INDUSTRY BRIEF

Using Machine Learning to Power Regulatory Compliance in Financial Services

Surface risks and maintain compliance efficiently and effectively

Introduction

The financial services industry has seen the evolution of machine learning technologies in compliance monitoring from their initial introduction, gradual adoption and now continued refinement. Initially introduced to the compliance monitoring space as a replacement for existing lexicon-based solutions, machine learning has come a long way.

Machine learning now offers clear advantages over purely lexicon-based solutions. However, it hasn't come without some challenges. Coming out of the research lab, machine learning technologies were typically optimized for analytic quality and not necessarily designed to meet the needs of a regulatory environment.

But we've hit an exciting milestone in recent years. By working together, applied machine learning engineers and compliance experts are identifying challenges and opportunities to continually refine machine learning-based solutions.

We've coined this refined approach, "regulatory-grade artificial intelligence." In this industry brief, we review the advancements, challenges and solutions that got us here.

More importantly, we highlight how large financial services firms now have proven and reliable tools to help them surface risks, maintain compliance and respond with agility to the ever-changing industry and regulatory landscapes.

The three generations of compliance analytics

To appreciate the tipping point of machine learning in financial services, we need to understand how it has evolved. We can view compliance monitoring analytics as having gone through three generations:

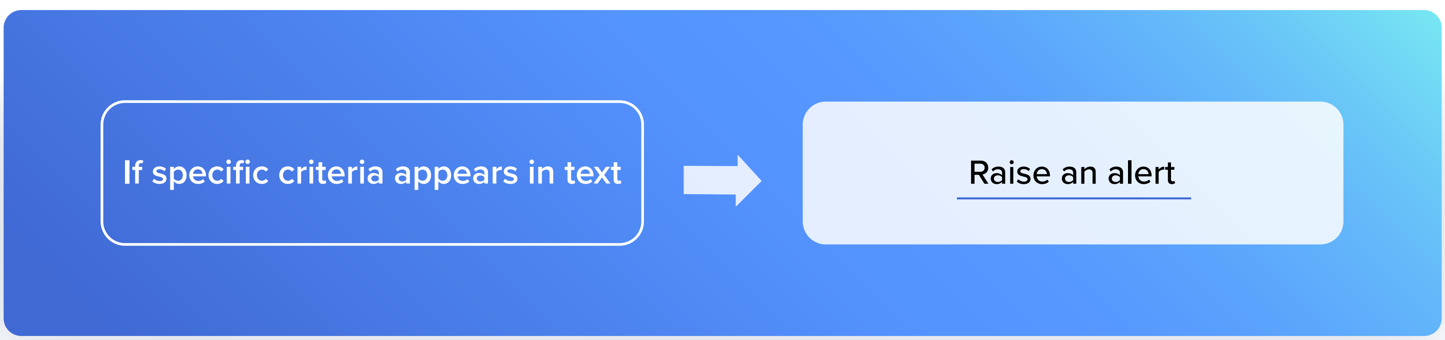
- First generation: Lexicon approach
- Second generation: Machine learning
- Current generation: Regulatory-grade AI

Let's explore how applied methods have evolved.

The lexicon approach

The first generation of conduct surveillance analytics relied on lexicons. Lexicons come in many shapes and sizes, from simple lists of keywords to complex collections of rules and patterns.

Whichever form they take, they use an "if-then" approach to raising compliance alerts:



The criteria (i.e., keywords, patterns, rules, etc.) are all defined by human subject matter experts, who often maintain and refine these lexicons over a period of years.

The lexicon approach has its limitations, which generally fall into two categories:

Too many false positives

Not all lexicons are created equal — some are carefully crafted to have much better precision than others. Compliance teams often complain that their lexicon-based policies alert on too much irrelevant content.

It's a tough problem. Real violations are rare, so most alerts are likely false positives. However, compliance teams also don't want to miss anything real. The result? Compliance teams are reluctant to narrow the scope of their lexicons too much, resigning themselves to following up on all alerts.

An inability to spot "unknown unknowns"

A lexicon can only raise an alert if it meets pre-defined criteria. In other words, a lexicon cannot "find" anything that a human has not explicitly told it to find. And humans can't define criteria to cover a situation that they don't know about or don't anticipate.

These "unknown" situations are not necessarily esoteric or rare. For instance, a misspelled or abbreviated word will foil a lexicon. Moreover, monitored employees will sometimes exploit this weakness intentionally if they know that keyword-based monitoring can be easily circumvented.

Analysts can always resolve these gaps as they find them by adding more complex rules to their lexicons. However, this maintenance process can be burdensome, and the risk of the "unknown unknowns" remains.

The machine learning approach

Then the research lab gave us the second generation of conduct surveillance analytics: machine learning.

In a way, machine learning models and lexicons are built for the same purpose. They both look at a given communication and decide whether to generate an alert. The difference comes in how they make this decision.

A lexicon uses a set of human-defined criteria to make a yes/no decision. A machine learning model generates its own set of criteria by looking at example data provided by humans.

The computer is far more effective at analyzing example data and creating optimized standards than human experts. This means the machine learning model ends up with a much better set of standards than the lexicon approach — and corrects the lexicon approach's two key challenges:

- A better set of criteria yields better analytical results: fewer false positive alerts and fewer missed alerts
- More nuanced criteria enable the machine learning model to overcome yes/no decisions: (0-100%), representing how likely the communication is to include a valid alert

For instance, a given chat message might score as 25% likely to have a valid market manipulation alert in the content, or an email might score as 85% likely to contain a customer complaint. This scored approach enables users to widen or narrow the aperture of their surveillance or focus their attention on specific bands of risk probability.

This probability score approach, coupled with the overall better analytic results, goes a long way to addressing the major limitations posed by lexicons. As a result, machine learning models are less susceptible to false positives and not as easily foiled by the "unknown unknowns," such as abbreviations or misspellings.

All of this sounds great, and study after study demonstrates the benefits of a machine-learning approach over traditional lexicon-based methods. But what happens in practice when we take this machine learning technology from the lab and apply it in real-world compliance monitoring?

Real-world machine learning challenges

Adopting machine learning in place of human-created lexicons essentially requires a trust fall into the arms of machine learning. We are asking analysts to give up the lexicons they have carefully crafted over the years and instead trust a "black box" of machine learning.

The analysts can no longer define (or in some cases even see!) the criteria used for alerting. And machine learning models can often produce unexpected results, leading analysts to wonder what the alerting criteria even is and undermining the analyst's trust in the model.

> **"If the model made this mistake, what other mistakes might it make?"**

This uncertainty is an uncomfortable place to be for an analyst working in a highly regulated environment. However, the possibility of looming audits casts a permanent shadow, and mistakes can have significant financial and legal impacts.

Let us assume for the moment that we have overcome this hurdle of trust. The next step is to refine the machine model so that it meets the unique needs of the compliance department in which it is being deployed. No organization is the same, having different regulatory needs, cultures, internal policies and more.

All of these differences need to be handled by the machine learning model. Knowing that the model will not work exactly how users want it to out of the box, the traditional machine learning response is to add more example data to the model until it learns the desired behavior.

We dub this the "monolithic model approach." Over the years, we've encountered challenges applying this approach in the compliance space.

Challenge example: Detect secretive behavior

The goal:
Compliance teams commonly monitor employee communications for secrecy behaviors because those are highly correlated with conduct concerns. Simply put, if people are being secretive, it may mean they are doing something they should not.

Proposing machine learning as the solution:
Machine learning seems like an excellent choice to find secretive behavior within conversations to detect phrases like, “Don’t mention this to anyone,” or “We can’t let anyone know about this.”

Machine learning greatly outperforms lexicons when it comes to finding this kind of nuanced human behavior in text.

Actual result of using machine learning:
When applying this model to real-world problems, we begin to see some issues:

1. The secrecy model thinks that all email disclaimer language is a secret

Text like, “The information contained in this email communication may be confidential. Do not distribute to unauthorized recipients,” quite reasonably seems like a secret to our model. This misunderstanding poses a significant problem in the field where most emails contain some disclaimer language, potentially resulting in thousands of useless secrecy alerts every day.

Addressing this problem using a traditional machine learning approach (again, the “monolithic model” approach), involves teaching the model that we want it to find secrecy, but not disclaimer secrecy. To accomplish this, we might add a few dozen or a few hundred samples to teach the model the new behavior we want.

Now, (hopefully) the disclaimer problem has gone away, and (hopefully) this work has not had any unintended consequences in our ability to find real secrecy.

It should be noted that even with these monolithic model problems, we do see the machine learning approach consistently outperform the legacy lexicon-based solutions in terms of the quality of alerts. However, this process is still frustrating for users and doesn't fully deliver on the promise of machine learning.

2. Analysts might want to tackle off-topic secrecy alerts

Alerts like, "It's a secret family recipe, so I can't share," or "Don't tell anyone I took the last cupcake, lol," are valid secretive language, but certainly not something that compliance teams are interested in seeing.

How do we teach the model that we want secrecy but not that kind of secrecy? Again, using the monolithic model approach, we might decide to feed those examples into the model as negative examples, thereby teaching the model to look for secrecy, but not disclaimer secrecy and not secrecy around topics like cupcakes or family recipes or surprise parties or Secret Santa or the secret to shiny hair, etc.

You can see that what began as a simple, clear concept of secrecy is now becoming quite complex. As we make the task more complex, we introduce more risk into our solution. The model may become confused and have a reduced ability to find the secrecy behaviors we do want.

The humans who are maintaining the model might also get confused. For example, they might have trouble remembering what "counts" as secrecy when providing examples to the model. As a result, they might provide conflicting training examples to the model, which degrades the model further.

3. In the compliance domain, machine learning models are not “black and white”

When it comes to compliance, machine learning’s greatest strength is also its greatest weakness. Machine learning doesn’t allow — at least not easily — an analyst to define deterministic “if-then” rules like they can with lexicons.

For instance, an analyst cannot tell a model, “If you ever encounter these exact words, always generate an alert.”

Instead, the machine learning model will learn its own rules and use them to determine probability scores. So, we might have a concerning email on which an analyst would definitely want to see an alert, but the model might assign a 79% probability of an alert. If the analyst has their alerting threshold set at 80% they won’t be notified.

Users appreciate the unpredictability of machine learning models when it means that the model returns something interesting that they hadn’t thought of before. But they don’t like that unpredictability when it means the model might not return something that they want.

In sum, machine learning models offer analytic quality improvements over lexicon-based solutions. Still, those quality improvements don’t negate users’ frustration and concern about lack of control, lack of explainability, and not having the confidence to really know for sure that the model will alert on certain phrases (all of which are critical in a regulated environment).

The regulatory-grade AI approach

Where does that leave us? How do we leverage the power of machine learning while at the same time providing the predictability, control, and explainability required in a regulatory environment?

At Smarsh, our answer is a new analytic framework that we call Smarsh Standard Scenarios. This approach represents all the insights we have learned in the field and the techniques we have developed in collaboration with compliance experts.

What we have done is decompose the problem into components. Instead of a monolithic model, we now leverage multiple components, each implemented with the most suitable approach, whether it is a machine learning model or a lexicon.

A Standard Scenario looks for a combination of signals to raise an alert and allows users to leverage both lexicons and models together, joined with Boolean logic and managed in a no-code user interface.

Traditional machine learning vs. Smarsh Standard Scenarios

The Traditional Approach	Smarsh Regulatory-Grade AI Approach
Monolithic model that is difficult to adapt • One model detects multiple types of risk • New use cases require retraining the model • Retraining the model can reduce accuracy	**Discrete models that are easily augmented** • Discrete models detect specific risk types • Augmented with lexicons and filters in cognitive scenarios • New use cases handled by augmentation, not model retraining • Enables greater accuracy even as use cases expand
Challenging to explain to regulators • Each refinement changes the model in a unique way • Explaining multiple iterations is difficult to impossible	**Easy to explain to regulators** • Cognitive scenarios built by Smarsh • No model training required • Scenario refinements handled in augmentation layer • Each augmentation is easy to explain • Augmentation layers are based on field-proven uses
Cumbersome to maintain • Every retraining requires internal MRM audit • Audits are time-consuming and costly	**Easy to maintain** • No re-verification with audit teams required • No need to pass model review boards repeatedly

Taking a step back, we know that a machine learning model is really good at finding nuanced human behavior in text, like the secrecy example above. What it is not good at is combining lots of other ideas into a single monolithic model, such as ignoring:

- Boilerplate disclaimers
- Surprise parties
- Secret recipes

It also doesn't accommodate the idea of deterministic rules (meaning, there are some cases where analysts will always want to see an alert).

With Standard Scenarios, we can let the machine-learning model do what it does best (find secrecy language), and allow other components (such as lexicons, rules, or additional models) to take care of the other tasks (such as filtering out disclaimers and non-work-related topics).

The Standard Scenario framework also empowers analysts to add their own deterministic rules and their own controllable, transparent lexicons on top of the models. This gives compliance teams the peace of mind that they will always get an alert when they know they want one.

Using this approach, we can leverage the power of machine learning while providing the predictability, control, and explainability required in a regulatory environment.

Standard Scenarios offer a pragmatic approach to enable analysts and compliance teams to leverage AI and machine learning and effectively own their risk in the field. At the same time, these tools won't encumber compliance teams with the ongoing maintenance of machine learning models.

Elevate compliance and risk management with Smarsh

Artificial intelligence and machine learning are continuing to provide value to financial services. They'll only continue to evolve with a pragmatic approach to surfacing risk and maintaining compliance in financial institutions.

Visit www.smarsh.com to learn how you can efficiently meet your supervision and surveillance needs.

smarsh®

Smarsh enables companies to transform oversight into foresight by surfacing business-critical signals in more than 100 digital communications channels. Regulated organizations of all sizes rely upon the Smarsh portfolio of cloud-native digital communications capture, retention and oversight solutions to help them identify regulatory and reputational risks within their communications data before those risks become fines or headlines.

Smarsh serves a global client base spanning the top banks in North America, Europe and Asia, along with leading brokerage firms, insurers, and registered investment advisers and U.S. state and local government agencies. To discover more about the future of communications capture, archiving and oversight, visit www.smarsh.com

Brief - 02/23

1-866-762-7741 www.smarsh.com @SmarshInc SmarshInc Company/smarsh