

차세대 AML이란?

로봇 공학과 의미 분석(semantic analytics) 및 인공 지능으로
한층 강화된 금융 범죄 방지 솔루션



목차

AML 기술 발전을 이끄는 원동력.....	1
세계 각지의 차세대 AML 솔루션 활용 현황.....	2
아시아 태평양 지역: 관심은 커지고 있으나 제도적 장치가 미비함.....	2
유럽: 발전하고 있으나 아직은 초기 단계.....	3
미국: 보다 엄격한 규제 환경 속에서 성숙 단계에 진입.....	3
차세대 AML로 성공하기 위한 10가지 요인.....	4
1. 핵심은 신중한 혁신.....	4
2. 엄격한 모델 거버넌스 구축.....	4
3. 국경 간 안전한 데이터 공유.....	4
4. 하이브리드 방식 고려.....	5
5. 철저한 데이터 기반 검토.....	5
6. 임계값 설정에 대한 미시적 접근.....	5
7. 중요도에 따른 경보 조치.....	6
8. 머신 러닝으로 희귀 사건 탐지.....	6
9. 베스트 프랙티스 활용.....	7
10. 금융 범죄 시스템과 프로세스의 통합.....	7
AML 솔루션 활용: 국가별 사용 사례.....	8
지능형 자동화.....	8
희귀 이벤트 탐지.....	8
인공 지능 기반 탐지.....	8
자연어 처리(NLP).....	9
맺음말.....	9

도와주신 분들:

Wallace Chow, AML Practice Lead, FCC Analytics

Amith Satheesh, Principal Solutions Architect and Global Lead, AML Analytics, Security Intelligence Practice, SAS

Beth Herron, Senior Solutions Architect and Americas AML Lead, Security Intelligence Practice, SAS

David Stewart, Director, Security Intelligence Practice for Financial Services, SAS

AML 기술 발전을 이끄는 원동력

마약 밀매, 밀수, 사기, 갈취, 부패의 공통점은 무엇일까요? 바로, 불법이지만 수익성이 높다는 것입니다. 이러한 범죄 행위로 거둬들이는 수익은 전 세계 GDP의 약 2~5%를 차지합니다. 유엔마약 범죄사무소의 조사에 따르면 이는 연간 8천억~2조 달러와 맞먹는다고 합니다.¹

자금 세탁은 이러한 자금의 출처와 목적지를 위장하여 금융 시스템 훼손과 같은 심각한 연쇄반응을 야기합니다. 테러리스트나 범죄 집단도 이렇게 해서 계속 돈을 벌어들입니다.

자금세탁방지(Anti-money laundering; AML)는 수십 년간 금융업계의 뜨거운 감자이자 규제 강화의 대상이 되어왔습니다. 대표적 예는 다음과 같습니다.

- 9/11 테러 직후 제정된 미국 애국자 법(USA PATRIOT Act)은 자금 세탁이나 테러 자금 조달 행위를 탐지 및 보고하는 데 필요한 요건을 확대함
- 2017년 발효한 뉴욕주 금융감독청 규정 제504호 'The Final Rule'에는 더욱 세밀하고 엄격한 요건이 추가되어 다국적 은행에서부터 동네 수표 교환점에 이르기까지, 뉴욕 은행법의 적용을 받는 모든 기관을 대상으로 시행됨
- 2020년 1월 발효하는 제5차 EU 자금세탁방지법(5EUAM)은 적용 대상 범위를 확대하고 실사 기준 및 공개 요건을 강화하며, 미국 규정에 맞춰 유럽 기업에게 일정 책임을 부과

금융 기관은 다음과 같이 4가지 기본 형태의 소프트웨어를 이용해 끊임없이 증가하는 AML 비즈니스 요건을 해결해왔습니다.

- 의심스러운 거래 패턴을 표시하는 **거래 모니터링 시스템(Transaction Monitoring System; TMS)**
- 거액의 현금 거래(미국의 경우 1만 달러 이상)를 보고하는 **고액현금거래보고제도 (Currency Transaction Reporting; CTR) 시스템**
- 고객 관계와 리스크를 명확하게 파악하는 **고객확인 의무(Customer Due Diligence; CDD), 고객알기제도(Know Your Customer; KYC) 시스템**
- 의심스러운 개인이나 조직을 식별하기 위한 **감시목록 스크리닝(Watchlist screening)**

현재 차세대 AML, 혹은 AML 2.0, AML 3.0 등 AML 솔루션 고도화 관련 논의가 활발히 이뤄지고 있습니다. 이러한 솔루션들은 로봇이나 의미 분석, 인공 지능과 같은 선진 기술에 바탕을 두고 있습니다.

차세대솔루션은 AML 프로세스의 효과와 효율성을 높이는 것은 물론, 기존의 규칙 기반 접근법을 보완하여 오탐율을 낮추고 조사 가치가 있는 활동을 더욱 정밀하게 탐지하는 데 큰 도움이 될 것입니다.

¹ 유엔마약범죄사무소, <https://www.unodc.org/unodc/en/money-laundering/globalization.html>, 2019년 3월 10일 조회

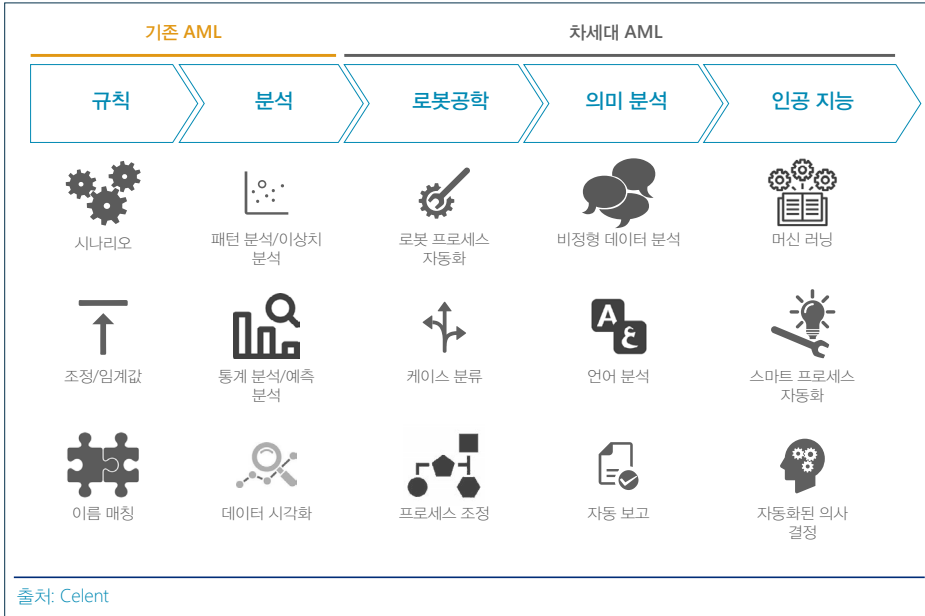


그림 1: AML 규정 준수 기술의 진화

여타 IT 분야와 마찬가지로 AML도 일단 단순 기능부터 시작해 고급 기능에 대한 검증을 해 나가며 단계적으로 진화하고 있습니다.

SAS에서 금융 서비스 보안 인텔리전스를 총괄하는 데이비드 스튜어트(David Stewart) 씨는 "지난 18개월간 이 분야에서는 주로 해결하기 쉬운 과제에 인공 지능을 적용해왔다"라며 "로봇 프로세스 자동화(Robot Process Automation; RPA) 기술을 이용해 케이스를 조사하고 속도를 높인 것이 그 예"라고 말합니다. "2018년 현재, 머신 러닝을 도입해 프로세스 자동화, 스코어링(scoring), 하이버네이션(hibernation) 등 다양한 용도로 활용하고 있으며, 의심스러운 활동을 탐지하기 위해 기존의 부울 논리를 보완하거나 아예 대체하기도 합니다."

세계 각지의 차세대 AML 솔루션 활용 현황

AML의 진화 속도는 지역에 따라 다른데, 규제가 엄격한 곳일수록 빠르고 규제가 느슨할수록 느립니다. 디지털 통화가 워낙 쉽고 빠르게, 또 교묘하게 국경을 넘나들다 보니 인공 지능 기반의 AML 기법에 큰 관심이 쏟아지고 있습니다.

아시아 태평양 지역: 관심은 커지고 있으나 제도적 장치가 미비함

FCC Analytics사 AML 프랙티스 리더인 월러스 차우(Walace Chow) 씨는 "홍콩의 금융 허브뿐만 아니라 아시아 태평양 지역 전체에서 AML이 급성장하고 있다"라며, "인공 지능을 갖춘 차세대 AML 솔루션을 도입하여 프로세스를 자동화하고 탐지 기술을 강화하고자 하는 SAS 고객이 많다"라고 강조합니다. 정부 기관은 은행에게 시나리오를 조정하거나 임계값을 설정할 때, 혹은 전체 AML 프로그램을 구성하는 다른 구성요소에 대해 보다 분석적인 방식을 사용하도록 요청하고 있으나 이에 대한 명확한 지침은 아직 내놓지 않은 상태입니다.

유럽: 발전하고 있으나 아직은 초기 단계

SAS에서 AML 분석 리더를 맡고 있는 아미스 새더쉬(Amith Satheesh) 씨는 AML 기술이 유럽에서 진화 중인 건 맞지만 아직 갈 길이 멀다고 지적합니다. 새더쉬 씨 설명에 따르면, 유럽 각지, 특히 발트 제국에서는 의심스러운 활동을 모니터링하고 탐지하는 데 집중하는 경향이 있다고 합니다. 근래 들어서 인공 지능과 머신 러닝이 화두로 떠오르고 있는데, 이는 뉴욕을 비롯해 미국 각지에서 영업하는 유럽 은행들에서 시작된 것이라고 합니다. 새더쉬 씨는 “현재 초기 발견 단계에서 응용, 구현 단계로 옮겨 가는 추세”라며, 실제 사용 사례와 더불어 일정 성과도 보고되고 있다고 설명합니다.

하지만 지금까지도 유럽 지역은 대체로 소극적인 태도를 보이고 있습니다. 아직은 기초를 정립하기에도 벅찬 상황이며, 이 때문에 유럽은 미국보다 10년 정도 뒤처져 있다고 새더쉬 씨는 지적합니다. “많은 은행들이 아직까지 흐름을 따라잡느라 급급한 상태이며, 세분화, 단일뷰, 계정 기반 모니터링에서 고객 기반 모니터링으로 이동하는 것, 임계값 설정, 효과적인 시나리오 구축, 오탐율 감소, 데이터 품질 문제 해결 등 기본적인 영역에 집중하고 있습니다.

따라서 예측 모델을 구현해 보다 나은 탐지 시나리오를 구축하려는 의지가 부족한 상태입니다. 하지만 조금씩 앞으로 나아가고 있습니다. 개념 검증과 더불어 활발한 논의가 진행 중이며, 사람들은 열린 마음으로 새로운 인공 지능과 머신 러닝 기법에 귀를 기울이고 있습니다.”

미국: 보다 엄격한 규제 환경 속에서 성숙 단계에 진입

SAS에서 미주 지역 AML 팀을 이끌고 있는 베스 허런(Beth Herron) 씨는 지난 10년간 미국 AML, 특히 거래 모니터링 분야에서 인공 지능과 머신 러닝이 급격한 성장세를 보였다고 설명합니다. ‘The Final Rule’ 시대가 도래하면서 AML 프로그램과 관련해 모델 리스크의 엄격함이 주목을 받게 됐습니다. 규제 준수 문화가 생겨난 것도, 미국의 분석 수준이 다른 시장에 비해 더 높아지게 된 것도 여기에 기인합니다.

“많은 사람들이 탐지 메커니즘에 주목하고 있으며 규칙과 시나리오가 조직이 직면할 수 있는 모든 리스크를 포착하는지에 대해 의문을 품고 있습니다. 샌드 박스나 분석 환경에서는 각종 실험, 즉 금융기관 자체적인 개념 검증 이 활발하게 진행되고 있으며, 이런 프로젝트 중 상당수가 실험 단계를 넘어 상용화 단계에 진입하여 성과를 보이고 있습니다.”

“일찍이 인공 지능 기술을 도입한 일부 선진 금융기관들은 파일럿 프로젝트를 상용화하며 큰 성과를 거두고 있습니다.”

베스 허런

미주 지역 AML 리더
Security Intelligence Practice, SAS

차세대 AML로 성공하기 위한 10가지 요인

1. 핵심은 신중한 혁신

지난 2018년 12월 미국에서는 연방준비제도 이사회를 비롯해 5개 연방 금융기관이 '자금 세탁과 테러 자금 조달을 막기 위한 혁신적 조치'를 요구하는 공동 성명서를 발표했습니다. 3페이지 분량의 성명서는 여러 가지 메시지를 담고 있었습니다. 성명서에서는 금융 기관이 AML 의무를 준수하기 위해 "혁신적인 조치를 고려, 검토하고 책임감 있게 이행해 나갈 것"을 권고하고 있습니다.

허런 씨는, "여기서는 혁신을 장려하는 어조와 함께, '책임감 있게'라는 말을 통해 이런 조치를 샌드 박스에서 시행하거나 당분간 기존 조치와 병행해서 시행하길 바란다는 뜻이 명확히 담겨 있다"라고 지적합니다. 물론 새로운 조치가 안정적이고 납득 가능한 기법이고 이를 통해 원하는 결과를 얻을 수 있다는 점도 입증되어야 합니다.

성명서에는 "시범 사업이 결국 실패로 돌아가더라도 관리·감독이라는 틀에서 은행을 비난하는 건 가급적 삼간다"라고 명시되어 있지만, 은행은 "가급적 삼간다" 보다는 더 명확한 표현을 원할 것입니다. 인공 지능을 활용하는 기법은 의심스러운 행위를 탐색하다가 멀쩡한 기존 시스템도 결함이 있는 것으로 인식할 가능성이 있기 때문입니다.

2. 엄격한 모델 거버넌스 구축

새더쉬 씨의 설명에 따르면 포괄적 AML 프로그램은 3중 방어선을 갖추고 있다고 합니다. 1차 방어선에서는 위험을 정량화하고 모니터링 대상 항목을 파악합니다. 2차 방어선에서는 제어 체계를 개발합니다. 3차 방어선에서는 이런 제어 체계를 시험해볼 수 있어야 합니다. 알고리즘이 계속 작동하는가? 케이스가 신속하게 처리되거나 하이버네이션 모드로 전환되는가? 각 모델을 모니터링하고 필요에 따라 조정하는가?

새더쉬 씨는 3차 방어선에 더욱 심혈을 기울여야 한다고 말합니다. "유럽은 이 분야에 더 엄격한 기준을 적용하여 혜택을 볼 수 있습니다. 현재로서는 지나치게 소극적입니다. 은행에 문제가 생기고 뉴스에 이런 소식이 보도되면 그제서야 모델 리스크 거버넌스에 관한 의문이 제기되고 있는 것이 현실입니다."

3. 국경 간 안전한 데이터 공유

영국 금융감독원은 여러 컨설팅 업체와 함께 해커톤(hackathon)을 열어 AML 분야에서 인공 지능과 머신 러닝을 효과적으로 활용하는 방안을 모색하는 한편, 국제 거래의 복잡성을 해결하고자 노력하고 있습니다.

스튜어트 씨는 "머신 러닝을 실질적으로 도입하는 것도 중요하지만, 그전에 먼저 GDPR(개인정보보호법)의 테두리를 넘어 데이터를 안전하게 공유할 수 있는 방법을 찾아야 합니다"라며, 이것이 해커톤을 진행하면서 모두가 함께 내린 결론이라고 강조합니다. 결국, 국경을 넘어 정보를 전송할 수 있는 동형 암호(homomorphic encryption) 같은 기술이 필요하다고 스튜어트 씨는 지적합니다. (동형 암호는 암호화된 데이터를 복호화 하지 않은 상태에서 바로 연산하는 기술을 가리키는데, 아직까지는 이론적인 수준을 벗어나지 못하고 있다.)

4. 하이브리드 방식 고려

규제 당국이 시인한 바와 같이, 위법 거래를 모니터링하는 기존 방식은 여러모로 부족한 부분이 많습니다. 규칙은 우리가 이미 아는 과거 패턴에 기반을 두고 있어서 시야가 좁고 한발 늦습니다. 규칙은 회피하기는 쉬워도 준수하는 것은 어렵습니다.

그런데 정작 이런 규칙 기반 시스템을 과감히 정리한다거나 분석 모델과 로봇으로 대체하려는 사람은 보이지 않습니다. 허런 씨는 하이브리드 방식을 언급합니다. 규칙만으로도 충분한 영역에서는 규칙을 사용하고 그렇지 않은 곳에서는 모델을 쓰는 방식입니다. 허런 씨의 설명처럼 패턴이 이미 잘 정의되어 있어 모델이 필요 없는 영역도 존재합니다. "이 영역은 충분히 제어 가능하다"라고 자신하는 경우도, 대상 지역이나 결과가 너무 희귀해서 전면 자동화하기 어려운 경우도 있을 수 있습니다.

분석 모델의 진가는 명확히 정의된 행동들 중에서 복잡한 패턴을 식별해야 하는 상황에서 드러납니다. 예를 들어 경보 스토어링이나 하이버네이션 모델은 단일 거래를 지켜보기만 하는 게 아니라 경고 패키지과 관련된 리스크까지 명확히 집어냅니다.

이런 방식을 활용하면 기존 거래 모니터링 시스템에서 더 많은 가치를 얻을 수 있다는 것이 허런 씨의 설명입니다. "멀쩡해 보이는 전신 거래(Wire Transaction)도 다른 행동이 눈에 띄면 수상쩍게 보이기 마련입니다. 실제로 개별 활동 대신 행동 프로파일이나 여러 활동을 살피는 경향이 눈에 띄고 있습니다. 이렇게 하면 케이스가 성공을 거둘 확률이 높아지거든요."

5. 철저한 데이터 기반 검토

데이터 사이언스는 본질적으로 'Garbage In, Garbage Out' 원칙에 기반합니다. 그런데 머신러닝이 등장하면서 이 원칙의 의미가 확장하고 있습니다. 시스템이 입력된 데이터를 학습해서 더욱 쓸모없는 데이터를 쏟아낼 수도 있기 때문입니다. 새더쉬 씨는 "인공 지능과 머신러닝에 질 낮은 데이터를 투입해봐야 얻을 게 없다"라고 말합니다.

데이터 관점에서 지난 10년간 진화했다고 할 만한 금융 기관은 찾아보기 어렵습니다. 적어도 유럽의 상황을 조망하는 새더쉬 씨가 보기엔 그렇습니다. "대부분의 은행들이 고객에 대한 단일 뷰를 갖추지 못한 실정입니다. 시나리오는 계정 수준에서 계속 경보를 발령하는데, 여러 명의 분석 담당자가 고객을 완전히 파악하지 못한 상태에서 경보를 제각기 분석하는 것입니다."

하지만 대형 은행이라면 상황이 다르지 않을까요? 새더쉬 씨는 "대형 은행, 특히 인수 합병으로 규모를 키운 은행은 규제 준수 데이터 허브에 들어가는 사일로화된 시스템과 데이터베이스가 그만큼 더 많으며, 이것이 다 연결되어 있는 것도 아닙니다."라며, "소스 시스템과 최종 목적지를 제대로 잇지 못해 데이터 품질에 문제가 생기는 경우가 허다합니다. 절대 쉬운 일이 아닙니다."라고 말합니다.

6. 임계값 설정에 대한 미시적 접근

개인 고객과 법인 고객을 나누는 것부터 고객 세분화(segmentation)가 시작됩니다. 기본적인 세분화 작업이 끝난 후 비즈니스 모델을 반영하기 위해 보다 상세한 분석이 필요해집니다. 표준편차를 이용해 법인 고객을 전체 거래 규모에 따라 소, 중, 대기업으로 분류하거나 상품이나 거래 유형별 세그먼트를 생성할 수 있습니다.

물론 더 정교한 방식도 있습니다. K-평균 군집화는 대표적인 비지도 머신 러닝 알고리즘으로서, 알려진 결과를 참조하지 않고 입력 변수에 근거하여 데이터에서 결과값을 추론합니다. 이 분석 기술을 통해 변수가 상호 작용하는 방식과 객체를 각기 다른 그룹으로 분류하는 방식을 알 수 있습니다. 그 후 시나리오를 확대하여 특정 클러스터에 어떤 리스크가 있는지 상세히 파악할 수 있습니다.

한 SAS 고객은 머신 러닝을 활용한 세분화 작업을 실시하여 생산성을 2.8%에서 6.8%로 높이고 오탐율과 경보 건수를 줄이는 데 성공했습니다. 세그먼트 별로 임계값을 조정하자 생산성률, 즉 조사 가치가 있는 경보 비율이 10.4 %로 대폭 상승했습니다.

7. 중요도에 따른 경보 조치

조사자가 모든 경보를 다 살펴보는 것이 정말 현실적일까요? 경보 중엔 그다지 심각하지 않은 경보도 있습니다. 만약의 경우를 대비해 이러한 경보를 안전하게 하이버네이션 처리할 수는 없을까요? 물론 가능합니다. 자동 추천이나 하이버네이션 기능은 광범위한 관련 정보를 활용하여 검토 작업에 앞서 경보 등급 상향 여부를 결정합니다.

차우 씨의 설명에 따르면, 경보가 발령되면 인공 지능 엔진이 리스크 점수를 산출하여 경보가 실제로 조사 가치가 있는지 확인한다고 합니다. 점수는 여러 위험 변수와 카테고리를 감안해 결정되는데, 점수가 매우 낮은 경보는 나중에 검토 가치가 있다고 판단될 때까지 하이버네이션 처리됩니다. "이런 기능 덕분에 조사자는 인력 소모를 크게 덜고 SAR(Suspicious Activity Reports)로 이어질 가능성이 있는 경보에 집중할 수 있습니다."

8. 머신 러닝으로 희귀 사건 탐지

모델은 비지도 학습을 통해 방대한 데이터를 분석하고 숨겨진 패턴을 파악하여 이상치를 발견합니다. 스투어트 씨의 설명에 따르면, 이 경우 특정 사람이 선의 또는 악의를 갖고 있는지 알기 어렵고, 단지 '경계 조건(edge case)'에 속하는 사람, 즉 일반인과 다르게 행동하는 사람을 찾을 수 있습니다.

허런 씨는 SAS가 "지도 학습 알고리즘을 이용해 17억 건에 달하는 거래 속에서 극히 일부에 해당하는 불법 행위를 잡아냈다"라며, "이 작업에 규칙과 지식을 적용하는 일은 어려운 작업이기에 기존에 없던 선진 기법을 동원할 수밖에 없었다"라고 말합니다.

"일단 타겟 변수로 FinCEN에 등록된 특정 고객 집단의 행동을 사용했습니다. 이 모델은 나머지 모집단에서 등록되어 있지 않지만 타겟 변수처럼 행동하는 사례를 탐색했습니다. 이 모델을 통해 불법 행위를 무더기로 잡아냈습니다."

감식 팀이 모래사장에서 '바늘'을 찾고 있다면 이 기법이 엄청난 위력을 발휘할 것입니다. 데이터가 아무리 많아도 리스크 요소를 감쪽같이 속아낼 수 있기 때문이죠. 기존 방식으로는 어렵습니다."

오탐인줄 알면서도 추적하시겠습니까?
시스템은 저위험 경보를 하이버네이션 처리한 다음, 케이스 관련 정보를 계속 수집하여 유의미한 정보가 발견되면 경보를 발령합니다.

9. 베스트 프랙티스 활용

스튜어트 씨는 파일럿 프로젝트를 진행하며 다수의 모범 사례들을 골라, 일명 '적응형 학습 및 지능형 에이전트 시스템(Adaptive Learning and Intelligent Agent System; ALIAS)'이라고 하는 패키지 봇(packaged bot)에 투입했다고 합니다.

이 기능은 머신 러닝 모델의 생성, 게시, 재학습 기능을 자동화합니다. 또, 데이터 사이언티스트와 컴플라이언스 분석 담당자에게 사전에 선정한 변수를 제공하고, 희귀 이벤트 샘플링 결과를 토대로 가장 적합한 모델을 추천해줍니다. 따라서 데이터 사이언티스트는 손쉽게 데이터를 더 효과적으로 분석하고 의심스러운 활동을 신속하게 탐지할 수 있습니다.

10. 금융 범죄 시스템과 프로세스의 통합

오늘날 금융 기관은 사기, 치안, 사이버 보안, 신용 리스크, AML 등 다양한 위험에 대응하고 있으며, 활용하는 시스템, 인력, 프로세스도 그만큼 다양합니다.

새더쉬 씨는 향후 이처럼 제각기 다른 요소들이 더욱 포괄적인 워크플로우에 통합될 것으로 내다보고 있습니다. 이는 'AML'이나 '사기'라는 각각의 관점 별로 고객을 보는 대신, 고객의 전반을 아울러본다는 뜻입니다.

앞으로는 데이터 조정은 물론 애널리틱스 개발, 의사 결정, 케이스 관리, 보고, 거버넌스까지 모두를 포괄하는 통합 환경에서 이루어질 것입니다.

"통합 환경이 구현되면 이런 방식이 더욱 효과적이고 경제적이라는 사실을 체감하게 될 것입니다."

“지금 우리는 개념 검증과 그에 수반하는 구현 작업을 거듭하며 실험을 진행하고 있습니다. 앞으로는 견고한 애널리틱스 기반을 확보한 Plug & Play 방식의 기능이 대세가 되리라고 생각합니다.”

아미스 새더쉬,
AML Analytics Lead
Security Intelligence Practice
SAS

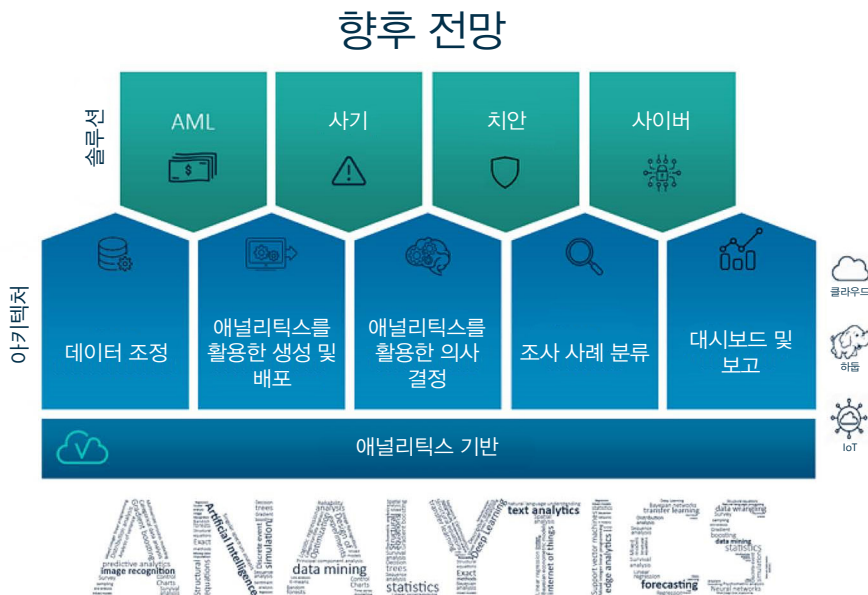


그림 2. 금융 범죄와 사이버 보안의 미래

AML 솔루션 활용: 국가별 사용 사례

지능형 자동화

아시아 태평양 지역에서 소매·상업 금융 서비스를 제공하는 한 은행은 여러 가지 난관에 직면했습니다. 거래량이 급격히 늘면서 모든 경보를 일일이 검토할 수 없게 된 것입니다. 오탐율이 늘어나 조사자의 부담도 커졌습니다. 그래서 이 은행은 중요도가 높지 않은 항목 위주로 검토 작업을 자동화하고 분석을 통해 숨겨진 위험을 탐지하고자 했습니다.

SAS는 조사 생산성을 높일 수 있는 케이스를 예측하기 위해 경보 스코어링과 하이버네이션 기법을 개발했습니다. 스코어링 알고리즘은 거래 리스크와 개체 리스크, 네트워크 리스크, 시나리오 리스크, 그리고 고객 리스크에 기반했습니다. 이 모델은 그래디언트 부스팅과 심층 신경망(딥러닝 알고리즘)을 사용한 앙상블 기법을 적용하여 경보 검토 프로세스를 100% 자동화하고 오탐율을 33%나 줄였습니다. 그 결과 조사 비용과 소요 시간이 크게 줄었습니다.

희귀 이벤트 탐지

한 글로벌 은행은 인공 지능을 적용해 숨겨진 위험과 오탐을 탐지할 수 있는지를 검증하고자 했습니다. 이 은행은 자체 규칙 기반 거래 모니터링 시스템이 알려진 리스크에 치중돼 있다는 사실을 인지하고 있었고, 은행이 고위험 거래에 대해 감시 폭을 확대할 것이라는 규제 당국의 기대 심리도 파악하고 있었습니다.

은행은 SAS 솔루션으로 대규모 데이터 세트를 조사하여 기존에 미처 발견하지 못했던 문제점을 찾고자 했습니다. 구체적으로는 은행의 고객 데이터베이스에서 고위험 고객이 제대로 분류되어 있지 않은 케이스가 있는지 확인하고자 했습니다.

스튜어트 씨는 트리 개수가 200개인 새로운 랜덤 포레스트 알고리즘을 적용하여 자금 서비스 업체(Money Service Business; MSB)이나 실제 온보딩 단계에서는 신고는 고사하고 등록조차 하지 않은 MSB 업체를 다수 잡아냈다고 밝혔습니다.

"20억 건에 달하는 거래를 조사한 결과(집계 소요 시간: 약 10분) 미심쩍은 MSB 416곳을 비롯해 알려지지 않았거나 등록되지 않은 MSB 89곳을 찾아냈습니다. 추가 분류 결과 조치가 필요한 케이스가 수십 개나 포착된 것입니다."

인공 지능 기반 탐지

미국에 있는 한 지역 은행은 운영 중이던 규칙 기반 AML 거래 모니터링 시스템을 현대화하고자 했습니다. 이 은행이 보유한 거래 모니터링 시나리오는 200개에 달해 관리가 어려웠습니다. 중요도가 떨어지는 이벤트 수를 줄이고 감시 폭의 차이를 좁히는 한편, SAR 전환율을 높이는 것이 관건이었습니다.

은행은 SAS 신경망 모델을 프로덕션 환경에 배포하여 거래 모니터링 시스템에 있는 현금 활동 시나리오 10개를 교체했습니다. 데이터 사이언티스트들은 이 모델이 미심쩍은 현금 활동을 탐지하는 데 75~80개의 변수를 사용한다는 것을 알게 되었습니다. 기존의 규칙 기반 시나리오가 6~12개의 변수를 사용하는 것과 비교하면 현격한 차이입니다.

스튜어트 씨는 은행이 SAR 전환율을 3배로 높인 결과 월간 처리 항목 수가 50% 줄었다고 말합니다. “은행 측은 일부 고위험 항목과 관련해 우선 규칙 기반 시나리오 200개를 25~40개 정도의 분석 기반 시나리오로 바꾸는 한편, 나머지도 향후 18개월에 걸쳐 모두 머신 러닝 전략으로 대체할 예정입니다.”

자연어 처리(NLP)

한 글로벌 은행은 고객확인업무 검토 속도와 정확성을 높여 무역 금융 기업 고객의 만족도를 높이고자 했습니다. 하지만 무역 서류 심사 담당자는 서비스 수준 협약서(Service Level Agreement; SLA)를 제대로 이행하지 못했습니다. 전체 서류를 다 검토하는 데 시간이 너무 많이 걸렸기 때문입니다. 자동화는 직원의 업무 부담을 줄이고 비용을 줄여줍니다.

이 은행은 딥러닝 알고리즘을 이용해 검토 중인 서류의 유형을 분류하고 맥락 분석 기법을 적용해 서류 등급을 지정했습니다(예: 과대·과소 청구 송장 파악). 이렇게 머신 러닝으로 작업을 자동화하자 2주 가까이 걸리던 검토 시간이 1분 미만으로 단축되었습니다.

맞춤말

금융 서비스 산업이 급격한 디지털 전환을 겪고 있고, 규제 당국이 ‘합리적’ 규제의 수위를 계속 높이는 가운데 차세대 AML이 떠오르고 있습니다. 로봇 공학과 의미 분석, 인공 지능, 특히 머신 러닝은 진화의 중심으로 자리매김할 것으로 보입니다.

머신 러닝은 단순히 과거 정보에 반응하는데 그치지 않고, 미래 지향적인 다양한 장점을 제공합니다. 머신 러닝을 이용하면 애초에 고려하지 않았던 데이터나 AML 모델로는 가용하지 않은 데이터 요소도 새롭게 찾아낼 수 있습니다. 컴퓨터가 알아서 사람의 육안으로 찾을 수 없는 패턴까지 밝혀냅니다. 사용자는 결과물을 검증한 다음, 모델에 다시 주입하기만 하면 됩니다. 학습하는 피드백 데이터가 늘어날수록 모델은 더 똑똑해지고 조정할 필요성도 그만큼 줄어듭니다.

물론 신뢰하던 규칙을 포기하는 게 쉬운 일은 아닙니다. 머신 러닝을 도입한다고 해서 기존 방식을 반드시 포기해야 하는 것은 아닙니다. 규칙과 예측 모델을 적절하게 결합한 보완적 방식도 생각해 볼 수 있습니다.

본 백서에서 예로 든 SAS 고객은 혁신을 추구하는 성향이 강하고 SAS와도 이미 몇 년째 협력을 이어가고 있습니다. 달라진 점이 있다면 소규모 기업의 진입을 가로막던 장벽이 낮아졌다는 사실입니다. SAS는 AML 업무에 맞는 검증된 머신 러닝 기술을 바탕으로 반복되는 수작업 공정을 자동화하고 미심쩍은 활동을 정확히 탐지합니다. 이제 보다 많은 금융 서비스 기업들이 SAS 솔루션의 이점을 누릴 수 있게 된 것입니다.

자세한 내용은 https://www.sas.com/ko_kr/software/anti-money-laundering.html에서 확인하시기 바랍니다.

시스템 내 정형 데이터 필드만을 토대로 결정을 내린다면 전체 가용 데이터의 20%도 안되는 수준으로 의사 결정을 내리는 셈이 됩니다. 원하는 수준의 정확도를 달성하려면 기존 SAR이나 고객 서비스, 통화 기록에 남아 있는 의견 등 여러 출처로부터 비정형 데이터를 반드시 가져와야 합니다.

금융 기관은 통합된 플랫폼에 수준 높은 애널리틱스와 강력한 머신 러닝 알고리즘을 적용하여, 위험 요인을 총체적으로 파악하고 더 많은 금융 범죄 패턴을 식별하며, 오탐율을 줄이고 조사 효율을 높일 수 있습니다.

해당 지역의 SAS 사무소 정보는 sas.com/korea/offices을 참조하십시오.

