러닝 생성모델이다 [11].

VAE의 손실함수(Loss Function)는 다음과 같다.

$$L = E_{q\Phi(z \mid x)}[\log p_\theta(x \mid z)] - D_{KL}(q_\Phi(z \mid x) \parallel p_\theta(z)) \quad (1)$$

입력값 X, 잠재 변수 Z라고 정의할 때 식 (1) 우변의 첫 번째 항은 입력 값과 출력값의 차이인 복원 오차(Reconstruction error)을 의미한다. 파라미터 Φ를 갖는 encoder로부터 Z의 확률분포를 만들고, 이로부터 Z를 샘플링하여 파라미터 Θ를 갖는 decoder를 통해 X를 복원하여 손실을 계산하게 된다. 식 (1) 우변의 두 번째 항은 정규화 오차(Regularization error) 의미하며 이상적인 샘플링 확률분포 q(z|x)와 샘플링 확률분포 p(z)의 차이를 나타낸다.

Kihyuk 등은[12] VAE에 레이블을 추가하여 지도학습이 가능한 CVAE(Conditional VAE)를 제안하였으며 글자, 새, 사람 얼굴의 그림 데이터셋을 대상으로 모델의 성능을 검증하였다. Furkan 등은[13] VAE에 측정값의 확률 분포의 가능도(Likelihood)와 두 분포의 차이(KL-Divergence)를 조절하는 하이퍼파라미터(β) 개념을 더해 β-VAE를 제시하였고 납땜(Solder) 그림 데이터셋을 이용하여 이상 탐지 성능을 검증하였다.

2.3. VAE 오버샘플링

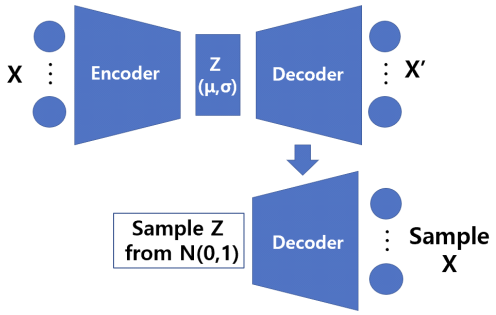

Fig. 1 VAE Oversampling

VAE 학습 후 표준정규분포에서 Z를 샘플링하고 해당 값을 Decoder에 넣으면 소수데이터에 추가할 샘플링 데이터를 생성 할 수 있다.

박종혁은[11] VAE를 불균형 데이터의 오버샘플링 기법으로 제안하였고 패션 스타일 그림 데이터셋을 이용하여 성능을 검증하였다. Huang 등은[14] 신용카드 사기 데이터의 불균형을 해결하기 위해 VAE 오버샘플링 기법을 제안하였고 GAN을 이용한 오버샘플링 대비 VAE가 우수함을 검증하였다. 해당 논문에서는 GAN과 비교하여 VAE를 오버샘플링에 이용 시 첫째, 많은 다양성을 가진 샘플데이터를 생성할 수 있으며 둘째, 인코더 및 디코더를 통해 데이터를 생성하기 편리하며 셋째, 텍스트 데이터를 생성하는데 제한이 덜한 장점이 있다고 하였다.

Ⅲ. W-VAE 오버샘플링 기법

3.1. Overview

본 논문에서 제안하는 샘플링 방법은 소수 클래스 간 밀도를 가중치로 VAE를 오버샘플링에 이용하는 가중치 VAE 오버샘플링 기법(W-VAE)이다.

먼저 K-NN을 이용하여 계산한 클래스 간의 밀도를 가중치로 두어 샘플링할 데이터의 개수 N를 정한다. 그

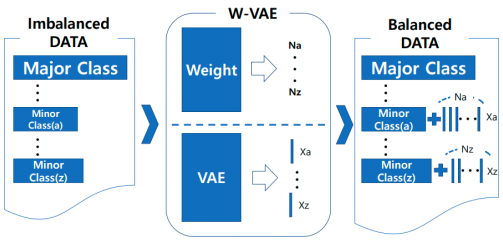

Fig. 2 W-VAE Process

후에 VAE를 이용하여 소수 클래스 데이터의 특징을 학습시키고 소수클래스 당 N개의 샘플링데이터 X를 추출하여 기존 데이터와 결합한다.

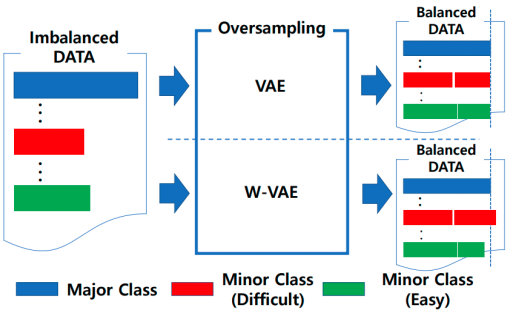

Fig. 3 W-VAE vs VAE

W-VAE 오버샘플링 기법은 기존 VAE를 이용한 오버샘플링 기법과 비교하여 클래스 간 밀도를 기준으로 구분하기 어려운 소수 클래스의 데이터를 더욱 많이 추출하고 구분하기 쉬운 데이터를 적게 추출하여 클래스마다 학습데이터 수가 다른 차이가 있다.

그 결과 구분하기 어려운 데이터를 인공지능이 보다 많이 학습하고 쉬운 데이터는 더 적게 학습하여 인공지능이 데이터를 분류하는데 이점이 있다.

3.2. 가중치 계산

만약 소수 클래스 중 다수 클래스와 경계값이 모호한 클래스가 있다면 해당 클래스는 다른 클래스에 비해 인공지능 모델이 분류하는데 어려움이 따른다.[6] 따라서 해당 클래스에 가중치를 두어 더 많은 샘플을 생성할 필요가 있다. 본 논문에서는 클래스의 가중치 계산을 위해 ADASYN에서 사용한 K-NN 알고리즘 계산 방식을 이용했으며 계산 방법은 그림 4와 같다.

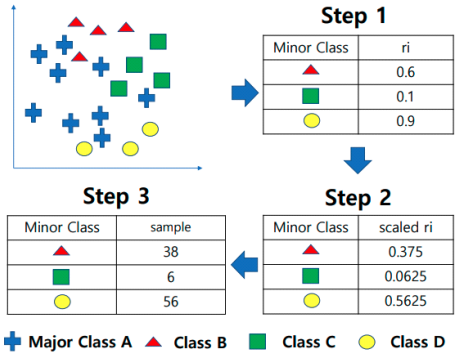

Fig. 4 Weight Calculation Process

$$Sample = (MJ_n - MN_n) \times \frac{r_i}{\sum_{i=1}^{m} r_i} \quad (2)$$

첫째, 소수 클래스를 기준으로 가까운 K개의 주변 데이터를 선택한다. 그리고 그중 다수 클래스의 수를 M이라고 할때 그 비율을 r_i=M/K라고 정의한다. 둘째, 모든 r_i 값을 [0,1]로 정규화한다. 셋째, 정규화된 r_i를 다수 클래스(MJ)와 소수 클래스(MN) 간의 데이터 개수 차이와 곱하고, 반올림한다. 이렇게 계산된 최종값은 오버샘플링 될 데이터의 개수가 된다. 이러한 과정을 식으로 나타내면 식 (2)와 같다. 이때 m은 클래스 개수이다.

소수 클래스 데이터 주변의 다수 클래스의 밀도(r_i)를 가중치 W라 했을 때 W가 높을수록 소수 클래스와 다수 클래스의 경계 값이 모호하다. 따라서 본 논문에서는 경계값이 모호하면 클래스간 판별의 불활실성 정도가 크다고 가정하여 W가 클수록 샘플링 데이터를 많이 생성하고 적을수록 불필요한 데이터를 적게 생성하여 인공지능 분류에 도움을 주고자 한다.

Ⅳ. 실험설계

4.1. 데이터셋

본 논문에서는 대표적인 보안데이터인 NSL-KDD를 검증데이터로 사용하였다. NSL-KDD 데이터 세트는 1999년 International Knowledge Discovery and Data Minig Tools 경진대회에서 사용된 KDD CUP99 데이터 세트의 단점을 개선한 데이터이다[15].

해당 데이터는 정상 포함 총 23개 유형의 정상 및 공

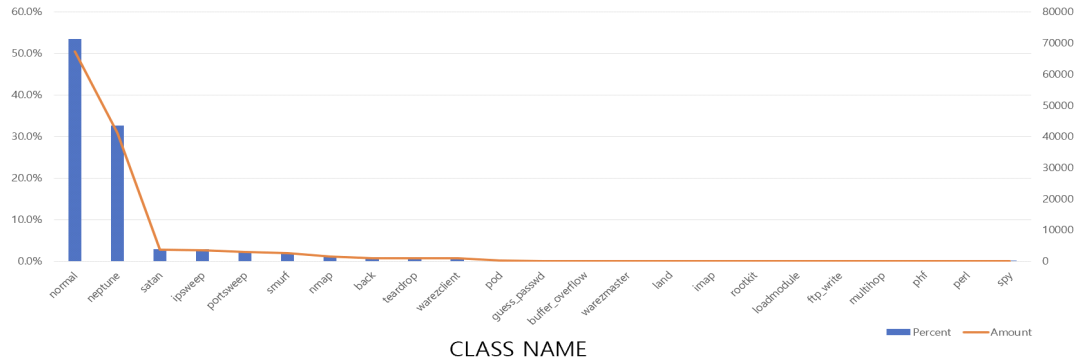

Fig. 5 Percentage, Amount of each class of data

격 네트워크 트래픽으로 구성되어 있는 보안데이터셋이며 학습데이터와 테스트데이터가 분리되어 구성되어 있다[16]. 또한 다양한 연구에서 해당 데이터를 실험 검증데이터로 사용하여 우수함을 입증하였다[17, 18, 19]. 본 논문에서는 학습데이터로 인공모델을 학습시킨 후 테스트데이터로 모델 성능을 검증하였다.

4.2. 데이터 분석

NSL-KDD의 학습데이터는 총 22개의 공격 클래스 데이터를 포함하고 있으며 각 클래스의 개수와 비율은 그림 5와 같다. 가장 많은 비중을 차지하고 있는 클래스는 정상 네트워크 데이터인 "Normal"로 전체 데이터의 약 53.5%를 차지하고 있다. 공격클래스 중에는 "Neptune"가 32.7%로 가장 많은 비중을 차지하고 나머지 공격 클래스는 3%미만의 비중을 차지한다.

공격데이터와 정상데이터 기준으로 보면 46.5%, 53.5%로 균형적인 데이터로 볼 수 있으나 공격 데이터 기준으로 보면 "Neptune"에만 편중되어 있는 매우 불균형한 데이터로 판단할 수 있다.

인공지능에서 모델이 분류해야 할 클래스 개수가 너무 많으면 인공지능 분류 성능이 떨어지고 소수클래스의 특징을 학습하기 위한 최소한의 데이터를 확보하기 위해 본 논문에서는 전체 데이터에서 1% 이상 비중을 가지고 있는 상위 클래스 7개를 선정하였다.(표 1)

Table. 1 Top7 Data Class Ratio

	amount	ratio
normal	67,343	53.5%
neptune	41,214	32.7%
satan	3,633	2.9%
ipsweep	3,599	2.9%
portsweep	2,931	2.3%
smurf	2,646	2.1%
nmap	1,493	1.2%

4.3. 데이터 전처리

NSL-KDD는 총 41개의 특징(Feature)를 가지고 있다. 그 중 "protocol_type", "service", "flag"는 문자형 Categorical 값을 가지고 있기 때문에 OneHot Encoding을 진행하였고 총 41개의 특징에서 총 122개의 특징으로 전환되었다. 또한 VAE를 통한 학습을 진행하기 위해 각 특징의 변수들을 표준정규분포(z-값 분포)로 정규화하였다.

4.4. W-VAE 모델 설계

본 논문에서 설계한 VAE는 그림 6과 같다. 입력 값인 특징 122개의 값은 Encoder의 은닉층(Hidden Layer)을 거쳐 20개의 Latent Dim 값으로 축소되고 다시 Decoder

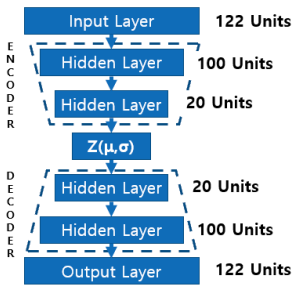

Fig. 6 Flow Chart of the VAE

의 은닉층을 거쳐 122개의 출력값을 만든다. 여기서 입력 값과 최대한 비슷한 출력값을 출력하는 Latent Dim 값의 평균 μ, 표준편차 σ를 추출한다. 여기서 학습된 Decoder에 표준정규분포표에서 추출한 임의값을 입력하여 샘플데이터를 만든다.

최적화 함수(Optimizer)는 Adam, 손실 함수(Loss Function)는 Negative ELBO를 사용하였으며 배치사이즈 32, 활성화함수(Activaton) Relu, epoch 50로 설정하였다. 소수 클래스의 가중치 계산을 위하여 Python의 Imblearn ADASYN 라이브러리를 이용하였으며 Random _state=0, K-NN의 K값은 기본값인 5로 설정하였다.

4.5. 평가지표

본 논문의 평가 지표로는 인공지능에서 일반적으로 사용되는 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1-score를 사용하였다.

인공지능이 정상과 공격을 분류한다고 했을 때 정확도는 인공지능이 예측한 전체 데이터 중 정상과 공격이 올바르게 분류된 데이터의 비율을 말한다. 정밀도는 인공지능이 공격이라고 예측한 데이터 중 실제 공격이라고 분류한 비율을 말한다. 재현율은 실제 공격 중 인공지능이 공격이라고 예측한 비율을 말한다. F1-score는 정밀도와 재현율의 조화평균을 말하며 일반적으로 모델의 성능을 비교할 때 자주 사용되는 지표이다.

Ⅴ. 실험결과

본 논문은 표 2와 같은 실험 환경에서 수행되었다. macOS에서 진행되었으며 Jupyter Notebook 6.4.5에서 환경에서 수행되었다.

Table. 2 Experiment Environment

OS	macOS Monterey
CPU	Apple M1
RAM	16GB
Language	python 3.9.7
Library	tensorflow 2.9.1, keras 2.9.0

데이터의 가중치 계산 결과 표 3과 같이 클래스의 가중치에 따른 오버샘플링 수가 도출되었다. 다수 클래스인 “normal”을 기준으로 볼 때 “satan” 클래스가 가중치가 가장 높으며 “portsweep”클래스가 가장 가중치가 낮은 것을 확인할 수 있다.

각 소수 클래스를 VAE에 학습시키고 학습된 VAE의 Decoder에 표 3에서 도출된 필요 오버샘플링 수만큼 표준정규분포 내 임의 수를 입력하여 샘플링을 수행하였다. ROS, SMOTE, VAE는 각 클래스의 수가 동일하며 ADASYN, W-VAE는 가중치에 의해 각 클래스의 수가 다르게 샘플링되었다.

오버샘플링이 수행된 데이터는 LightGBM 알고리즘을 이용하여 인공지능 학습을 진행하였고 LightGBM 설정값은 부스팅값 gbdt(Gradient Boosting Decision Tree), 최대 트리수 75, 최대 리프수 31, Random state 1337로 정하였으며 그 결과 표 4, 표 5와 같은 결과를 도출하였다.

표 4는 오버샘플링 기법 기준 4가지 평가지표를 통해 모델 성능을 비교한 결과이며 실험결과, F1-Score 기준 ROS < SMOTE < ADASYN < VAE < W-VAE 순으로 좋은 결과를 나타냈다. 본 논문에서 제안한 W-VAE 오버샘플링 기법이 F1-Score 96.23%로 타 오버샘플링 기법과 비교하여 우수한 오버샘플링 기법으로 확인되었다.

Table. 3 Weight sampling

	amount	Weight
normal	67,343	1
neptune	67,306	0.9985839
satan	67,349	1.0000942
ipsweep	67,334	0.9998588
portsweep	67,274	0.9989288
smurf	67,348	1.0000773
nmap	67,308	0.9994685

표 5는 데이터 클래스 기준 오버샘플링 기법 별 F1-Score를 비교해 본 결과이다. 오버샘플링이 적용되지 않는 원본데이터와 비교하여 W-VAE를 적용한 후 “neptune” 2.64%, “satan” 9.31%, “ipsweep” 1.04%, “portsweep” 15.39%, “nmap” 1.35% 상승한 것을 볼 수 있는데 기존 데이터에서 가장 분류성능이 낮게 나왔던 소수 클래스 “satan”, “portsweep”가 가장 많은 인공지능 분류 성능 상승을 보였다.

또한 단순히 VAE만을 적용한 기법과 본 논문에서 제안한 W-VAE 기법을 비교해 보아도 “satan”, “portsweep”가 각각 0.22%, 1.71%로 가장 많이 상승했는데 가중치

Table. 4 Comparison of oversampling technique performance(Algorithm : LightGBM)

	F1	Accuracy	Precision	Recall
Origin	92.52%	96.77%	87.94%	97.61%
ROS	94.01%	97.56%	90.34%	97.99%
SMOTE	94.43%	97.99%	90.93%	98.21%
ADASYN	94.04%	97.90%	91.65%	96.55%
VAE	95.95%	98.31%	93.16%	98.90%
W-VAE	96.23%	98.35%	93.42%	99.21%

Table. 5 Comparison of oversampling technique performance by Data Class(Evaluation Method : F1-Score)

	Origin	ROS	SMOTE	ADASYN	VAE	W-VAE
normal	98.64%	98.69%	98.67%	98.41%	98.68%	98.65%
neptune	97.23%	98.46%	97.84%	99.78%	99.80%	99.87%
satan	79.24%	83.84%	81.48%	89.38%	88.33%	88.55%
ipsweep	97.54%	96.84%	97.19%	95.50%	98.58%	98.58%
portsweep	73.18%	81.32%	77.04%	82.05%	86.86%	88.57%
smurf	97.79%	97.49%	97.64%	92.25%	97.79%	97.79%
nmap	98.65%	98.65%	98.65%	99.32%	100.00%	100.00%

에 의한 샘플링 수 변동량이 많을수록 좋은 성능을 나타냈다. 이는 가중치 적용을 통해 클래스의 밀도에 맞게 샘플링이 진행되어 성능 향상에 도움을 주었음을 증명한다.

Ⅵ. 결 론

본 논문에서는 인공지능 분야에서 불균형한 보안데이터의 분류 성능을 향상시키기 위해 이미지 데이터 분야에 일반적으로 사용된 VAE 알고리즘과 K-NN을 이용한 가중치 기법을 결합한 W-VAE 오버샘플링 기법을 제안하였다.

또한 제안한 기법을 검증하기 위해 불균형 보안데이터인 NSL-KDD에 제안한 W-VAE 오버샘플링 기법을 적용하여 ROS, SMOTE 등 기존 다른 샘플링 기법보다 제안한 오버샘플링 기법이 가장 효과적인 샘플링 기법임을 검증하였다.

보안데이터는 현실적으로 정상적인 데이터는 얻기 쉬운 반면 공격데이터는 얻기 어려워 인공지능 학습 데이터셋 구성 시 불균형 데이터를 초래할 수밖에 없다. 따라서 딥러닝 생성 모델인 VAE 및 가중치를 적용하여 균형데이터를 만듦으로써 인공지능 분류모델의 성능을 보다 효과적으로 상승시킬 수 있음을 증명했다는 점에서 본 실험은 의미가 있다.

하지만, CVAE, β-VAE 등 기존 VAE를 개량한 알고리즘이 연구되고 있고 또 다른 생성 모델인 GAN을 이용한 오버샘플링에 대한 연구도 많이 이루어지고 있기 때문에 추후 연구에서는 이런 다양한 생성 모델에 K-NN을 이용한 가중치 기법을 적용하여 성능비교를 진행할 예정이다. 또한 본 논문에서는 K-NN의 K값과 VAE의 Latent Dim 값을 임의로 5와 20으로 고정하여 성능비교를 진행하였으나 차후에는 해당 값의 조정에 따른 오버샘플링 성능비교를 진행하고자 한다. 더 나아가 실무에 적용하기 위해 실제 보안데이터를 통한 효과성 검증 연구 또한 필요하다.

REFERENCES

[1] S. H Seo, Y. J. Jeon, J. S. Lee, H. J. Jung, and J. T. Kim, “An Over-sampling Method based on Generative Adversarial Networks for Effective Classification of Imbalanced Big Data,” in *Proceedings of Korea Software Congress 2017*, Busan, Korea, pp. 1030-1032, 2017.

[2] M. J. Son, S. W. Jung, and E. J. Hwang, “A Deep Learning Based Over-Sampling Scheme for Imbalanced Data

Classification," *KIPS Transactions on Software and Data Engineering*, vol. 8, no. 7, pp. 311-136, Jul. 2019.

[3] J. H. Yang, "Comparison of the Classification Algorithms Using a Sampling Technique in Imbalanced Data," M. S. thesis, Dongguk University, Korea, 2017.

[4] I. O. Jung, J. W. Ji, G. H. Lee, and M. J. Kim, "A study on intrusion detection performance improvement through imbalanced data processing," *Jouranl of Information and Security*, vol. 21, no. 3, pp. 57-66, Sep. 2021.

[5] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, Jun. 2002.

[6] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning," in *Proceedings of IEEE International Joint Conference on Neural Networks*, Hong Kong, pp.1322-1328, 2008.

[7] K. Lee, "Oversampling based on Gaussian Mixture Model for Imbalanced data classification," M. S. thesis, Hanyang University, Korea, 2019.

[8] Y. H. Choe and K. W. Oh, "A Study on the Introduction of CTGAN Oversampling Algorithm to improve Imbalance Problem in Intrusion Detection Data," *The Journal of Korean Institute of Communications and Information Sciences*, vol. 45, no. 12, pp. 2114-2122, Dec. 2020.

[9] S. T. Yoo and K. S. Kim., "Comparison of Anomaly Detection Performance Based on GRU Model Applying Various Data Preprocessing Techniques and Data Oversampling," *Journal of the Korea Institute of Information Security & Cryptology*, vol. 32, no. 2, pp. 201-211, Apr. 2022.

[10] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," *arXiv:1312.6114v10*, 2013.

[11] J. H. Park, "Improving Fashion Style Classification Accuracy using VAE in Class Imbalance Problem," *The Journal of Korean Institute of Information Technology*, vol. 19, no. 2, pp. 1-10, Feb. 2021.

[12] K. Sohn, H. Lee, and X. Yan, "Learning Structured Output Representation using Deep Conditional Generative Models," in *Proceedings of Advances in neural information processing systems (NeurIPS)*, Montreal: QC, Canada, pp. 3483-3491, 2015.

[13] F. Ulger, S. E. Yuksel, and A. Yilmaz, "Anomaly Detection for Solder Joints Using β-VAE," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 11, no. 12, pp. 2214-2221, Oct. 2021.

[14] H. Tingfei, C. Guangquan, and H. Kuihua, "Using Variational Auto Encoding in Credit Card Fraud Detection," *IEEE Access*, vol. 8, pp. 149841-149853, Aug. 2020.

[15] S. C. Hsiao, D. Y. Kao, Z. Y. Liu, and R. Tso, "Malware Image Classification Using One-Shot Learning with Siamese Networks," in *Procedia Computer Science*, Budapest, Hungary, vol. 159, pp. 1863-1871, 2019.

[16] University of new brunswick, *NSK-KDD dataset* [Online]. Available: https://www.unb.ca/cic/datasets/nsl.html.

[17] P. Devan and N. Khare, "An efficient XGBoost-DNN-based classification model for network intrusion detection system," *Neural Computing and Applications*, vol. 32, pp. 12499-12514, Jan. 2020.

[18] C. Yin, Y. Zhu, J. Fei and X. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," *IEEE Access*, vol. 5, pp. 21954-21961, Oct. 2017.

[19] K. J. Ryu, "Study for Solving Network Traffic Data Imbalance And Rare Class Problems Using a Similarity Neural Network," M. S. thesis, Sejong University, Korea, 2021.

강한바다(Hanbada Kang)

2013년 3월 한양대학교 도시공학과 공학학사
2019년 8월 ~ 현재 중앙대학교 융합보안학과 석사과정
※관심분야 : 인공지능, 딥러닝, 정보보안

이재우(Jaewoo Lee)

2006년 2월 서울대학교 컴퓨터공학부 학사
2008년 2월 서울대학교 컴퓨터공학부 석사
2017년 8월 University of Pennsylvania, Ph.D in Computer and Information Science
2018년 3월 ~ 현재 중앙대학교 산업보안학과 조교수
※관심분야 : Cyber Physical System Security