GENIANS WHITE PAPER

새로운 세상이 온다

부제 : 머신러닝을 이용한 악성코드 탐지의 새로운 변화

mkt@genians.com
www.genians.co.kr

Genians

Introduction

2016년 3월, 구글의 알파고(AlphaGo)와 이세돌의 대결은 인공지능에 대한 가능성과 두려움을 심어주기에 충분하였습니다. 그로부터 1년 반 후 구글은 새로운 알파고제로(Alphago Zero)를 선보이며 또 한번 세상을 놀라게 했습니다. 기존의 알파고와의 대국에서 100전 100승을 거두었기 때문 입니다. 더욱 놀라운 사실은 알파고제로의 학습방법에 있습니다. 과거의 알파고가 인간의 기보를 반복 학습하였던 데 반해, 알파고제로는 바둑의 규칙을 기반으로 스스로 학습이 이루어 졌기 때문입니다. 더 이상 사람의 지도감독(Supervised) 없이 스스로 강화학습(Reinforcement Learning)을 통해서 '축'에 대한 이해 등 바둑의 기본 지식을 깨닫고 기존의 실력을 크게 뛰어 넘는 수준에 이르게 되었습니다.

이제 알파고의 확장가능성이 주목 받고 있습니다. 바둑이 아닌 범용학습을 통해 다양한 분야에 인공지능 및 머신러닝이 적용될 수 있다는 가능성을 충분히 보여주었기 때문입니다.
정보보안 분야에 적용된다면 어떨까요? 많은 보안 업체들은 이러한 가능성을 바탕으로 이미 인공지능 및 머신러닝을 적용한 기술 및 솔루션을 선보이거나 대규모 투자를 진행하고 있습니다.

Symantec
"DLP · 보안관제시스템 · IoT 보안에 머신러닝 활용"

IBM Watson
"보안 인텔리전스 · SIEM에 '머신러닝 왓슨' 적용"

NORTHROP GRUMMAN
"Sandbox로는 악성코드 문제를 해결할 수 없는 것을 확인하고, 머신러닝을 악성코드 탐지에 활용한 기술개발 및 자회사 설립"

DARKTRACE
"머신러닝을 보안에 적용, 대규모 투자 유치 (삼성 SDS 포함, 영국정부 지원)"

SentinelOne CYLANCE
"머신러닝을 Endpoint 보안에 접목, 성공적인 사업전개"

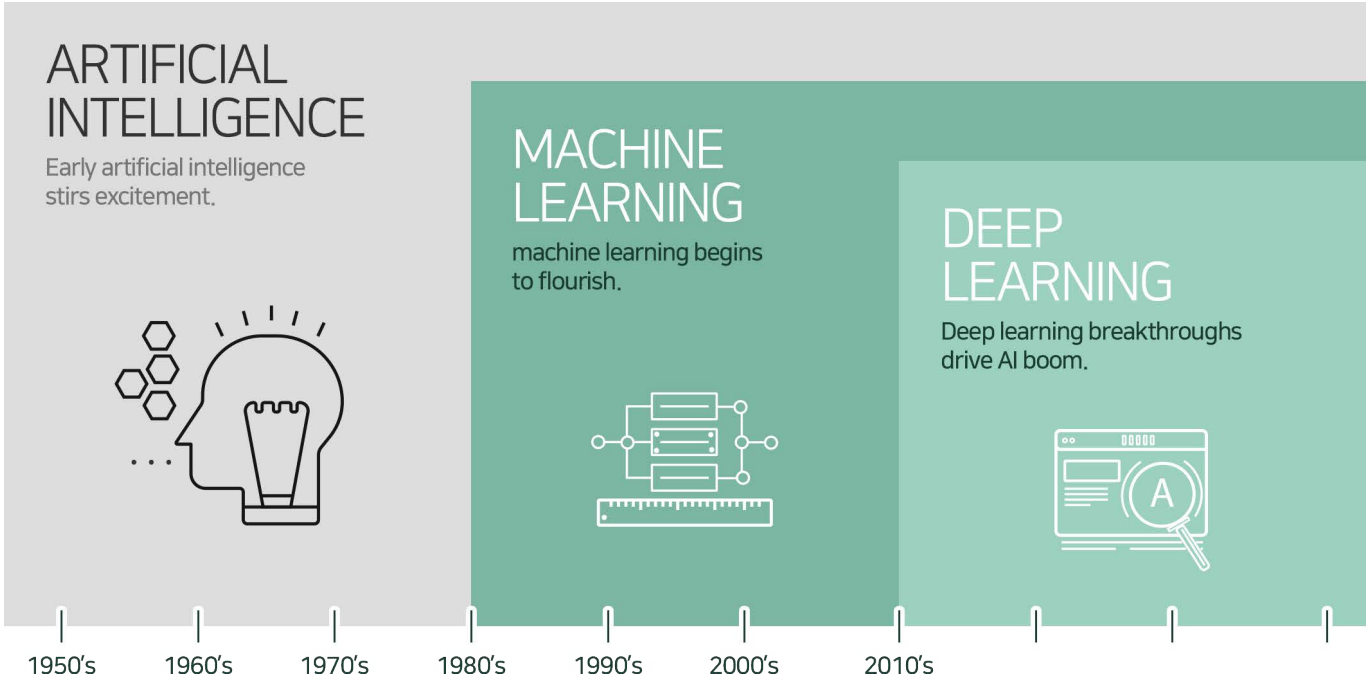
SECURITY MATTERS
"머신러닝을 제어보안에 접목, 글로벌 고객 확보 및 투자 유치"

머신러닝은 정보보안 영역에서 큰 역할을 수행하고 있습니다. 이미 악성코드 탐지 기술을 활용하여 APT 및 랜섬웨어 등 악성코드를 탐지, 차단할 뿐 만 아니라 네트워크 및 사용자 행위의 모니터링을 통해 오용(Anomaly)을 감지하여 위협을 예방하는 등이 대표적인 사례라고 하겠습니다.

본 문서는 Genian EDR(지니안 EDR)가 악성코드 탐지를 위해 사용하는 머신러닝에 대해 소개합니다. 이와 더불어 딥러닝 등 새로운 기술 및 제품에 대한 이해를 높이는 것을 그 목적으로 합니다.

머신러닝(Machine Learning)의 이해

머신러닝을 이해하기 위해서는 먼저 인공지능(Artificial Intelligence)과 머신러닝(Machine Learning) 그리고 딥러닝(Deep Learning)의 관계를 이해할 필요가 있습니다. 아래그림은 이들간의 관계를 잘 보여주고 있습니다.



[인공지능, 머신러닝, 딥러닝의 관계 - 엔비디아]

머신러닝(Machine Learning)의 이해

인공지능은 오래 전에 등장한 개념입니다. 당시에는 인간의 지능과 유사한 특성을 가지는 컴퓨터를 꿈꾸었습니다. 즉 인간의 사고력을 지니고 인간처럼 생각하는 일반AI(General AI)를 목표로 하였습니다. 그러나 많은 어려움에 직면하면서 일반AI는 실현되지 못하였습니다. 현재의 수준은 이미지를 분류하거나 얼굴 등을 인식하는 등의 특정 작업을 인간 이상의 수준으로 처리할 수 있는 수준 입니다. 좁은AI(Narrow AI)의 범주라고 할 수 있습니다.

머신러닝(Machine Learning): 좁지만 구체화된 인공지능

머신러닝은 인공지능을 구현하는 구체적인 접근방식 이라고 할 수 있습니다. 머신러닝은 알고리즘을 이용해 데이터를 분석하고 분석을 통해 학습하며 학습한 내용을 바탕으로 판단이나 예측을 합니다. 즉 구체적인 방향이나 지침을 코딩하는 것이 아니라 대량의 데이터와 알고리즘을 통해 학습을 진행하는 방식이라고 할 수 있습니다.

딥러닝(Deep Learning): 심층학습, 현재까지 가장 뛰어난 머신러닝

딥러닝은 인공신경망(ANN, Artificial Neural Networks)을 기반으로 하는 머신러닝의 한 분야 입니다. 인공신경망 역시 부침을 거듭하다가 기술적 한계의 극복, GPGPU(General-Purpose computing on Graphics Processing Units) 등 하드웨어의 발전 그리고 빅데이터(Big Data)와 어울리면서 엄청난 발전을 이루게 됩니다. 이후 각종 머신러닝 대회 등에서 압도적인 성능을 보이며 단연 두각을 보이게 됩니다. 최근에는 영상처리 및 음성인식 분야 역시 딥러닝에 대한 연구가 활발히 진행되고 있습니다.

머신러닝의 학습과 적용

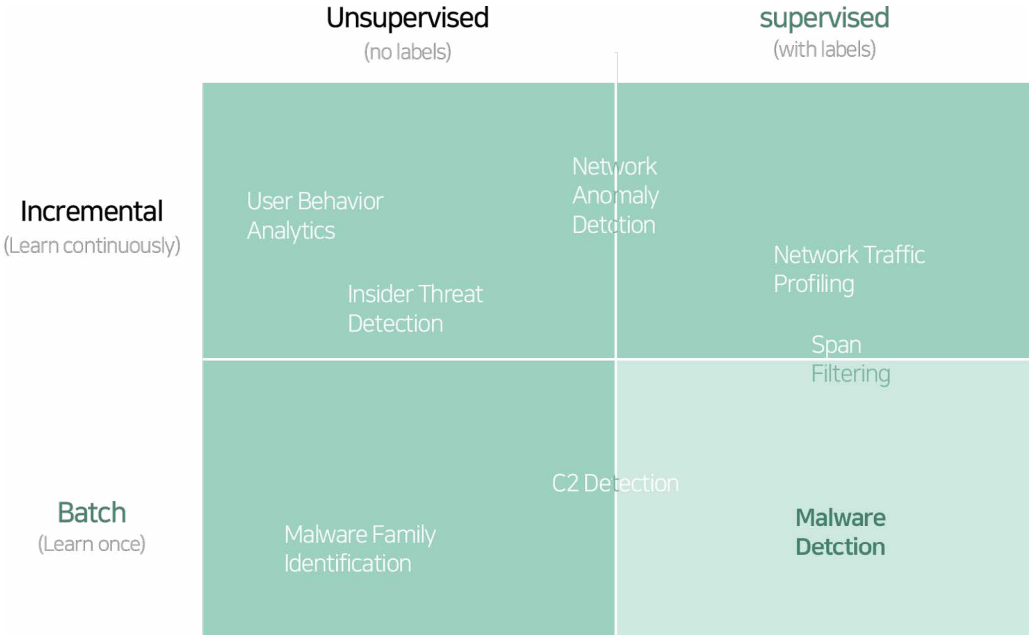
과거 인공지능의 학습방법은 인간의 지식을 저장하고 이를 추론하는 하향식 접근방식이었습니다. 그러나 우리는 어떤 지식을 다양한 경험과 데이터를 통한 학습과정으로 축적하는 경우가 더 많습니다. 머신러닝은 학습능력을 기계를 통해 구현하는 방법으로 환경과의 상호작용에 기반한 데이터로부터 스스로 성능을 향상시키는(기계가 학습할 수 있는) 알고리즘 및 기술을 개발하는 상향식 접근방식이라고 할 수 있습니다.

머신러닝은 학습하는 방식에 따라 ①지도학습(Supervised Learning), ②비지도학습(Unsupervised Learning), ③강화학습(Reinforcement Learning)으로 구분할 수 있습니다.

구분	내용
지도학습 (Supervised)	· 문제와 답을 동시에 주고 학습 (Labeled Data) · 주로 '인식, 분류, 진단, 예측' 등의 문제 해결에 적합 · 좋은 결과를 위해 시간과 비용이 증가 · 얼굴인식, 음성인식, 언어변역 등에서 활용
비지도학습 (Unsupervised)	· 문제만 주고 학습 (Unlabeled Data) · 주로 군집화, 밀도추정, 차원축소, 특징추출 등의 문제에 적합 · 지도학습 대비 학습 데이터 구축이 용이하고 비용이 절감 · 인간(어린이)의 학습형태와 유사하여 향후 발전가능성 높음
강화학습 (Reinforcement)	· 결과에 대한 피드백을 통하여 학습 · 특정 행동에 대하여 외부 환경에서 보상/피드백이 주어지며 보상이 최대화 하는 방향으로 학습이 진행 · 게임, 로봇주행 등에서 활용

[머신러닝의 학습방법 비교]

정보보안 분야 역시 머신러닝의 연구 및 적용이 활발하게 이루어 지고 있습니다. 스팸필터링(Spam Filtering)은 지도학습이 적용된 가장 대표적인 사례라고 할 수 있습니다. 이외에도 학습방법과 특징에 따라 사용자행위분석(User Behavior Analytics), 이상행위탐지(Anomaly Detection), 악성코드 탐지(Malware Detection), 인증(행위분석을 통한 개인 식별), 보안관제, 포렌직 등의 광범위한 사이버보안 분야에서 연구 및 적용이 진행되고 있습니다.



[정보보안 분야의 머신러닝 활용]

머신러닝과 새로운 플레이어의 등장

악성코드탐지(Malware Detection) 분야에서의 딥러닝의 활용 및 발전은 혁명에 가깝다고 할 수 있습니다. 최근 APT(지능형 표적공격) 및 랜섬웨어(RansomWare) 등 지능형 위협이 기하급수적으로 증가함에 따라 패턴(Signature) 기반의 안티바이러스(Anti-Virus)제품 군의 탐지 및 대응능력이 한계에 다다르고 있습니다. 이러한 변화에 따라 시만텍(Symmantec) 등 전통적인 보안업체의 머신러닝 도입이 가속화 되고 있으며 차세대 단말보안(NGES: Next Generation Endpoint Security) 또는 차세대백신(NGAV: Next Generation Anti-Virus) 등의 새로운 단말보안의 영역과 함께 플레이어들이 주목을 받고 있습니다.

CYLANCE

Founded in 2012

- Machine learning 기반 악성코드 탐지
- $177M funding
- Investment values company at $1B

SentinelOne

Founded in 2013

- Machine learning 기반 악성코드 탐지
- $109.52M funding

BLUVECTOR

Founded in 2015 (by Northrop Crumman)

- Machine learning 기반 악성코드 탐지
- Acquired by LLR Partners on January 9, 2017
- $50M funding (Private Equity LLC Partner Acquisition)

invincea

Founded in 2009

- Machine learning 기반 악성코드 탐지
- $47.4M funding
- Acquired by Sophos on February 8, 2017 ($100M)

이들은 머신러닝을 이용하여 악성코드를 탐지하고 시스템의 익스플로잇(Exploits)등 비정상 행위를 감지하여 위협을 제거합니다. 뿐만 아니라 네트워크의 트래픽 과 패킷을 분석하고 흐름(flow)을 학습하여 오용(Anomaly)을 탐지하고 위협을 예방할 수도 있습니다. 탐지된 위협(threat)의 근본원인(Root Cause)과 파일, 프로세스, 네트워크 등의 상호 연관관계를 분석하여 대응의 범위를 확장하고 정밀한 대응이 가능하게해 줍니다. 체인이벤트, 어택 타임라인 등의 다양한 시각화 기법을 제공하여 위협에 대한 가시성과 대응의 적시성을 보장해 줍니다.

기존의 백신과 단말보안 제품이 제공하는 기능과 효용을 크게 뛰어넘었다는 평가를 받습니다. 시장의 평가도 긍정적 입니다. 투자금과 기업의 가치 평가가 이를 증명해 줍니다. Cylance의 경우 2012에 설립되었지만 무려 1조의 가치를 평가 받습니다.

어떻게 불과 수 년 사이에 이러한 변화가 가능해 진 것일까요? 변화의 중심에 머신러닝이 있다고 볼 수 있습니다. 특히 머신러닝 중 딥러닝(심층학습, Deep Learning)은 다른 머신러닝의 학습방법과 비교할 때 프로그래밍의 수고를 크게 덜어 줍니다. 과거 머신러닝을 활용하는데 있어 가장 큰 걸림돌은 바로 피쳐엔지니어링(Feature Engineering)이었습니다. 이것은 분석가(Analyst) 또는 데이터 과학자(Data Scientist)가 특정 데이터에서 특징(feature)을 추출하고 재가공하는 일련의 작업을 의미합니다. 딥러닝은 이러한 특징의 추출과 학습이 자동으로 이루어지는 학습 방법 입니다. 따라서 많은 데이터와 컴퓨팅파워가 제공된다면 충분히 신뢰할 수 있는 결과를 기대할 수 있게 되었습니다. 정리하자면 아래와 같은 요인이 딥러닝의 발전과 함께 새로운 플레이어의 출현을 가속화 했다고 할 수 있습니다.

✕ 데이터 비용의 감소

빅데이터 이슈와 함께 데이터의 양(Quantity) 과 질(Quality)이 크게 발전하였습니다. 과거에는 고작 손글씨 데이터 (e.g, MNIST)정도가 전부였으나 현재는 수천만 장의 고해상도의 이미지는 물론(e.g, ImageNet) 유투브, SNS 등도 활용할 수 있습니다. 특히 랜섬웨어 등의 악성코드의 경우 사이버위협인텔리전스(CTI. Cyber Threat Intelligence)의 발전과 함께 공유 및 협업이 더욱 중요해 지고 있습니다. 바이러스토털(VirusTotal), 멀웨어스닷컴(malwares.com), 멀코드(Malc0de)등 악성코드의 분석 및 평가 등 협업플랫폼이 확대 되면서 매우 양질의 분류데이터(Labeled Data)를 획득 및 재처리 하는 비용이 감소하였고 이를 바탕으로 발전이 가능하게 되었습니다.

✕ 하드웨어의 발전

머신러닝의 학습과정은 엄청난 연산능력을 요구합니다. 그러나 범용 컴퓨터의 CPU는 물리적 코어(Core)의 수가 한정되어 있고 순차적인 연산에 특화되어 있습니다. 이와 비교해 GPU는 수십 개 이상의 코어를 보유할 수 있으며 이를 병렬로 처리하는 경우 다중연산, 특히 숫자나 알고리즘을 처리하는데 매우 유용합니다. 또한 이를 효율적으로 이용할 수 있는 언어구조(e.g, CuDA)가 개발되고 가격이 저렴해지면서 딥러닝은 그 컴퓨팅 시간을 수십 분의 일로 줄일 수 있었습니다. 과거 구글이 범용 서버 1,000대를 병렬로 연결해 시도한 '구글브레인' 프로젝트를 현재는 GPU 가속화 서버 3대로 처리할 수 있을 정도로 하드웨어는 비약적으로 발전하고 있습니다.

✕ 오픈플랫폼의 약진

구글(Google), 마이크로소프트(Microsoft) 등의 글로벌 IT 기업들과 학계 연구그룹들이 머신러닝 관련 플랫폼(프레임워크 및 라이브러리 등)을 무료로 공개하고 있습니다. 이러한 플랫폼은 사용자의 기술적 진입장벽을 획기적으로 낮추어 어플리케이션과 효용(Value)에 집중할 수 있게 해 줍니다. 특히 구글이 공개한 텐서플로우(TensorFlow)는 가장 대표적으로 이미 지메일의 스팸필터링, 이미지 검색 등에 사용되고 있으며 이를 이용한 악성코드 탐지, 신용카드 오용탐지 등 다양한 영역에서 활용되고 있습니다.

딥러닝은 어떻게 동작하는가?

최근 딥러닝이 크게 주목 받고 있습니다. 몇 년 전부터 머신러닝이 일반의 관심을 받기 시작하더니 지금은 머신러닝의 한 종류인 딥러닝이 머신러닝을 대표하다시피 이야기 되고 있습니다. 기업들은 관련 인력의 확보에 사활을 걸고 있습니다. 구글이 딥마인드를 인수하고 페이스북이 딥러닝의 대가인 얀 러쿤(Yann LeCun) 교수를 인공지능 센터장으로 모셨으며 중국의 구글이라고 불리는 바이두에서도 앤드류 응(Andrew Ng)교수를 모셔가는 등 인재전쟁에 가까운 모습입니다.

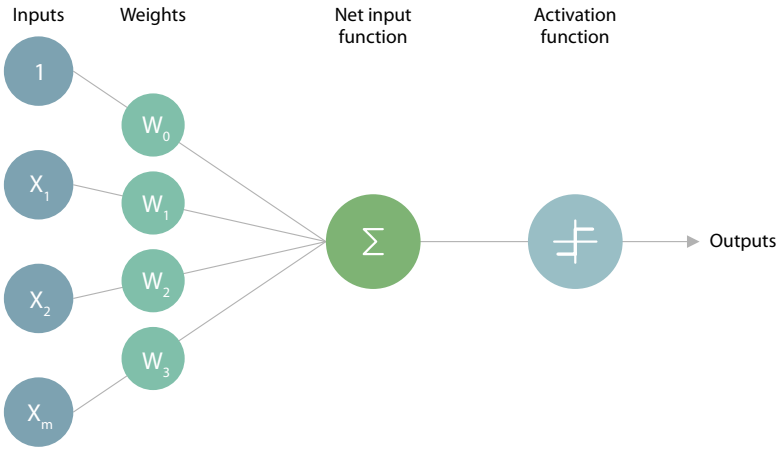
그렇다면 딥러닝은 어떻게 동작할까요?

여기 다각형(Polygon)을 구분할 수 있는 딥러닝을 만든다고 가정하고 그 동작방식을 개념적으로 이해해 보도록 합시다. 이 중 왼쪽 파란색 도형은 무엇인가요?

사람은 왼쪽의 파란색 개체(object)가 정사각형임을 바로 인지할 수 있습니다. 왜냐하면 사람은 구체적으로 추상화된 정사각형의 개념을 지식으로 가지고 있기 때문 입니다. 그래서 왼쪽의 개체가 다각형이며 그 중에 정사각형임을 바로 결정할 수 있습니다. 반면 머신, 즉 기계의 경우는 어떨까요? 아쉽게도 사람과 같은 지식의 저장과 이를 바탕으로 하는 결정(판단)이 불가능합니다. 결국 하나하나 특징(Feature)을 인지하고 이를 조합하여 결정을 내리는 방식을 취하게 됩니다. 그 결과 아래와 같은 단계가 필요하게 됩니다.

- 직선의 존재 유무
- 직선의 개수
- 직선의 길이
- 직선의 연결 수
- 직선의 연결 각도 등

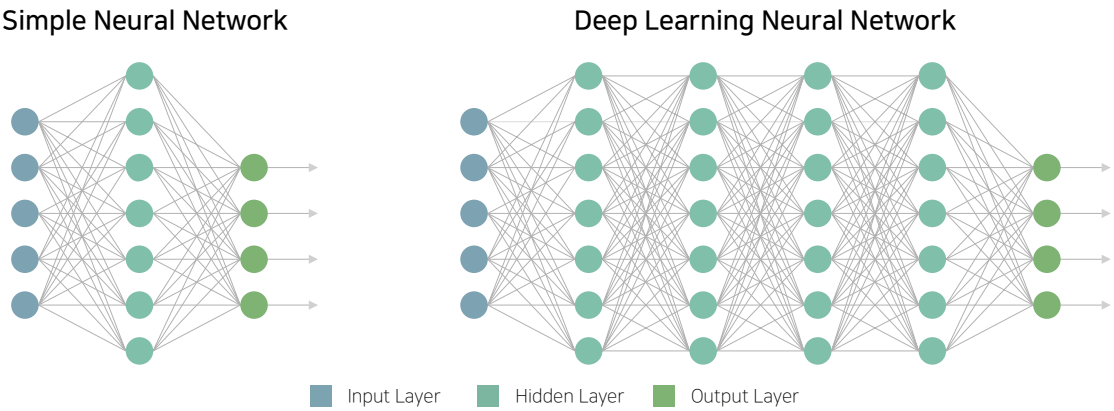
이러한 특징이 입력데이터로 딥러닝에 전달 됩니다. 딥러닝은 여러 개의 층(Layer)으로 이루어진 신경망을 의미 합니다. 한 층은 다시 여러 개의 노드로 이루어져 있습니다. 노드에서는 실제로 연산이 일어나는데 이 연산 과정은 인간의 신경망을 구성하는 뉴런에서 일어나는 과정을 모사하도록 설계되어 있습니다.



[노드의 연산 - 입력데이터와 가중치를 통해 활성여부가 결정됨]

노드는 일정크기 이상의 자극을 받으면 반응을 하는데 그 반응의 크기는 입력 값과 노드의 계수(또는 가중치, Weights)의 곱에 비례 합니다. 일반적으로 노드는 여러 개의 입력을 받으며 입력의 개수만큼 계수를 가지고 있습니다. 따라서 이 계수를 조절하는 것으로 여러 입력에 서로 다른 가중치를 부여할 수 있습니다. 최종적으로 곱한 값들은 모두 더해지고 그 합은 활성함수(Activation Function)의 입력으로 들어가게 됩니다.

위에서 언급한 특징(Feature)들은 입력데이터로 첫 번째 층(Layer 1)의 입력이 되며 그 이후 각 층의 출력(결과)이 다시 다음 층(Layer 2)의 입력이 됩니다. 층이 거듭될수록 복잡하고 추상적인 학습이 이루어 집니다. 계수(Weights)는 학습 과정에서 미세하게 조정되며 결과적으로 각 노드가 어떤 입력을 중요하게 판단하는지는 결정합니다. 결국 학습(Learning)은 최적화된 결과를 도출할 수 있도록 이 계수를 최적화, 업데이트 하는 과정이라고 할 수 있습니다.

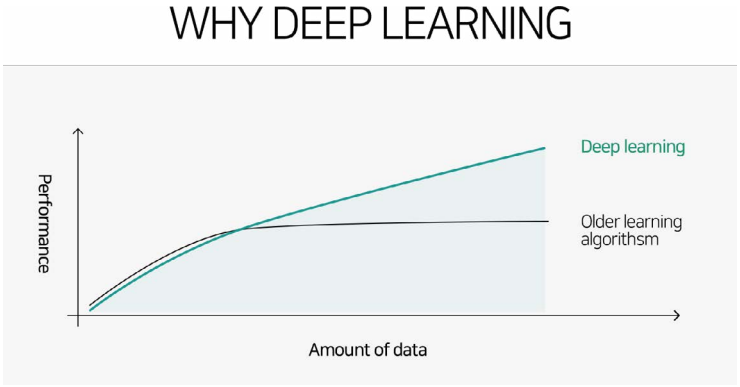


[단일(Simple) 신경망과 심층(Deep) 신경망]

이후 모든 결과를 종합하여 어떠한 형태의 다각형인지를 판단할 수 있게 됩니다. 이것은 마치 의사결정트리(Decision Tree)와 유사해 보입니다. 그러나 이전 단계에서 다음단계로의 입력변수가 2개 이상일 수 있으며 이것은 딥러닝의 근간인 신경망(Neural Network)의 특징입니다.

딥러닝은 인공신경망(ANN, Artificial Neural Network)에 기반하여 입력층(Input Layer) 과 출력층(Output Layer) 그리고 다수의 은닉층(Hidden Layer)의 계층 구조를 가지는 심층신경망(DNN, Deep Neural Networks)을 학습의 주요 방식으로 사용하는 머신러닝의 한 분야 입니다. 실제 딥러닝의 동작은 선형맞춤(Linear Fitting) 과 비선형변환(Nonlinear Transformation)의 반복이라고 할 수 있습니다. 즉 간단한 학습 구조를 쌓아 올려가며 순차적으로 학습하는 계층적 구조의 학습법이라고 할 수 있습니다.

딥러닝의 가장 큰 특징은 최적화된 결정을 위한 이러한 특징(feature)의 추출과 학습이 함께 이루어진다는 점 입니다. 이는 앞서 설명한 피쳐엔지니어링의 수고를 크게 덜어 줍니다. 따라서 대량의 정제된 데이터(Labeled Data)가 제공된다면 충분히 신뢰할 수 있는 학습모델을 얻을 수 있습니다.



[Why Deep Learning? _ Andrew Ng]

위의 그림은 데이터의 증가와 딥러닝의 효과(Performance)와의 관계를 보여 줍니다. 이것이 딥러닝이 최근 가장 주목을 받은 이유라고 할 수 있습니다.

머신러닝과 악성코드 탐지

최근 랜섬웨어가 큰 이슈가 되고 있습니다. 악성코드의 일종인 랜섬웨어를 이용하여 공격자들은 금전적 이득을 취하고 있습니다. 시만텍(Symantec)의 인터넷위협동향보고서(ISTR)에 따르면 2015년 발생한 악성코드의 수는 약 4억3천만개라고 합니다. 2009년 한 해 발생한 악성코드의 개수가 약 236만개 라고 하니 2015년에는 하루에 약 118만개의 악성코드가 발생한 샘 입니다. 매년 30배씩 증가했다고 볼 수 있습니다.

급격한 악성코드의 증가원인 중에 변종이 있습니다. 대다수의 악성코드 제작자들은 백신을 피하기 위해 변종코드를 만들어 유포하고 있습니다. 독일의 보안회사 지데이터(G-Data)에 따르면 올해 1분기 감지된 신종, 변종 악성코드는 185만개에 이릅니다. 4초에 1개 꼴로 새로운 악성코드가 나타나고 있으며 이 가운데 60% 이상은 랜섬웨어라고 합니다.
랜섬웨어는 누구나 쉽게 입수해 변종을 만들 수 있고 가상화폐의 등장으로 추적 받지 않고 돈을 벌 수 있어 빠르게 유포되고 있습니다. 이러한 환경의 변화 속에서 머신러닝이 악성코드의 탐지와 관련해서 주목을 받는 이유를 다음과 같이 정리할 수 있습니다.

✕ 탐지 방식의 한계

안티바이러스(Anti-Virus) 제품의 가장 기본적인 탐지 방법은 시그니쳐(Signature)와의 비교 입니다. 2016년 한 해, 매일 약 1백만 건의 악성코드가 신규로 발견되었습니다. 이 중 안티바이러스 제품에 적용될 수 있는 수는 수백 건 정도 입니다. 결국 모든 악성코드를 시그니처로 관리하는 것은 불가능 하다는 결론에 이르게 됩니다.

✕ 동작 방식의 한계

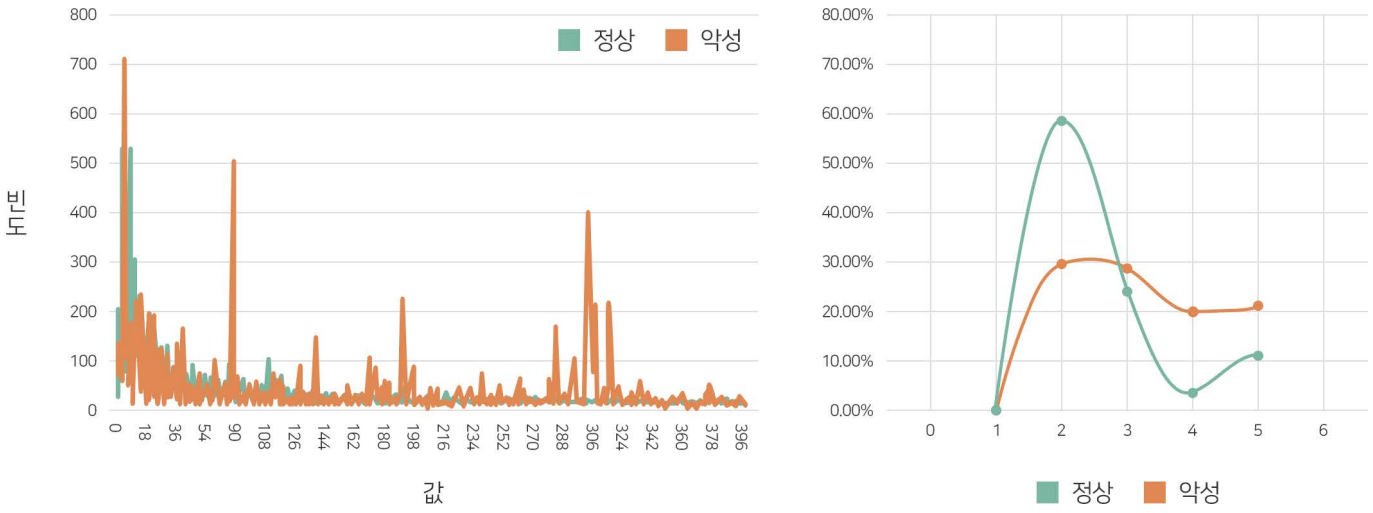
최신의 시그니처를 유지하기 위해서는 빈번한 업데이트가 반드시 필요 합니다. 네트워크를 이용한 업데이트는 불행히도 폐쇄망에서는 이용할 수 없습니다. 이것은 클라우드를 이용하는 동작방식에도 큰 걸림돌이 됩니다. 얼마 전 발생한 국방부 해킹사건은 이러한 폐쇄망의 한계를 잘못된 방식으로 해결하려는 시도가 얼마나 큰 결과를 초래하는지를 보여 준 대표적인 사례라고 할 수 있습니다.

✕ 백신 등 보안 소프트웨어 우회

악성코드를 탐지하기 위한 방법을 공격자가 역으로 이용할 수 있습니다. 자신이 작성한 악성코드를 바이러스토털(VirusTotal)이나 쿠쿠샌드박스(Cuckoo Sandbox)등을 이용하여 테스트하거나 이를 우회하는 기술적인 방법을 적용할 수 있습니다.

이러한 이유로 증가하는 악성코드의 탐지를 위해 머신러닝이 실질적인 대안으로 평가 받고 있습니다. 머신러닝은 시그니처(Signature)가 아닌 특징(Feature)을 기반으로 악성코드를 탐지하는 기술 입니다. 따라서 악성코드의 양(Quantity)과 탐지율(Detection Rate)의 관계가 없으며 유사한 변종의 탐지에 유리 합니다. 그렇다면 머신러닝은 어떻게 악성코드를 탐지할 수 있을까요? 앞에서 언급한 다각형을 인지하는 딥러닝 모델을 떠올리면 이해하기 쉽습니다. 우리는 다각형을 인지하기 위해서 '직선, 연결성, 각도' 라는 3가지 특징(Feature) 과 이에 따른 몇 가지 가중치(Weight)를 이야기 했습니다. 악성코드의 탐지 역시 유사 합니다. 중요한 것은 악성코드로 판단하기 위해 어떠한 특징(Feature)을 사용할 것인가와 학습을 위해 어떠한 알고리즘을 사용하는가에 있습니다.

어떠한 특징(feature)이 실행프로그램을 유해한(악성코드) 것과 정상인(정상코드) 것으로 잘 구분해 줄 수 있을까요? 사용할 수 있는 많은 특징이 있습니다. 파일의 이름과 해쉬값부터(유용하지는 않습니다.) 헤더정보, 호출함수, 레지스트리키, DLL 등이 이에 해당 합니다. 그러나 불행히도 악성코드와 아닌 것을 딱 잘라 구분할 수 있는 단일한 특징은 존재하지 않습니다. 예를 들어 실행프로그램의 헤더(Header)에는 SizeOfInitializedData 라는 필드(Field)가 존재합니다. 이것은 프로그램에서 사용되는 변수들이 초기화 되어 있는 영역의 총 합을 의미 하는데, 다수의 악성코드와 정상코드를 대상으로 해당 특징의 분포를 분석해 보면 아래와 같습니다.



[파일사이즈에 따른 SizeOfInitializedData값의 분포]

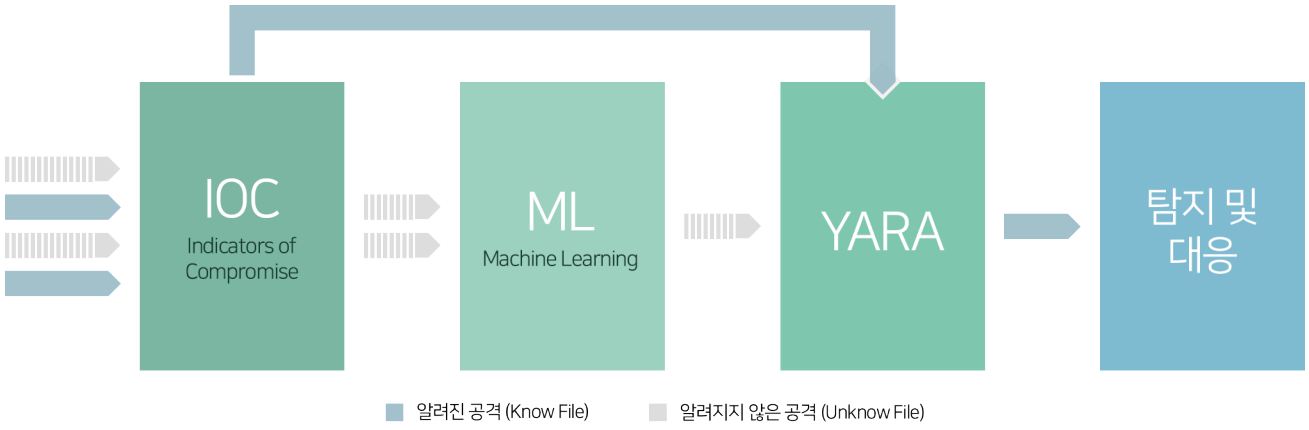
'3구간'을 기준으로 분포가 역전되는 현상이 발생합니다. 아쉽게도 이 특징은 유용해 보이지는 않습니다. 이러한 특징은 대부분의 악성코드와 정상코드에서 동일하고 반복적으로 나타나게 됩니다. 따라서 실제 악성코드 탐지에는 '수백개 ~ 수천개' 의 특징(Feature)과 가중치(Weight)를 조합하여 이용하게 됩니다. 이것이 바로 머신러닝이 필요한 이유 입니다.

결국 탐지 성능은 악성프로그램과 정상프로그램을 대상으로 어떠한 특징을 추출/사용하여 어떠한 알고리즘으로 어떻게 학습시켰느냐에 달려 있습니다. 머신러닝을 이용한 다양한 벤더가 출현할 수 있는 이유이기도 합니다. 현재는 정적인 특징(Static Feature)뿐 아니라 동적인 특징(Dynamic Feature)을 추출하여 사용하며 다양한 알고리즘 또는 다양한 데이터를 함께 사용하여 예측 성능을 높이는 앙상블(Ensemble) 방법 등이 사용되고 있습니다.

Genian EDR 와 머신러닝

지니언스㈜의 EDR은 '단말기반 지능형 위협탐지 및 대응솔루션'으로 국내최초로 개발된 EDR(Endpoint Detection & Response)솔루션 입니다. APT 와 랜섬웨어 등 지능형위협을 탐지하고 공격에 대한 가시성(Visibility)을 확보할 수 있습니다. NAC와 긴밀한 협업을 통해 위협을 조기에 발견하고 대응하여 위협으로 인한 피해(Risk)를 최소화 할 수 있어 이미 NAC를 사용하고 있는 환경에서 주목을 받고 있습니다.

Genian EDR는 지능형 위협을 탐지하기 위해 머신러닝을 포함한 다단계 탐지 방식을 지원하고 있습니다.



[Genian EDR의 위협 탐지 단계]

✕ IOC(침해사고지표, Indicators of Compromise)

이미 알려진 악성코드의 해쉬(hash), 분류, 위험성, IP 등 관련 정보를 기반으로 악성코드를 탐지 합니다. 안티바이러스 제품의 시그니쳐와 유사하며 실제 다수의 악성코드가 이 단계에서 사전 탐지 됩니다.

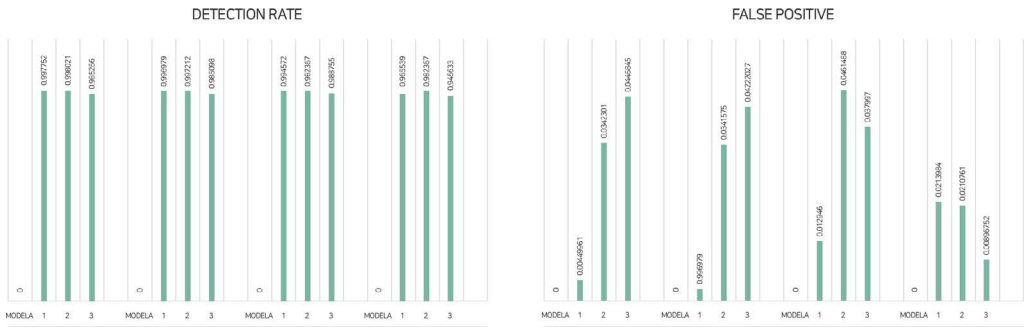
✕ ML(머신러닝, Machine Learning)

IOC에 의해 탐지되지 않은 실행파일의 경우 머신러닝에 의해 추가 탐색이 이루어 집니다. 1,000개 이상의 특징(Feature)을 추출하여 정교하게 학습된 모델을 적용하는데 채 1초가 걸리지 않습니다. 탐지 정확도는 99% 이상 입니다.

✕ YARA(야라)

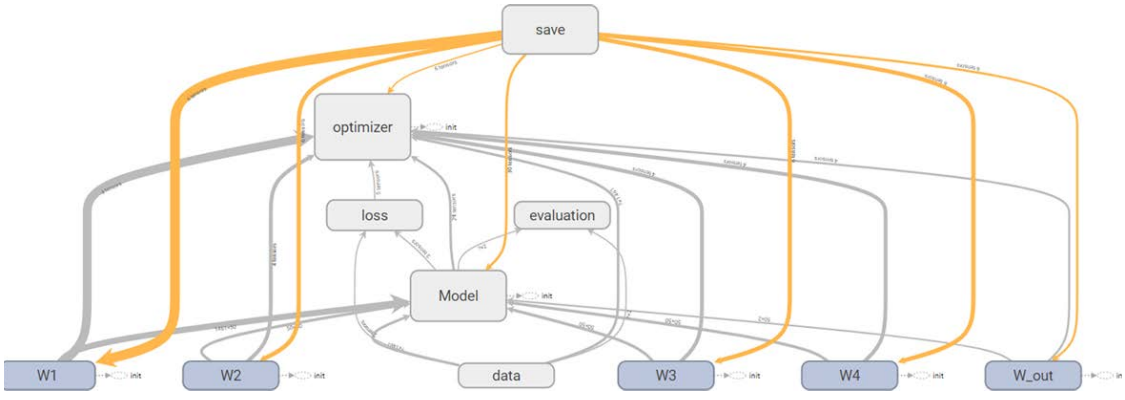
추가로 실행파일 내부에 악성코드의 흔적(String)을 규칙(Rule)을 기반으로 탐지 합니다. 이미 알려졌거나 또는 알려지지 않은 유사변종 악성코드를 탐지 할 수 있습니다.

Genian EDR는 악성코드와 정상코드의 구분을 위하여 약 1,500개 이상의 특징(Feature)을 사용하며 딥러닝을 기반으로 하는 다양한 학습 방법을 연구하고 있습니다. 이 중 4개의 학습모델(Model A, B, C, D) 바탕으로 약 10만개의 파일을 대상으로 3회 검사한 결과는 아래와 같습니다. (아래의 결과는 내부 측정결과이며 Training Set 으로 학습시킨 후 약 10만개의 Test Set을 검증하여 도출한 결과 입니다.)



[Test Set을 이용하여 측정한 탐지율(정탐, 오탐) 결과]

악성코드 탐지율(Detection Rate)에 있어 4개 모델 평균 98.61%의 탐지율을 보였으며 가장 뛰어난 결과를 보인 Model A 의 경우 99.36% 였습니다. 정상파일을 악성코드로 탐지하는 비율(False Positive, Type I Error)에서는 4개 모델 평균 2.6%의 결과를 보였으며 가장 뛰어난 Model D의 경우 1.7%로 확인되었습니다. 결론적으로 악성코드의 탐지율에서는 Model A 가, 실제 적용을 위한 오탐율에서는 Model D 가 가장 뛰어난 모델임이 확인되었습니다.



실제 Genian EDR에 적용되는 머신러닝은 이보다 훨씬 복잡하고 정교하게 최적화(Optimized)된 학습모델이 탑재 됩니다. 앞의 예에서와 같이 서로 다른 모델이 동시에 학습, 사용되거나 또는 판단결과를 다시 재 학습하는 등의 다양한 방법이 적용됩니다. 이러한 노력은 탐지율의 고도화와 오탐율의 감소로 이어져 실제 악성코드로 인한 보안위협을 제거하는데 획기적인 역할을 할 것으로 기대하고 있습니다.

[신경망(Neural Network)의 학습 – 수 많은 특징(Feature)의 입력값(Input)과 가중치(Weight)를 반복 업데이트하면서 악성코드와 정상코드를 구분할 수 있는 최적의 모델이 완성된다.]

머신러닝, 과연 만능입니까?

많은 업체들이 머신러닝을 이야기 합니다. 악성코드의 증가 특히 랜섬웨어와 변종의 출현에 대한 대안으로 빠르게 자리를 잡아가는 모양새입니다. 심지어는 랜섬웨어를 100% 탐지할 수 있다고 선전하며 탐지율 경쟁으로 치닫는 모습도 볼 수 있습니다. 그러나 머신러닝의 실제 적용에 있어서는 아래와 같은 한계가 존재 합니다. 이러한 특징을 정확하게 이해하고 올바르게 사용하는 것이 머신러닝의 적용에 있어 매우 중요합니다.

✕ 탐지 결과의 해석

머신러닝으로 악성코드가 탐지되는 경우 그 결과값은 확률(%)로 표기 됩니다. 즉 탐지결과는 'foo.exe 라는 파일이 90%의 확률로 악성이라고 판단됨' 과 같습니다. 구체적으로 어떠한 이유 때문에 악성코드로 판단되었는지를 확인(해석)할 수 없습니다. 이러한 해석을 위하여 의사결정트리(Decision Tree), 선형회기(Linear Regression) 등의 추가적인 모형을 이용할 수 있으나 이 역시도 추정에 가깝다고 할 수 있습니다.

✕ 오탐(False Positive)

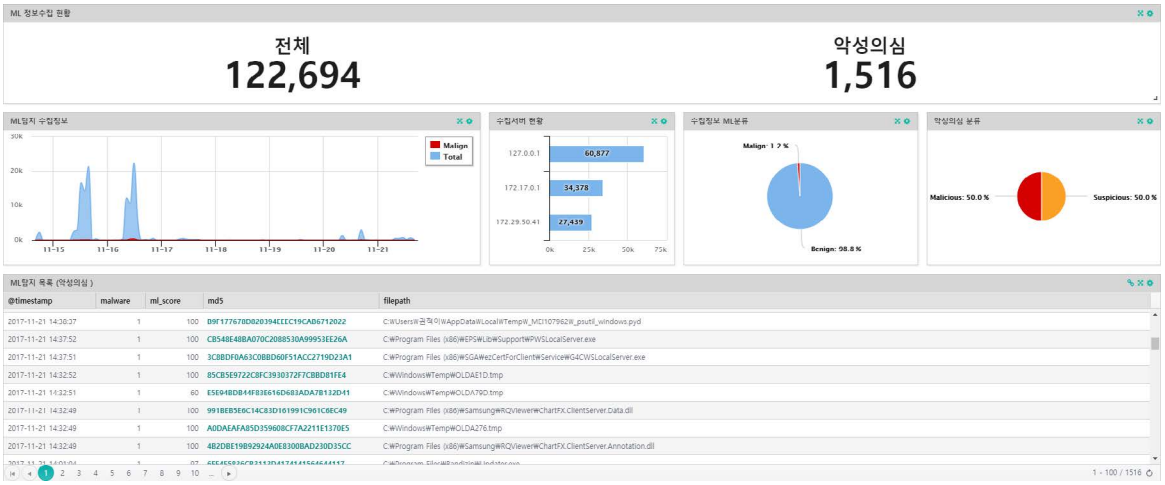
'높은 탐지율' 보다 더욱 중요한 것은 '낮은 오탐율' 입니다. 특히 정상파일을 악성파일로 판단하는 오류(False Positive, Type I Error)의 관리가 매우 중요합니다. 오탐율 5%는 숫자로는 낮아 보입니다. 그러나 1,000개의 파일을 검사했을 때 50(5%)개의 파일을 삭제할 수 있다는 의미와 같습니다. 이러한 오류는 실제 적용에 있어 심각한 피해를 초래할 수 있습니다.

✕ 탐지 결과와 대응(Response)의 관계

'abc.dll' 이라는 파일이 55%의 확률로 악성코드로 탐지되었다면 어떠한 조치를 취하시겠습니까? 그냥 두어야 할까요? 아니면 삭제해야 할까요? 만일 삭제 후 시스템이 정상적으로 부팅하지 못하거나 어플리케이션에 장애가 발생하면 그 책임은 누구에게 있을까요? 아무리 뛰어난 머신러닝 이라도 그 결과를 즉각적인 대응으로 연결하기에 무리가 있습니다. 단일한 머신러닝으로만 이루어진 솔루션의 해결과제라고 할 수 있습니다.

✕ 업데이트(Update)

머신러닝 역시 업데이트가 필요 합니다. 그 주기는 수개월 ~ 수년 일 수 있습니다. 전혀 다른 종류의 악성코드가 출현하게 되면 탐지 및 대응이 어려울 수 있습니다. 악성코드가 변화하는 것에 맞추어 추가적인 학습이 필요합니다. 따라서 새로운 악성코드를 지속적으로 수집, 분석하고 업데이트 할 수 있는 에코(Eco)시스템이 필요 합니다.



[탐지율과 오탐율에 대한 지속적인 관리가 필요함]

지니언스㈜ 이러한 기술적인 문제와 적용상의 어려움을 잘 인지하고 있습니다. 또한 Genian EDR에는 이러한 문제를 해결하기 위한 다양한 방법들이 이미 적용되어 있습니다. 지니언스는 Genian EDR의 지속적인 고도화를 통해 기술적인 문제의 해결을 그리고 NAC(Network Access Control)와의 긴밀한 협업을 통해 적용상의 어려움을 해결할 수 있다고 확신하고 있습니다. 이것이 Genian EDR가 NAC와 함께 운용될 때 최적의 효과를 거둘 수 있는 이유 이기도 합니다.

Conclusion *

딥러닝(Deep Learning)은 오래 전부터 연구되어 왔습니다. 오랜 기간 부침을 거듭하였지만 꾸준한 연구가 지속되면서 알고리즘이 거듭 개선되었으며, 하드웨어의 발전 그리고 빅데이터의 발전과 맞물리면서 최고의 성능을 가진 머신러닝의 방법으로 평가 받고 있으며, 정보보안 분야를 포함하여 미래 인공지능의 희망으로 떠오르고 있습니다.

특히 악성코드 탐지 분야에서 딥러닝의 발전은 경이롭기 까지 합니다. 그러나 실제 사용(Usage) 관점에서 보면 아직은 '환상' 또는 '실망' 이라는 이분법적인 평가가 주를 이루는 것 같습니다. 왜 그럴까요? 바로 새로운 기술에 대한 정확한 이해와 적용이 없었기 때문 입니다. 단순히 높은 탐지율 같은 왜곡되고 단편적인 잣대로만 해당 기술을 평가했기 때문이라고 생각합니다.

새로운 세상이 오고 있습니다. 전통의 강자가 한 순간에 몰락하고 신기술과 신생업체가 새로운 패러다임을 제시할 수 있는 세상이 되었습니다. 이제 시스템과 사용자가 융합되고 네트워크와 엔드포인트의 구별이 없는 환경이 도래하고 있습니다. 정보보안 역시 머신러닝 등의 새로운 기술로 인해 영역이 파괴되고 있습니다. 그러나 걱정할 필요 없습니다. 결국 사람을 위하는 기술과 변화만이 살아남고 확대될 것이기 때문 입니다.

About Genians

지니언스는 내부보안 분야를 선도하는 기업입니다. 지니언스는 내부 네트워크 접속하고 활동하는 사람 그리고 단말의 행동에 집중합니다. 그것이 보안을 위한 가장 중요한 부분임을 잘알기 때문입니다. 우수한 개발자와 탁월하고 건전한 기업문화를 바탕으로 Genian NAC(Network Access Control) 과 Genian GPI(Genian PC Inspector) 그리고 Genian EDR를 출시하였으며 2017년 코스닥 상장을 통해 명실상부한 국내 보안 선도기업으로 자리매김 하였습니다.

국내 NAC 분야 1등을 넘어 글로벌 시장에서도 통하는 보안SW 기업을 목표로 클라우드 기반 서비스를 통해 미국 보스톤에 현지법인을 설립하고 미국, 유럽, 동남아로 서비스를 확장해 나가고 있습니다.