Vol.2019-CSEC-86 No.6
Vol.2019-SPT-34 No.6
2019/7/23

金融サービスにおける機械学習システムの適切な活用について：セキュリティと品質に焦点を当てて

清藤武暢† 宇根正志†

概要：日本銀行金融研究所では，金融業界が機械学習システムを活用する際に，セキュリティ対策や品質保証の観点から留意すべき点や課題について議論するために，第 20 回情報セキュリティ・シンポジウムを開催した．本稿では，同シンポジウムにおける講演やパネル・ディスカッションで示された意見を紹介する．

キーワード：機械学習システム，金融サービス，セキュリティ対策，品質保証

Appropriate Use of Machine Learning Systems in Financial Services: From Viewpoints of Security and Quality

TAKENOBU SEITO† MASASHI UNE†

Abstract: The Institute for Monetary and Economic Studies of the Bank of Japan held the 20th Information Security Symposium in order to discuss open issues from the viewpoints of security countermeasures and quality assurance in use cases of machine learning systems for the financial sector. This paper will present opinions expressed in the presentations and panel discussion of the symposium.

Keywords: financial service, machine learning system, security countermeasure, quality assurance

1. はじめに

近年，金融分野では，AI（artificial intelligence）を利用して既存の業務の効率化や新しいサービスの提供を検討する動きが活発化しており，そのコアとなる技術として機械学習（machine learning）が注目を集めている．もっとも，機械学習を実装したシステム（機械学習システム）には，機械学習に特有の脆弱性が存在するほか，品質（要件の充足度合い）の評価に際して従来のソフトウェア工学に基づく手法だけでは十分に対応できないケースが存在する．こうした状況を踏まえると，今後，金融業界において機械学習システムを活用するに当たり，セキュリティ対策や品質保証をどのように実施していくかは重要な課題となる．

こうした問題意識のもと，日本銀行金融研究所では，2019 年 3 月 27 日，「金融分野における機械学習システムの適切な活用に向けて」をテーマとして第 20 回情報セキュリティ・シンポジウムを日本銀行本店で開催した[1]．本稿では，今次シンポジウムで行われた講演やパネル･ディスカッションの内容を紹介する．

なお，本稿におけるシンポジウムの内容は，すべて著者たちの責任で取りまとめたものであり，日本銀行の公式見解を示すものではない．また，ありうべき誤りはすべて著者たち個人に属する．

† 日本銀行金融研究所情報技術研究センター
Center for Information Technology Studies, Institute for Monetary and Economic Studies, Bank of Japan

2. 第 20 回情報セキュリティ・シンポジウム

2.1 概要

第 20 回情報セキュリティ・シンポジウム（以下，単に，シンポジウムという）では，キーノート・スピーチ，3 件の講演，パネル・ディスカッションを行った．これらのタイトルや講演者・パネリストは以下のとおりである（敬称略）．各参加者の所属や役職名はシンポジウム開催時点のものであることに留意されたい．

- キーノート・スピーチ「金融分野における機械学習システムの適切な活用に向けて」（横浜国立大学教授　松本勉）
- 講演 1「機械学習システムのリスクとセキュリティ対策」（日本銀行金融研究所　井上紫織）
- 講演 2「機械学習システムの品質評価」（日本銀行金融研究所　清藤武暢）
- 講演 3「機械学習システムの品質保証ガイドラインの動向」（国立情報学研究所准教授　石川冬樹）
- パネル・ディスカッション「金融機関が機械学習システムを金融サービスで効果的に活用するための留意点や課題」
  - モデレータ：横浜国立大学教授 松本勉
  - パネリスト：国立情報学研究所准教授　石川冬樹
    日本 IBM 東京基礎研究所部長
    細川宣啓
    シティグループ証券株式／シティバンクエヌ・エイ東京支店　マ

Vol.2019-CSEC-86 No.6
Vol.2019-SPT-34 No.6
2019/7/23

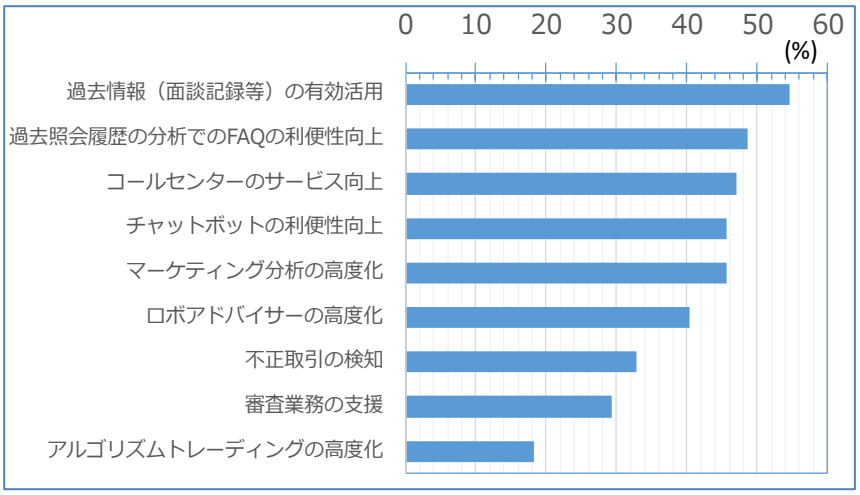

図 1　金融機関による AI 活用の目的と回答率
（備考）参考文献[2]を元に作成.

ネージング・ディレクター　オペレーション・テクノロジー　ヘッド　日高寛公

2.2 背景と問題意識：キーノート・スピーチ

まず，キーノート・スピーチでは，以下の主旨の発表が行われた.

近年，金融分野における AI の活用が広がりをみせている. 金融情報システムセンターのアンケート調査によると，「AI（自然言語処理，機械学習，ロボティクスといった要素技術を用いるもの）」を「導入中」,「準備段階」または「検討中」と回答した金融機関の割合は，2015 年度は 10%程度であったが, 2017 年度には 50%程度にまで増加している[2]. これらの金融機関による AI 活用の目的をみると，社内の過去情報の有効活用，チャットボットによる顧客応対の向上，マーケティング分析の高度化，不正取引の検知，審査業務の支援等，多岐にわたっている（図 1 参照）. 主要なクラウド・ベンダーが MLaaS の提供を開始するなど，金融機関が AI を活用しやすい環境が整いつつあることも，こうした動向の背景の 1 つになっているとみられる.

機械学習システムにはさまざまな構成が想定される. 今次シンポジウムでは，比較的シンプルな構成のシステムを前提に議論を行う. まず，訓練データを学習アルゴリズムに適用して訓練を実施し，判定・予測を行うソフトウェア（判定・予測モデル）を生成する. 機械学習システムを利用する際には, 判定等を行うデータをシステムに提示する. そうすると，そのデータが判定・予測モデルに入力され，そのモデルの出力に基づいて判定・予測結果が示される.

金融機関における顧客応対向けのチャットボットの利便性を向上させるために機械学習システムを活用する事例では，照会対応のノウハウや金融商品に関する情報等を訓練データとするほか，顧客からの照会情報を判定・予測モデルに入力しその出力に基づいて顧客への回答を提示する. 審査業務の効率化を企図して機械学習システムを活用する事例では，金融機関が有する顧客の信用度にかかる情報等を訓練データとして判定・予測モデルを生成したうえで，顧客が提示した審査に必要なデータを判定・予測モデルに入力し，その出力に基づいて審査結果を生成する. また，不正な金融取引の検知に機械学習システムを用いる事例では，まず，金融取引にかかる既存のデータを訓練データとして判定・予測モデルを生成する. そのうえで，不正か否かの判定の対象となる取引のデータを判定・予測モデルに入力し，その出力に基づいて当該取引が不正か否かを判断する.

新しい技術を金融分野で活用する際には，予めそのセキュリティ上のリスクを考慮し，対策を適切に行うことが求められる. この点，機械学習システムには，機械学習に特有のリスクが存在することに留意が必要である.

機械学習システムの品質をどう確保するかも重要な課題である. 従来の IT システムでは，その振舞いを概ね把握可能であり，一定の品質を保証することができた. 一方，機械学習システムでは，その振舞いを事前に把握することが困難であり，従来のソフトウェア工学による手法のみでは対応が難しい場合がある. その結果，有用な技術であっても品質を保証できず，技術の差別化を図ることが困難となりうるほか，想定外の事故が発生した際に無過失であることを説明できないというリスクが発生しうる.

こうした課題に対応するための取組みが各所でなされている. 例えば，産業技術総合研究所サイバーフィジカルセキュリティ研究センターでは，産業界と連携しつつ，機械学習システムの品質にかかる基準の策定と品質を確認・検査する手法等の研究開発を進めている.

今次シンポジウムでは，機械学習システムのセキュリティ対策や品質保証にかかる研究開発動向を把握するとともに，金融サービスにおいて機械学習システムを適切に活用するうえでの留意点や課題等を議論していきたい．

2.3 各講演の概要

(1)機械学習システムのリスクとセキュリティ対策

講演 1 では，[3]に基づき，機械学習システムに特有の脆弱性，金融サービスで活用される機械学習システムにおいて想定されるリスクとセキュリティ対策のあり方について，以下の主旨の発表が行われた．

機械学習システムにおいては，通常の IT システムと同様に，そこで取り扱われるデータ（訓練データ等）やシステムを構成するソフトウェア（学習アルゴリズム，判定・予測モデル）等が保護の対象となり，それらの機密性・一貫性・可用性の確保がセキュリティ目標となる．これを実現する手段として，通信路上のデータを暗号化する，各エンティティが有するデータやソフトウェアへのアクセス制御を実施するなどの一般的なセキュリティ対策を利用することができる．

もっとも，機械学習システムの場合，そうした対策のみでは対応が困難な脆弱性が存在する点に留意する必要がある．例えば，判定・予測モデルの入出力から，訓練データや判定・予測モデルにかかる情報が漏洩する場合がある．また，判定・予測モデルへの入力に微小なノイズが加わると，その（ノイズ付きの）入力に対して誤った判定・予測が出力される場合があるほか，訓練データにノイズが加わると，判定・予測モデルの精度が大きく低下しうる．

各脆弱性への対応の要否は，脆弱性が顕在化した場合の影響の大きさに基づいて判断することになる．訓練データや判定・予測モデルにかかる情報の漏洩に対しては，訓練データの機密性や判定・予測モデルの資産性が判断材料となる．判定・予測の精度低下に対しては，誤った判定・予測によるリスクの多寡が判断材料となる．

顧客の照会にチャットボットを用いて応答する機械学習システムにおいて，訓練データとして公開情報（一般的な照会とそれに対する回答等）が用いられる場合，訓練データの機密性は低く，判定・予測モデルの資産性も高くないと想定される．一方，判定・予測モデルの精度低下により，多くの顧客が誤った回答を受信すれば，金融機関のレピュテーションが低下する可能性がある．こうしたリスクを無視できないならば，何らかの対策を講じる必要がある．

スマートフォン・アプリ等を用いて顧客の信用度を評価するシステムでは，訓練データとして，年齢や収入，勤務先等，個人情報が含まれる場合があるほか，システム自体の資産性も高いと考えられる．したがって，訓練データや判定・予測モデルの漏洩への対策を講じる必要がある．判定・予測モデルの精度低下についても，顧客の信用度の不適切な評価は本来よりも緩い条件での貸出の実行等に繋がりうるため，対策の検討が求められる．

訓練データ等にかかる情報の漏洩に対しては，判定・予測モデルの出力を変換して提示したり，判定・予測の確からしさを示す値（確信度）を丸めて提示したりするなど，推定に必要な情報を攻撃者が入手できないようにすることが考えられる．また，個人情報が訓練データに含まれる場合，個人を特定できないように加工するなど，推定時の影響を軽減することも対策方針として挙げられる．さらに，訓練データの推定が困難な学習アルゴリズムの採用も有用である．

判定・予測モデルの精度低下に対しては，誤った判定・予測を誘発する入力等を事前に検知可能な判定・予測モデルを別途生成して利用するという手法が挙げられる．また，誤った判定・予測を誘発する入力の影響を低減させる学習アルゴリズムを採用する手法も対策として挙げられる．

攻撃手法や対策手法は日々進化している．金融機関が機械学習システムのリスクを把握しセキュリティ対策を検討する際には，最新の研究動向を注視することが必要である．また，いったん導入したセキュリティ対策に関しても，その効果が失われていないかを定期的に確認し，必要があれば対策の内容を見直すことが肝要である．

(2)機械学習システムのソフトウェアの品質評価

講演 2 では，[4]に基づき，機械学習システムの品質評価にかかる課題とそれへの対策手法の研究動向，金融分野で活用される機械学習システムの事例において品質を評価し保証する際の留意点や課題について，以下の主旨の発表が行われた．

機械学習システムの主たる機能は判定や予測の実行であり，判定等を行うデータに対して，期待される判定・予測結果が一定以上の確率で得られることが求められる．機械学習システムの実用に供する際には，こうしたビジネス要件の充足度合いを事前に確認しておくことが望ましい．期待される判定・予測結果が一定以上の確率で得られるという特性は，機能正確性と呼ばれ，従来のソフトウェアにおける品質特性の 1 つとして知られている．機械学習システムについて，どのような品質特性を設定するか，設定した品質特性をどう評価するかが課題となっている．

従来のソフトウェアは，通常，人間が期待する入力と出力の関係を定式化したうえで，処理の流れを明確化して生成される．したがって，その品質の評価は，実際の入出力が定式化した関係と整合的か否かを確認することによって行われる．こうした評価の方法論がソフトウェア工学として長年研究されているほか，JIS X 25010 等の標準規格が策定されている．一方，判定・予測モデルについては，訓練データと学習アルゴリズムを用いて直接生成され，入力と出力の関係が必ずしも明確ではないため，従来の方法論に

よる評価が難しい場合がある．

近年，こうした課題に対応するための研究が活発化している．入力と出力の関係が把握できない場合であっても，判定・予測モデルの特性を利用することによって，品質評価に用いるデータ（テストデータ）を生成する手法が複数提案されている．例えば，判定・予測モデルに入力するテストデータの一部を変化させた場合の，出力の変化の方向性が明確であるケースがある．こうしたケースでは，入力を変化させたときの実際の出力における変化が予想した方向か否かを確認することで，機能正確性を評価することが考えられる（こうした手法はメタモルフィック手法と呼ばれる）．

顧客の照会にチャットボットを用いて応答する機械学習システムにおいては，音声やテキストによる照会に対して，従来のシステムと同程度に正確な情報をより効率的に回答することが主たる目的となる．例えば，正確な照会結果を提示するというビジネス要件に関しては，顧客からの照会内容が一部変化した際にチャットボットの回答がどう変化するかをメタモルフィック手法等によって評価することが考えられる．また，照会結果の根拠を顧客に提示するというビジネス要件を設定する場合には，その根拠を人間が解釈しやすい形式で出力する学習アルゴリズム等を活用することが有用である．根拠の提示が困難であるならば，金融機関の職員への直接の照会・問合せを推奨することが考えられる．

スマートフォン・アプリ等を介して顧客の信用度を評価する機械学習システムにおいては,顧客の属性等に応じて，従来のシステムと同程度に正確かつ公平な信用度をより効率的に評価することが求められる．特に，信用度評価の結果が公平性を満たすこともビジネス要件の 1 つとして設定されうる．特定の顧客が不利益となる（公平性を損なう）データが訓練データとして使用されていないことを確認する，あるいは，判定・予測モデルの出力の偏りを検知・排除する手法（フェアネス・アウェア・データマイニングと呼ばれる）を活用するなどの対応が挙げられる．

今後，機械学習システムの品質保証を実現していくうえで，判定・予測モデルの品質評価にかかる新しい手法を取り入れることを視野に入れつつ，運用面での対応を検討していくことが重要である．また，最近では，機械学習システムの品質保証にかかるガイドラインの策定が進められていることから，それらを活用することも有用であろう．

(3)機械学習システムの品質保証ガイドラインの動向

講演 3 では,機械学習システムの開発や品質保証の現状や,機械学習システムの品質を評価するためのガイドライン策定に向けた取組みについて,以下の主旨の発表が行われた．

通常の IT システムは，計算や判断を行うための知識・規則を人間が決定し，それを実現するプログラム（判定・予測モデル）を作成するという流れで開発される（演繹的システム開発）．一方，機械学習システムは，知識・規則を人間が決定するのではなく，訓練データから獲得してプログラムを作成するという流れで開発される（帰納的システム開発）．そのため，開発された機械学習システムの特性や限界をエンジニアですら把握困難な場合がある．

こうした課題に対処するためには，工学的な視点が不可欠である．日本ソフトウェア科学会では，2018 年に「機械学習工学研究会」を設置し，技術者や研究者による研究発表や議論の場を提供している．先般，機械学習を業務に用いている技術者等を対象にアンケートを実施したところ，機械学習システムの開発における「顧客との意思決定」や「テスト，品質の評価・保証」において，通常の IT システムの開発とは異なる対応が必要であるとの回答が多かった．機械学習システムの場合，どれだけテストを実施すれば十分かが不明確な場合が多く，不確実性が高いシステムを発注者やユーザーがどの程度受け入れられるかが大きな課題となっている．上記のアンケート結果はこうした問題を反映しているといえる．

機械学習システムの主要ベンダーでは，品質保証にかかる原則や指針を独自に策定している．そうした指針においては,例えば,機械学習システムの入出力の傾向を分析し，問題として顕在化する前に不適切な入出力を検知することや，入出力の範囲や分布が予想と合致しているかを評価することなどが盛り込まれている．これらは，機械学習システムの振舞いが事前に想定したとおりとなっているか否かをチェックするという考え方である．わが国においては，AI プロダクト品質保証コンソーシアム（QA4AI）が，機械学習システムの品質保証に関する技術の調査研究やその体系化，機械学習システムの活用の支援等を行っており，品質保証に関するガイドラインの策定も進めている．

QA4AI のガイドラインでは，機械学習システムの品質を評価するうえで重要となる 5 つの概念を定める予定である．まず，①訓練データと判定・予測モデルの入力の整合性等の「データの品質」や，②正解率等の性能，学習アルゴリズムの妥当性等の「判定・予測モデルの品質」が挙げられる．また，③機械学習システムの価値，インシデントの発生度合いとリスク,説明可能性等の「システム全体の品質」，④品質向上のためのチェックを行う周期の短さ等の「プロセスの迅速さ」が挙げられる．さらに，⑤顧客が機械学習システムへ期待する程度や品質・リスクに関する理解の度合い等の「顧客による期待の高さ」も重要な要素であることから，ガイドラインに含まれる見込みである．

機械学習システムの品質保証にかかる研究開発のスピードは非常に速くなっており，新しい品質評価の手法も次々と提案されている.ガイドラインの策定に向けた取組みは,QA4AI 以外の組織や団体においても活発化しているが，こうした状況を踏まえると，当面，これらガイドラインは頻

繁に更新されるであろう．今後，機械学習システムの活用を検討していく際には，こうしたガイドライン等を参照しつつ，機械学習システムの特性や限界を十分理解するとともに，活用の目的についてビジョンを明確にすることが重要である．

2.4 パネル・ディスカッション

パネル・ディスカッションでは，機械学習システムにおける品質保証，機械学習システムの活用の範囲や開発形態，機械学習システムに関連する国際標準化の動向等について，議論が行われた．以下では，各論点に関するパネリストやフロア参加者による主な意見やコメントを示す．

(1)金融分野で活用される機械学習システムにおける品質の考え方

- 金融サービスに関連する IT システムの品質として，長期間の安定稼動の実現が重視されるケースが多い．機械学習システムにおいても，重大なインシデントが発生しないことなどを品質として評価していくことになるのではないか．
- インシデントが発生しないことを保証するための対策について，チャットボットを用いた金融取引のサービスのケースでは，例えば，顧客に無断でチャットボットが資金移動を行わないように，資金移動の際には必ず顧客が最終確認を行う設計とするなどの対応が考えられる．
- 機械学習システムによる判断の結果が一定の基準を超えた場合に，そのシステムが自動的に停止する機能を実装することや，人間がシステムを強制的に停止できる仕組みを導入することが考えられる．
- 機械学習システムでは，自分の動作を自律的に停止させる機能を実装することは難しいとの研究結果もあり，機械学習システムを停止させるためには人間の介入が必要である．
- 現在の機械学習の技術では，すべての判断の基礎となる「方針」をも適切かつ自律的に学習することが困難である．
- 誤った判断によるインシデントの発生は，人間が判断する際にも起こりうる．しかし，機械学習システムにおける対策の必要性ばかりが過剰に意識されてはいないか．
- 金融機関では，人間の判断についても誤りがないかを慎重に確認し，問題がある場合には業務のプロセスを停止する体制を整備している．金融機関における業務のプロセスをすべて機械学習システムによって実現するとすれば，既存のプロセスと同程度の安全性や信頼性を維持するための対策を講じる必要がある．

(2)機械学習システムの活用の範囲

- **既存の業務の代替可能性**
  - 代替可能性を考えるうえで，機械学習システムの機能だけでなく，そのシステムを用いたサービスを顧客がどう受け止めるかという視点も重要になる．例えば，銀行の窓口業務において，すべての応対を機械学習システムが行う場合よりも，銀行員が笑顔で接した場合の方が，顧客満足度が高く，結果として金融商品の売上が増えるというケースが考えられる．
  - 現時点では，人間が判断する際の補助機能として機械学習システムを活用することが現実的である．
  - すべての業務を機械学習システムに代替させる場合，トラブル発生時の責任の所在が不明確になる可能性がある．
- **誤った判断にかかる責任の所在**
  - （フロアからの質問）機械学習システムを人間の判断の補助機能として活用する場合であっても，その出力が引き金となって誤った判断が発生するケースが想定される．この場合の責任の所在をどう考えればよいか．
  - 金融機関内部における責任の所在については，機械学習システムの導入を判断した部門を中心に，システム部門やコンプライアンス部門等，社内の複数の部署が責任を負うことになる可能性がある．
  - 発注者と開発者の間の責任分担について，機械学習システムを開発する際には，発注者と開発者が緊密に連携しながら共同で作業を進めていくケースが多い．したがって，発注者と開発者が共同で責任を負うことになる場合がある．
  - 機械学習システムのリリースを優先した結果，最終的にビジネス要件が満たされなかったり，リリース後に性能劣化が発生したりしたケースでは，発注者と開発者の責任分界点を定めることが難しくなる．
- **機械学習システムを採用するか否かの判断の基準**
  - （フロアからの質問）機械学習システムを採用した結果，誤った判断によって損害が発生したという事例が生じうる一方で，社会全体として捉えればトータルの損害よりも利益が上回るという状況がありうるが，こうした場合に機械学習システムを活用すべきか否かをどのように判断すればよいか．
  - 機械学習システムの活用を社会的な潮流と捉え，リスクも含めて積極的に活用するという考え方がある一方，誤った判断によるリスクを重視して活用しないという考え方もある．個人がそれぞれのリスク選好に基づき判断していくことになるのではないか．
  - 機械学習システムを活用することによるリスクを踏まえつつ，どのような用途で活用していくかを予め明確にしておくことが必要であり，それに基づいて判断することになる．

- 機械学習システムが少子化に伴う人手不足等の社会問題を解決する手段の 1 つとして注目されており，その積極的な活用が社会のトレンドとなっている．
- 誤った判断を下すリスクは人間の場合でも存在するが，機械学習システムにおいてそうしたリスクをいかにして低下させるかを考えることが重要である．

● **誤った判断のリスクを低下させる手法**

- 機械学習システムが誤った判断を下すリスクを低下させるうえで，発生しうるインシデントを予め網羅的に洗い出すことができれば有用である．
- インシデントの洗出しは可能であり，実際に自動運転の分野においては，そうした対応の必要性を指摘する声が聞かれている．
- 複数の機械学習システムを用いて，異なる観点から総合的に判断することが有用ではないか．
- そうした手法の 1 つであるアンサンブルでは，判断にかかる責任の所在が不明確になるという課題がある．例えば，融資判断を行う機械学習システムでは，システムが，融資可能な金額だけでなく，融資に伴うリスクに関する情報を出力するなど，判断に関連する付加的な情報を出力する機能を実現する方が有用ではないか．
- 金融分野では，判断を行う際に合理性や論理性が重視される傾向にある．判断の根拠が明確でない機械学習システムを複数用いたとしても，判断にかかる合理性や論理性を明確にすることは困難である．合理性があると認められる判断の候補を機械学習システムが複数提示し，最終的な判断を人間が選択するというアプローチがありうる．
- 機械学習における重要な要素として，現時点では（学習の対象となる）データが中心的な位置を占めているが，今後，人間とのインタフェースも重要になってくるであろう．例えば，曖昧な回答や（二者択一ではない）中間的な回答等，より高度な意思決定を支援する機械学習システムをどう実現するかといった点も重要な課題である．

(3)機械学習システムに関する標準化等の動向

- 近年，機械学習システムの品質評価において，人間が不快に感じるか否かといった人間の感性に関わる要素が重要であるとの認識が広がりつつある．例えば，公平性（fairness），説明可能性（accountability），透明性（transparency）を充足する機械学習システムが注目を集めており，その開発や評価に関する研究が活発化している．
- AI の国際標準化を担当する ISO/IEC JTC1/SC42 において，AI の信頼（trustworthiness）に関する検討を行う作業部会が設置された．同作業部会では，AI の信頼に関連する問題の抽出・整理や事例の収集等が行われており，収集した事例をもとに国際標準にかかる審議が進められている．
- 中国や米国等では，機械学習システムの運用開始時点で十分な品質の確保を要求されるわが国と異なり，問題が発生する都度，改善しながら機械学習システムを運用するケースが多い．QA4AI のガイドラインの策定等を通じて，わが国で通用する品質保証を確立することができれば，他の産業分野と同様に，高い品質が機械学習システムの分野における国際競争力の面でのわが国の強みとなるのではないか．

(4)機械学習システムの開発形態

- 金融機関が有するデータは機密性の高いデータが多く，それらを社外のベンダー等に渡して判定・予測モデルを構築することは容易でない．
- 判定・予測モデルの生成を外注したいというニーズは少なくないものの，セキュリティ上のリスクに配慮した結果，訓練データを十分に用意できず，生成した判定・予測モデルの精度が低く，実用に耐えないといったケースもある．
- 比較的少ない訓練データを活用するケースとして，製造業の不良品検出の分野では，自社の数十件の訓練データを既存の判定・予測モデルに適用することによって，自社の訓練データに適合した判定・予測モデルを生成することができたという事例がある．
- 欧州では，EU 一般データ保護規則によってデータ保護の動きが広がっているものの，本人の同意があれば，個人から取得したデータの二次利用や転売が認められており，自社で大量のデータを保有していない企業であっても，他社からデータを購入して活用することが可能である．
- 一部の企業がサーバ上で提供している学習アルゴリズムを用いることで，判定・予測モデルを自社開発することもできる．
- 判定・予測モデルを他社に提供することに伴うセキュリティ上のリスクに関して，機械学習システムを提供するクラウドサービスの入出力から，その判定・予測モデルの推定に成功した事例が報告されている．
- 正規の判定・予測モデルに電子透かしを埋め込んでおくことで，そのモデルを推定して生成された不正な判定・予測モデルを判別する技術等，こうした攻撃に対する対策も研究されている．

3. おわりに

第 20 回情報セキュリティ・シンポジウムでは，3 件の講演において，機械学習システムのセキュリティや品質保証にかかる最新動向が紹介されたほか，パネル・ディスカッ

ションでは，金融分野での活用を展望した際の課題や技術的な限界についてさまざまな意見が示された. 現時点では，機械学習システムは，既存の金融機関業務をサポートするツールとして位置づけられているが，今後の本格的な活用を展望していくうえで，セキュリティの確保や品質保証は避けて通れないテーマである．パネル・ディスカッションにおいて提示された課題や限界を十分認識したうえで，機械学習システムの適切な活用のあり方を検討していくことが望まれる．品質保証ガイドラインの策定や国際標準化の動向等も引き続きフォローするとともに，今回示された課題への金融機関における対応等に注目していきたい.

参考文献

[1] 日本銀行金融研究所,「情報セキュリティ・シンポジウム（第20回）の模様：金融分野における機械学習システムの適切な活用に向けて」，IMES Discussion Paper Series，no. 2019-J-11，日本銀行金融研究所，2019年.
[2] 金融情報システムセンター,「平成30年度金融機関アンケート調査結果」，金融情報システム，no.345，金融情報システムセンター，2018年，1〜200頁
[3] 井上紫織・宇根正志,「金融分野で活用される機械学習システムのセキュリティ分析」，IMES Discussion Paper Series，no. 2019-J-1，日本銀行金融研究所，2019年.
[4] 宇根正志・清藤武暢,「機械学習システムにおけるソフトウェアの品質評価の現状と課題」，IMES Discussion Paper Series，no. 2019-J-6，日本銀行金融研究所，2019年.

付録　金融機関の実務者からのアンケート

シンポジウム当日の参加者(約100名)を対象にアンケート（無記名，所属組織の業態のみを選択）を実施し，金融機関の実務家から36件の回答を得た（全体では89件）.

主な質問事項は，今後の情報セキュリティ・シンポジウムで取り上げてほしいテーマを問うもの（質問**イ**），足許の情報セキュリティ上の課題を問うもの（質問**ロ**），金融サービスを提供するシステムにおいて先行き攻撃対象となりうる部分を問うもの（質問**ハ**）である．各質問の内容は以下のとおりである.

- 質問**イ**：今後シンポジウムで取り上げてほしいトピックを選択肢から3つ以内でお選びください.
- 質問**ロ**：情報処理推進機構による「情報セキュリティ10大脅威2019」における各脅威について，貴社においても同様の課題があると思われる項目や，影響が大きいと思われる項目を選択肢から 3 つ以内でお選びください.
- 質問**ハ**：今後，貴社においても脅威となりうると考えられる項目（金融サービス等を提供する情報システム全体における攻撃箇所に着目した整理）を選択肢から3つ以内でお選びください.

質問の選択肢や回答の集計結果を表 A-1 に示す．主な結果を整理すると，以下のとおりである.

(1)今後取り上げてほしいトピック：FinTech

質問**イ**では，「FinTech（暗号資産，オープンAPI等）（**イ-1**）」が最も高い回答率（56%）であった．これに次いで高い回答率（44%，33%）となったのが，それぞれ，「サイバー攻撃（**イ-6**）」と「認証技術（生体認証等）（**イ-8**）」であった.

(2)2019年のニュース：標的型攻撃による被害

質問**ロ**では，「標的型攻撃による被害（**ロ-1**）」が最も高い回答率（47%）となった．次いで，「インターネットサービスからの個人情報の窃取（**ロ-7**）」と「ビジネスメール詐欺による被害（**ロ-2**）」が高い回答率（それぞれ28%，25%）となった.

(3)今後の脅威となりうる対象：顧客の端末，社員の端末

質問**ハ**では，「顧客の端末（PC，スマホ）への攻撃（**ハ-1**）」と「社員の端末（行員のPCやタブレット等）への攻撃（**ハ-5**）」が，最も高い回答率（39%）となった．次いで，「金融機関対外接続系システム（WEBサービス，インターネットバンキング等）への攻撃（**ハ-2**）」と「金融機関情報系システム（電子メール，EPR 等）への攻撃（**ハ-3**）」が共に高い回答率（36%）となった.

表 A-1　シンポジウムでのアンケート集計結果

			金融機関 (36)	ベンダー (32)	大学・研究機関等(4)	その他 (17)	全体 (89)
イ．今後取り上げてほしいテーマ（数字は回答割合）							
	イ-1	FinTech（暗号資産、オープンAPI等）	56%	69%	75%	53%	61%
	イ-2	インターネット・バンキング	11%	13%	25%	12%	12%
	イ-3	リテール金融取引（QRコード決済等）	22%	34%	25%	24%	27%
	イ-4	IoT機器	17%	34%	75%	41%	30%
	イ-5	クラウド	22%	22%	25%	24%	22%
	イ-6	サイバー攻撃	44%	34%	25%	24%	36%
	イ-7	暗号技術（高機能暗号、暗号の移行等）	19%	25%	25%	18%	21%
	イ-8	認証技術（生体認証等）	33%	25%	25%	18%	21%
	イ-9	モバイル端末	19%	3%	25%	29%	16%
	イ-10	分子・DNAを活用する情報技術・コンピューティング	14%	16%	0%	18%	15%
ロ．情報セキュリティ10大脅威2019で同様の課題があると思われるものは？（数字は回答割合）							
	ロ-1	標的型攻撃による被害	47%	28%	75%	53%	43%
	ロ-2	ビジネスメール詐欺による被害	25%	16%	0%	47%	25%
	ロ-3	ランサムウェアによる被害	11%	16%	50%	24%	17%
	ロ-4	サプライチェーンの弱点を悪用した攻撃の高まり	8%	22%	25%	18%	16%
	ロ-5	内部不正による情報漏えい	19%	19%	25%	12%	18%
	ロ-6	サービス妨害攻撃によるサービスの停止	19%	25%	0%	24%	21%
	ロ-7	インターネットサービスからの個人情報の窃取	28%	25%	25%	29%	27%
	ロ-8	IoT機器の脆弱性の顕在化	6%	34%	25%	24%	20%
	ロ-9	脆弱性対策情報の公開に伴う悪用増加	8%	34%	25%	24%	0%
	ロ-10	不注意による情報漏えい	22%	13%	0%	6%	9%
ハ．今後脅威となり得ると考えられるものは？（数字は回答割合）							
	ハ-1	顧客の端末（PC、スマホ）への攻撃	39%	22%	0%	12%	26%
	ハ-2	金融機関対外接続系システム（WEBｻｰﾋﾞｽ、ｲﾝﾀｰﾈｯﾄ･ﾊﾞﾝｷﾝｸﾞ等）への攻撃	36%	16%	50%	29%	28%
	ハ-3	金融機関情報系システム（電子メール、EPR等）への攻撃	36%	19%	75%	12%	27%
	ハ-4	クラウド上のシステムへの攻撃	31%	31%	25%	29%	30%
	ハ-5	社員の端末（行員のPCやタブレット等）への攻撃	39%	22%	0%	29%	29%
	ハ-6	金融機関勘定系システムへの攻撃	11%	22%	0%	29%	29%
	ハ-7	金融機関の設備制御系システム（空調、監視ｶﾒﾗ、IoT機器等）への攻撃	8%	25%	25%	0%	13%
	ハ-8	FinTech企業への攻撃	19%	41%	25%	18%	27%
	ハ-9	AIを用いたサービスに対する攻撃	22%	44%	75%	29%	34%

（備考）表中の括弧内の数字は，回答者の各分野におけるサンプル数（回答数）を示す．