

# 防災応用の観点からの機械学習の研究動向

宮本 崇<sup>1</sup>・浅川 匡<sup>2</sup>・久保 久彦<sup>3</sup>・野村 泰稔<sup>4</sup>・宮森 保紀<sup>5</sup>

<sup>1</sup>山梨大学 工学域スマート社会基盤創造研究ユニット (〒 400-8511 山梨県甲府市武田 4-3-11) (Corresponding Author)  
E-mail: tmiyamoto@yamanashi.ac.jp

<sup>2</sup>株式会社 安藤・間 建設本部 土木技術統括部 (〒 107-8658 東京都港区赤坂 6-1-20) (Equally Contributed Author)

<sup>3</sup>防災科学技術研究所 地震津波火山ネットワークセンター (〒 305-0006 つくば市天王台 3-1) (Equally Contributed Author)  
E-mail: hkubo@bosai.go.jp

<sup>4</sup>立命館大学 理工学部 環境都市工学科 (〒 525-8577 滋賀県草津市野路東 1-1-1) (Equally Contributed Author)  
E-mail: y-nomura@fc.ritsumei.ac.jp

<sup>5</sup>北見工業大学 工学部 社会環境系 (〒 090-8507 北海道北見市公園町 165) (Equally Contributed Author)  
E-mail: miyamoya@mail.kitami-it.ac.jp

深層学習に代表される機械学習の手法は近年に大きな性能の発展を遂げており、防災上の各種タスクへの応用も期待されている。一方で、データ数の本質的な不足やタスク処理過程の説明性・解釈性など、防災上の意思決定に用いる上で計算モデルに対処が求められる課題は多い。本稿では、機械学習モデルにおける研究動向からこの2点へ個別に対処するための方法論、およびそれら2点の課題に同時に対処する数理モデルとデータ駆動モデルを統合したアプローチについて、それぞれの考え方や具体的な手法、応用事例を調査する。

**Key Words:** 深層学習, 防災応用, データ不均衡, 解釈性, 数理モデルとの統合

## 1. はじめに

深層学習に代表される機械学習手法は、そのタスク処理性能が飛躍的に発展したことから、様々な分野における応用が近年に進んでいる<sup>1)</sup>。自然災害に対する被害の防御や軽減を担う防災の分野においても、機械学習は災害事象の予測<sup>2)</sup>や被害状況の把握<sup>3)</sup>といった諸場面において利用が試みられており、少子高齢化の進む我が国において人命や国土の安全を保持し続ける上で、今後も有効な活用が期待される重要な技術と考えられる。

深層学習の成功は、大規模な学習を通じて現実のデータが有する性質や分布特性を再現することにその理由があると考えられており<sup>4),5)</sup>。そのような学習のためには多量のデータが存在していることが前提となる。一方で、自然災害は一般に発生頻度の少ない稀少事象であり、平常時に関するデータに比して災害関連データを十分な量で取得することが難しい、データの不均衡性が実用上の問題となっている<sup>6)</sup>。また、深層学習はその原理から計算過程の解釈に難点がある他、計算の安定性や信頼性に対する問題が指摘されている<sup>7)</sup>。機械学習モデルによる事象の予測や判断結果を元に防災上の意思決定を行う上では、計算プロセスに妥当な解釈がなされることは重要である。

データの不均衡性と計算過程の解釈性は、防災上の課題に対して機械学習の応用を図る上で、多くの場面で共通して対処が求められる問題と考えられる。そこ

で本稿では、機械学習分野における研究動向からこの2点へ個別に対処するための方法論、およびそれら2点の課題に同時に対処する数理モデルとデータ駆動モデルを統合したアプローチについて、その考え方や具体的な手法を整理すると共に、それらの具体的な応用事例を調査する。

## 2. データの不均衡性への対応

### (1) サンプルバランシング

データサンプルの不均衡とは、学習データにおけるクラスごとのサンプル数に差があることであり、これにより分類精度や学習の収束や汎化性能に悪影響が生ずることが知られ、多くのレビューや研究が行われている<sup>8),9),10)</sup>。防災や土木構造物の維持管理などでは、通常時と災害時(異常時)などクラスごとで収集できるデータ量が異なるため、サンプルの不均衡は必然的に起こり得る。Sunら<sup>11)</sup>はビッグデータと人工知能を用いた橋梁の構造健全度診断技術におけるレビューを行い、データセットの偏りについて述べている。例えば、**図-1**においてDataset1は通常時と災害時のデータをバランスよく含み量も十分である。Dataset2はデータ量が少なく、バイズ誤差の範囲内で境界を誤って決定する恐れがある。そしてDataset3は、災害時のデータが極端に少なく、最悪の場合は災害時のデータを含まないデータセットで学習を行ってしまう恐れもある。このように防災に関するデータでクラスごとのサンプル

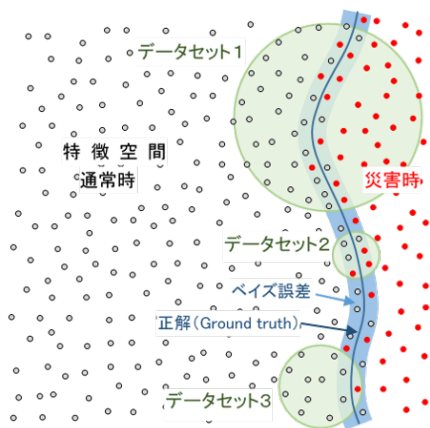


図-1 特徴空間における分類 (Sun ら<sup>11)</sup> を改変)

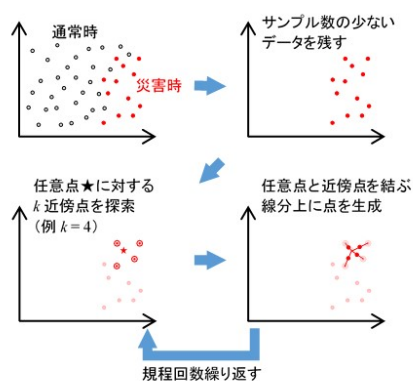


図-3 SMOTE によるオーバーサンプリング

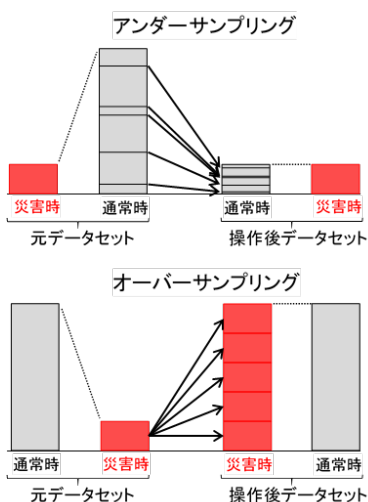


図-2 アンダーサンプリングとオーバーサンプリング

数が不均一だと、異常や被災の見落としなど重大な間違いにつながる恐れがある。

この問題への対策として、第一に学習データ側で調整するデータレベル (サンプリングレベル) の対策があり、図-2のように多い方のクラスのデータを減らすアンダーサンプリングと、少ない方のクラスのデータを増やすオーバーサンプリングがある。このうち、アンダーサンプリングは大量のデータを含むクラスの中から少量のデータのみを選択して使用するものである。アンダーサンプリングでは多数派のクラスの有用なデータを捨ててしまったり、全体のデータ数が不足する懸念があり、サンプル数の減少による分類器の分散が増加したり、データの事前確率の変化による事後分布のゆがみが発生したりする問題が指摘されている<sup>12)</sup>。

一方、少数派のクラスのデータ量を増やすデータ拡張によるオーバーサンプリングは様々な手法が研究されている。データ拡張において簡単な方法は、データをわずかに変化させて新しいデータとして学習データ

に加えるものである。振動・音声などの時系列データの場合では、従来からデータに時間シフトを与えることや、少量のノイズを追加して新しいデータを生成することが行われてきた<sup>13),14),15)</sup>。画像の場合でも、切り取りや回転、色調の変化などでサンプルを拡張することがよく行われている<sup>16),17)</sup>。このようなデータ拡張によってもモデルがまったく同じデータを2回学習するようなことは避けられ、汎化性能が向上することが知られている。

より複雑なオーバーサンプリングの方法としては、以下に挙げるような方法で新しいデータを作り出すものがある。SMOTE (Synthetic Minority Over-sampling Technique)<sup>18)</sup> は、図-3のように少数クラスのデータからランダムにデータを選択し、そのk近傍点とデータを結んだ線の上に新しくデータを生成する方法である。SMOTEはオーバーサンプリング手法としてよく知られており、現在では多数の拡張手法が提案されている<sup>19)</sup>。防災分野では熊本地震において、土砂崩壊地の特徴を評価するためや、地中埋設管の被害予測に適用した例などがある<sup>20),21)</sup>。

## (2) コスト考慮型学習

サンプルの不均衡に対する対策として、これまで述べた学習データ側で調整する対策のほかに分類器側の対策もある。このうち、分類器でデータの誤分類に対してより重いペナルティを課すコスト考慮型学習について述べる<sup>22),23)</sup>。通常の機械学習では、誤分類に関する評価は等しいが、データが不均衡な場合には問題が発生する可能性がある。コスト考慮型学習では、少数派クラスの誤分類に対するペナルティを多数派クラスの場合より重くすることで再現率が向上することを期待するものである。ここで、再現率 (recall) は実際に正であるもののうち、正であると予測されたものの割合である。

HEM (Hard Example Mining)<sup>24)</sup> は間違っ て検出し やすい対象の画像 (ロスの大きい画像) のみを選んで 学習する方法で, 防災に 応用が期待される例として空 中写真から自動車の検出に適用した例が報告されてい る<sup>25)</sup>.

また, 分類器を操作する方法として, 転移学習はある データ領域で学習させたモデルを, 異なる領域のデー タに 適応させる方法である<sup>26)</sup>. 大規模な教師付きデー タセットで学習させたパラメータを, 学習データに限 られる実際の問題に転用することができる. CNN を簡 単に利用する方法として, 学習済みネットワークの適 当な中間層が出力する値をそのまま特徴ベクトルとし て識別器に用いるものがある. 防災における研究例と しては, 合成開口レーダ (SAR) 画像を用いたりモート センシングにおいて災害前後の SAR 画像と, 標高デー タ (DEM) を用いて土砂災害地域の識別, 検出を行っ た例がある<sup>27)</sup>. この例ではもとの広域画像では災害を 受けていない地点が多数のため, データセットでは被 災箇所と非被災箇所の画像が同数になるようにし, 転 移学習として ImageNet と呼ばれる大規模な画像デー タセットで訓練した VGG16 と呼ばれるモデルを識別器と している. また, 鉄道構造物の維持管理のため, レー ル表面傷の判定を転移学習で行った例も報告されてい る<sup>28)</sup>.

さらにサンプルの不均衡によらず, 少ない学習デー タから高精度な予測結果を得るためブースティングと いう方法を用いる場合がある. ブースティングは精度が 中程度のいわゆる弱学習アルゴリズムを組み合わせる ことで, 高精度な強学習アルゴリズムに変換するもの である<sup>29)</sup>. このうち, AdaBOOST (Adaptive Boosting) は弱学習アルゴリズムとして繰り返し同じ分類器を呼 び出し, 訓練データのセットに対する重みのセットの アンサンブルを用いる. 各ラウンドにおいて, 誤って 予測された事例の重みが増やされ, 弱学習アルゴリズム がより難しい事例に集中して学習することで, 最終 的に訓練誤差を低減するものである. 土木分野では構 造健全度診断における実験的研究で, 構造物や部材の 損傷検出に適用された例がある<sup>30),31)</sup>.

### (3) 異常検知

本節では, 近年大きな発展を遂げている深層学習技 術による異常検知技術について概観する.

異常検知問題は, 検知する対象により異常の概念が 変わってくる. Chalapathy ら<sup>32)</sup> は, 深層学習に基づく 異常検知の研究について体系的かつ包括的にレビュー し, 様々な応用領域への適用性や有用性の評価を行って いるものであるが, そこでは, 異常の概念を, 点異常, 文脈依存型異常, 集団型異常の 3 つに分類している.

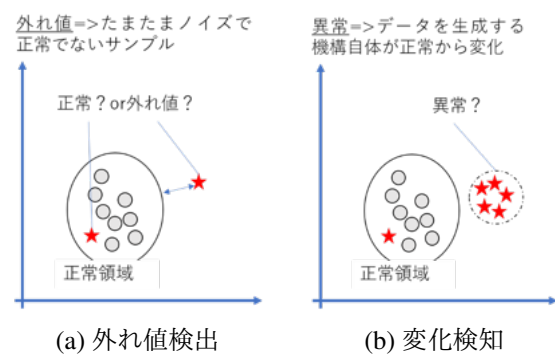


図-4 外れ値検出と変化検知

点異常は, ある確率分布 (多峰性分布も含む) から 大きく外れたデータであり, 代表的な応用例として外 れ値検出問題がある. 文脈型依存型異常は, データの 変化の傾向から予想できる値から外れたデータ等が該 当する. 代表的な応用例として, 時系列上に現れる急 激な変化を, 時系列モデルを仮定して検出する変化点 検知がある. 集団型異常は, 個々のデータポイントに 注目すれば異常ではないが, ある一連のデータを集団 として捉えた場合に, 本来, 同時に発生することがな い状況下で, 集中的に発生する状態などを言う. 代表 的な応用例として, 異常行動検出などがある.

異常検知の問題は, 時系列信号や画像 (輝度値) な どの数値データだけでなく, スпамメールの判定など にも定義できる. 機械学習に基づく異常検知は, デー タの性質に応じて確率分布をどのように学習するかが 重要であり, 通常, 教師あり型と教師なし型の 2 つの タイプに分けられる.

教師あり型は, 正常と異常のサンプルがこれまでに 十分に獲得されており, 教師データとして利用できる 場合においては有効である. 「正常」や「異常」といっ たラベル付きデータを教師データとする識別問題とし て学習ができる. また, 異常のサンプルが十分になく ても, 検知したい異常が入出力関係の不一致として定 義できる場合, センサデータ等を用いた回帰問題とし て, 従来からある時系列モデルや教師あり学習の各種 方法を用いて異常検知ができる. 具体的には, 正常サ ンプルのみから  $y = f(x)$  を推定し, 新しい入力 が得られると, その応答 (出力) を予測できるシステムを 構築する. そして, システムの予測結果と観測データと の逸脱の度合い (予測誤差) から異常を検知するもの である.

一方, 教師なし型の異常検知は, 正常サンプルに比 べて異常サンプルが極端に少ない場合, 正常クラスの凝 集性を学習し, 外れ値検出問題に適用される (図-4 (a)). また, この外れ値の蓄積から, データを生成する機構 自体がある状態 (正常) から変化したと判断するような変

化検知問題に展開できる(図-4(b)). なお, 教師なし型異常検知は, ほとんどのサンプルを占める正常サンプルを用いることから, 半教師あり異常検知と呼ぶ場合もある. 特に教師なし型異常検知技術の中核となるのが, 次元削減器の一種である自己符号化器(Autoencoder: AE)である. AEは入力と出力が一致する( $x = f(x)$ )ように学習するもので, 入力空間を低次の空間に射影するencoderと, 低次の空間から元の空間に戻すdecoderから構成される. AEの中間層を入力次元より少なくしておくことで, 入力空間を低次元空間で表現できるようになる. 大多数の正常サンプルに対してAEを学習することは, 中間層に正常サンプルの表現空間(潜在変数)をつくることに相当する. このようにして作成されたAEに対して, 異常なサンプルが入力されると, decoderが入力ベクトルを復元できず, 入出力のベクトル(あるいは行列)に大きな誤差が生じる. この再構成誤差から異常を検知することができる. 内部のencoderとdecoderのアーキテクチャは入力データが画像の場合, 畳み込みニューラルネットワークが利用されることが多く, 時系列データなどの系列データの場合, 全結合型ニューラルネットワークや長短期記憶アルゴリズム<sup>33)</sup>(Long short-term memory:LSTM)が利用される.

また, 生成モデルの代表的な方法である敵対的生成ネットワーク<sup>34)</sup>(Generative adversarial network: GAN)と変分自己符号化器<sup>35)</sup>(Variational autoencoder: VAE)を基本とした方法が教師なし異常検知問題に有効であるとして利用されている. 生成モデルとは, 観測されたデータが何らかの確率モデルから生成されていると仮定し, その生成過程を確率分布によってモデル化し, そのモデルから全く新しいデータ(画像・信号・テキストなど)を作ることを目的としたものである.

GANは二種類のニューラルネットワークから構成され, 一方は生成器と呼ばれ, ランダムノイズを入力としてデータを生成する役割を果たし, もう一方は識別器と呼ばれ, 生成されたデータが, 事前に与えた教師データであるのか, あるいは生成器から出力されたデータであるのか識別する役割を果たす. 生成器は, 識別器が教師データであると誤認識するようなデータを生成するように学習を進め, 識別器は, 生成器から出力されるデータを誤認識することなく正しく生成器からの出力であると認識するように学習を進めていく. このようなプロセスを応用し, 画像データの異常検知問題に適用したものがAnoGAN<sup>36)</sup>である. AnoGANでは, まず, 教師データとして正常画像を与え, 上記のGANモデルを学習させる. 次に, 異常かどうかテストしたい画像(テスト画像)に対して, よく似た画像を生成できるノイズ(生成ノイズ)を求め, それを生成器に入力し画

像を生成させる. このとき, 生成された画像が入力されたテスト画像を的確に再構成していれば正常であると判断し, そうでなければ異常と判断するというのがAnoGANの処理の流れである. AnoGANではテスト画像とよく似たデータを生成する生成ノイズを求める際に, テスト画像と生成画像の誤差をバックプロパゲーションして生成ノイズを更新学習するため, 異常検知に時間が掛かるという問題があり, 現在, この方法を解決するための方法の一つとしてEfficient GANという方法が提案されている. この詳細はZenatiら<sup>37)</sup>を参照されたい. 一方, もう一つの代表的な生成モデルであるVAEはAEと同様に入力データと出力データが一致するように学習するものであるが, 潜在変数の取り扱いが異なる. VAEのencoderは入力 $X$ が得られると, 潜在変数 $z$ の変分事後分布 $q_\phi(z|X)$ における平均ベクトルと標準偏差ベクトルの二つの要素を出力し, decoderは, 潜在変数 $z$ が得られた上での $X$ の条件付き確率 $p_\theta(X|z)$ における平均ベクトルと標準偏差ベクトルの二つの要素を出力する. なお, 変分事後分布 $q_\phi(z|X)$ と条件付き確率 $p_\theta(X|z)$ は多変量正規分布としてモデル化される. VAEでは, 潜在変数を確率分布としてモデル化することで, decoderの出力を連続的に変化することが可能となる. なお, 損失関数には, 再構成誤差と正則化損失が用いられる. この再構成誤差を異常度として定義すると, 複雑な構造を有する工業製品の画像データによる教師なし異常検知が高精度に行えることが報告されている<sup>38)</sup>.

また, 教師なし型異常検知には, ハイブリッド型のものもある. 例えば, 事前に訓練された転移学習モデルやAEを特徴抽出器として利用し, 集約された情報をone-class SVMなどの従来の機械学習アルゴリズムに入力し, 外れ値検出を行う方法も広く利用されている. ただし, 転移学習モデルやAEで獲得される特徴ベクトルは異常検知を的確に行えるように表現されているわけではなく, つまり, 異常検知用に特化されていないことから, 外れ値の検出に失敗してしまうことがある. この問題を克服するため, 近年, one-class SVMでの考え方を踏襲したone-classニューラルネットワーク<sup>39)</sup>が提案されている. 原著論文では, one-class SVMにおけるカーネルトリックを使用して高次元空間に写像する部分を全結合型ニューラルネットワークの一層目に置き換え, 他の層の重みも含めバックプロパゲーションにより最適化するとともに, これらと交互に, 原点と教師データ(正常データ)のマージンを最大化するように境界を決定することで, 異常検知に適した特徴表現を獲得できるとしている. 産業分野への応用はされていないものの, いくつかのベンチマーク問題において, 有効性が示されていることから, 注目すべき

方法論として考えられる。

本節では、異常検知問題における従来の機械学習技術については触れなかったが、Chalapathyら<sup>32)</sup>らによって述べられているように、深層学習は従来の機械学習技術と比較して、データの規模が大きくなるにつれて性能が向上する傾向がある。従来の機械学習技術を用いた異常検知に関しては、井出・杉山<sup>40)</sup>によって詳細にまとめられているので、そちらを参照されたい。

### 3. 計算過程の説明性・解釈性の向上

#### (1) 説明可能 AI

近年著しい発展を遂げた機械学習モデルは人間よりも高い精度で予測や識別を行うことができる一方で、その根拠を説明することは苦手としている。このような「ブラックボックス性」への不安が機械学習の社会実装を阻む壁になっているケースがよく見受けられる。特に防災や医療など人の一生を左右するような場面で機械学習を使おうとした際に、機械学習モデルが導き出す答えに説明を求められるケースが出てくるのが十分に想定される。

このような状況を背景として、「機械学習モデルの説明・解釈」が近年重要視され始めている。アメリカ国防総省の研究機関である国防高等研究計画局 (DARPA) は、Explainable AI (XAI, 説明可能 AI) プロジェクトを2017年から立ち上げている<sup>41)</sup>。このプロジェクトは単に性能がよい機械学習モデルを作るのではなく、人間から見て理解でき、そして信頼できる説明可能な機械学習モデルを作っていくことを目指している。日本においても、総務省のAIネットワーク社会推進会議が2019年に公表したAI利活用ガイドラインにおいて、「透明性の原則」や「アカウントビリティの原則」などがうたわれている<sup>42)</sup>。また内閣府統合イノベーション戦略推進会議が発表したAI戦略2019の中核基盤研究開発の一つとして「説明できるAI技術」が取り上げられている<sup>43)</sup>。

このような社会的要請を受けて、説明可能AIに関する研究が盛り上がりを見せている。本節では、説明可能AIに関する簡単なレビューを行っていくが、情報分野における現状は原<sup>44),45)</sup>、やMolnar<sup>46)</sup>にまとめられているので、興味のある読者はぜひ一読願いたい。なお本稿での説明可能AIは、機械学習モデルの出力に関する補助情報（出力根拠の説明、モデルの解釈など）を提供することなどにより、機械学習モデルが導き出す答えやそれに至るプロセスが人間にとって説明可能・解釈可能な形になっており、脱ブラックボックス化（ホワイトボックス化）した機械学習モデルを指す。

説明可能AIの一番のメリットは、ユーザーの不安を

解消させ、新たな機械学習技術を社会に受け入れやすくする点にある。また機械学習モデルによる出力がこれまでの科学的知見と整合する形で行われているのかを説明に基づいて調べることによって、機械学習モデルの妥当性を検証することができる。データが少ない、もしくはデータがない部分での予測や識別を機械学習は苦手としていることが多いが、そのような状況での機械学習モデルの信頼性を担保するには既往の知見を含めた検証が欠かせない。さらに説明可能AIから得られる説明は、なぜ機械学習が高い精度での予測や識別を実現しているのかを知る手掛かりともなるため、新たな科学的発見につながることも期待される。

次に機械学習モデルの「説明」について考えていく。大きくは機械学習モデルの挙動に関する説明と、機械学習モデル自体の仕組みの説明に分けられる。挙動の説明は、入力と出力の関係のみを説明するものであり、機械学習モデルの中がどのようなようになっているかは気にしないという意味でブラックボックス的な説明である。他方で、仕組みの説明はモデルの中でどのような作用が働いて出力に至っているのかを説明するものであり、ホワイトボックス的な説明である。

挙動の説明として代表的なものは、データのどの特徴が重要だったのかを調べることである。データの特徴量（入力変数）が少ない場合は、個々の特徴量の重要度を調べるのがよく行われている。決定木などの木系モデルや単純なニューラルネットワークの場合、モデル内部において特徴量が出現する頻度などから、個々の特徴量の重要度を求めることができる。また学習済みモデルによる出力結果から各特徴量の重要度を評価するPermutation Importance<sup>47),48)</sup>は、モデルに依存せずに重要度を求めることができるため、複雑な機械学習モデルに対しても適用可能である。

それぞれの特徴量が機械学習モデルの出力結果にどのような影響を与えているのかという説明も行われている。Partial Dependence<sup>49)</sup>は個別の予測結果に基づき、1つの特徴量が増減すると予測がどう変わるのかを調べることができる。またLIME (Local Interpretable Model-agnostic Explanations)<sup>50)</sup>やSHAP (SHapley Additive exPlanations)<sup>51)</sup>は深層ニューラルネットやアンサンブル木などの複雑な機械学習モデルの予測をより単純で解釈しやすいモデルで近似することで、個別の予測結果に対して特徴量がどう寄与するのかを評価することができる。

画像のように入力変数が多いデータを用いている場合は、入力した画像において機械学習モデルが注目する領域をハイライトさせることによって、データのどの特徴が重要だったのかに関する説明が行われている。代表的な手法としてはGrad-CAM (Gradient-weighted Class



Activation Mapping)<sup>52)</sup>が挙げられる。他にも、予測プロセスを自然言語で記述することによる説明（例えば、画像のどの部分が予測・識別において重要だったかを説明文として提示）や、学習したデータの中で特に重要であったものを提示する説明などが提案されている。

仕組みの説明としては、深層ニューラルネットワークなどの複雑な機械学習モデルを、可読性の高いモデル、例えば単一の決定木やルールモデルで近似的に表現することでモデルへと置き換えることが試みられている。また設計段階から可読性の高い解釈可能なモデルを採用し、それを用いて学習モデルを作成することによって説明可能 AI を作り出す研究もおこなわれている。

実際の問題に対する説明可能 AI の適用も進められており、防災分野でもいくつか実例がある。久保ら<sup>53)</sup>は地震動指標を予測するランダムフォレストモデルの解釈・説明を試みており、各入力変数の重要度の評価、SHAP による説明変数と予測変数の関係の評価、アンサンブル学習における弱学習器の可視化を通じた機械学習モデルの解明などを行っている。青島ら<sup>54)</sup>は深層ニューラルネットワークによる画像認識を用いたコンクリート構造物の変状検出において、Grad-CAM による学習モデルの着目領域の可視化を行うことで分類結果の検証を行っている。

ここまで見てきたように説明可能 AI に向けた動きが加速しているが、説明可能 AI が提供するものはあくまでも補助的な情報であり、それだけで機械学習モデルの全てを説明できるわけではないことに留意してほしい。また対象とする問題ごとに必要となる説明の程度は異なっており、それに合わせた形での機械学習モデルの説明・解釈を行っていくことが重要である。説明可能 AI で得られる説明が誤っているリスクがあることや、説明可能 AI の取り組みには相応のコストやリソースが必要であることを考えると、どこまでの説明を求めるべきなのか、そもそも説明が必要なのかを念頭に置くべきである。また、説明可能 AI から得られるものはあくまで「学習済みモデルの説明」であり、「解きたい問題の説明」ではないことにも注意する必要がある。

## (2) ベイズ学習

機械学習モデルの解釈性を向上させる別のアプローチとして、ベイズ学習の方法がある。ベイズ学習は、観測データと未観測の変数の関係性を各種統計分布などの確率分布を組み合わせることで事前分布を作成し、ベイズ推定により事後確率分布を算出し、未知の変数の学習を行う計算モデルである。

ベイズ学習の枠組みでは、物理的な計算プロセスの中で、同定したいパラメータをベイズ推定によってデータから確率的に推定する。計算全体は物理的な知見に

基づいているためにブラックボックス的な部分がなく、計算の解釈性や説明性に優れたモデルである。また、ベイズ学習によるパラメータの推定では、事前分布の設定が正則化の役割を果たすために、原理的に過学習が生じにくいことが知られており、観測データ数が少ない場合や、モデルに含まれるデータの次元数や種類の数が多い場合に適しているとされる<sup>55)</sup>。

ベイズ学習（ベイズ推定）の基本的枠組みは、モデルの未知パラメータ  $\theta$  について、確率分布による統計的モデリングを適用した事前分布  $p(\theta)$  を設定し、ベイズの定理  $p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)}$  を適用することによって条件付き分布  $p(\theta|D)$  の事後確率分布を求めることと言える。ここで、ベイズ学習において使用されるベイズの定理は、条件付き分布の以下の定義

$$p(x|y) = \frac{p(x,y)}{p(y)} \quad (1)$$

$$p(y|x) = \frac{p(x,y)}{p(x)} \quad (2)$$

と、確率分布の周辺化  $p(y) = \int p(x,y)dx$  から導かれる、下記の等式である。

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\int p(x,y)dx} \quad (3)$$

$y$  に対して特定の値が決められた時の  $x$  の条件付き分布は、 $x$  に対して特定の値が決められた時の  $y$  の条件付き分布と確率分布  $p(x)$  の積を確率分布  $p(y)$  で除した値と等しい。これは、原因  $x$  から結果  $y$  が得られる確率  $p(y|x)$  から、結果  $y$  が得られたときの原因  $x$  の確率  $p(x|y)$  を逆算する計算となっている。

ベイズ学習によるモデルの構築は、多項式回帰などの解析的に計算が実施できる場合、以下のような確率分布が事前分布として選択される。2値を取る離散変数に対しては離散確率分布として、ベルヌーイ分布が用いられる。例えば、 $N$  個のデータ点  $\mathbf{X} = \{x_1, \dots, x_N\}$  を観測した後の事後分布は以下ようになる。

$$p(\mu|\mathbf{X}) = \frac{\{\prod_{n=1}^N p(x_n|\mu)\}p(\mu)}{p(\mathbf{X})} \quad (4)$$

より一般の  $K$  値を取る事象に対しては、ディリクレ分布が用いられる。連続確率分布として一般的に用いられるのは、次元や次元のガウス分布である。

この際、事後分布や予測分布の効率的に計算する方法として共役事前分布を使用することが行われている。共役事前分布は、事前分布と事後分布が同じ種類の確率分布を有するように設定された事前分布のことで、例えば、尤度関数がベルヌーイ分布で記述される場合、事前分布として共役事前分布であるベータ分布を使用すると、事後分布もベルヌーイ分布になる。

機械学習で対象とされる一般的な画像や自然言語などの応用分野では、より複雑な確率モデルが必要とされ、解析的に事後確率分布や予測分布を計算すること

は一般的に困難となる。解析的に計算できない未知の確率分布を計算するための手段は2種類あり<sup>56)</sup>、1つがサンプリングである。ベイズ推論の分野では、MCMC (Markov chain Monte Carlo) と呼ばれるサンプリング手法群が広く利用されている。また、ギブスサンプリングやハミルトニアンモンテカルロ、逐次モンテカルロなど様々な手法が提案されている<sup>57),58)</sup>。またもう一つのアプローチは、計算のしやすい簡単な式を使って周辺化計算を近似的に実行する方法である。その方法には、ラプラス近似、変分推論 (変分ベイズ)、期待値伝播といった手法がある。変分推論では、最適化問題を解くことによって未知の確率分布の近似的な表現を得るが、これは例えばKLダイバージェンスの基準を用いた以下の最小化問題の解として得られる。

$$q_{\text{opt}}(z) = \arg \min_q KL[q(z)|p(z)] \quad (5)$$

ベイズ学習は、対象データが含む不確実性を量的に取り扱いながら、物理的・数理的な知見に観測データをなじませる自然な手法の1つであり、土木分野や防災分野においても様々な応用がなされている (例えば<sup>59),60)</sup>。

#### 4. 数理モデルとデータ駆動モデルの統合

データの不足や不均衡性、計算モデルの解釈性といった問題を解決する上で、前章までに述べた手法に加えて数理モデルとデータ駆動モデルを統合的に利用した方法論が徐々に応用される動向にある。それらの方法論は、数理モデルの有する物理的知見との親和性や計算過程の透明性といった利点と、データ駆動モデルの有する表現能力や柔軟性といった利点を相補的に利用するものであり、両者を部分的に混成した解析過程に加えて近年は新しい計算モデルも提案されている。

そこで本章では、前述した問題への対処という観点から、こうした方法論のいくつかについてその具体例を説明する。

##### (1) シミュレーション学習

大規模なデータの取得が難しい事例などにおいて、機械学習のための教師データを数値シミュレーションによって生成し利用する研究が行われている。数値シミュレーションは、多量のデータの生成やそのアノテーションが容易であり、特に深層学習に必要とされる大規模な教師データの生成の観点において有効であると考えられる。本稿ではそのようなアプローチをシミュレーション学習と称する。

観測の機会の少ない稀少事象のデータを補完するために物理シミュレーションを用いた事例として、Matsuokaら<sup>61)</sup>は数値計算によって仮想的に生成した台風の発生・

発達的气象条件を学習させた深層学習モデルを利用して、台風の発生する前段階を高精度に検知する手法を提案している。また Araya-Polo ら<sup>62)</sup>は、ある地盤構造に対して地震波動伝播を計算した大量の数値シミュレーション結果を元に、入出力関係を逆転させたデータを学習した深層学習モデルを逆問題の解法として利用し、地震波トモグラフィーに応用している。また、数値シミュレーションの入出力関係を直接再現する機械学習モデル (サロゲートモデル) を用いて、計算の高速化やパラメトリックスタディを行うアプローチは、流体解析<sup>63)</sup> や構造計算<sup>64)</sup> などの分野で活用されている。

##### (2) データ駆動モデルへの物理的知識の反映

近年における機械学習手法の動向の1つに、物理的な知識を機械学習モデルの計算プロセスに何らかの形で統合・反映させることにより、モデルの学習効率や予測性能の向上を図ると共に、物理的・数理的事実や事前知識に反しない結果を与えるデータ駆動モデルを構築しようとするものがある。

一言ら<sup>65)</sup>は、物理的なモデルによる河川流量の予測をニューラルネットワークを組み合わせた河川水位予測を行うことによって、特に機械学習モデルの学習時に未経験であった規模の事象の予測精度が改善する可能性があることを示している。また、Greydanus ら<sup>66)</sup> や Sanchez-Gonzalez ら<sup>67)</sup> は、ニューラルネットワークにハミルトニアンを学習・予測させることによって、エネルギー保存則などの物理法則に反しない良好な予測結果が得られることや、学習時に未経験だった事象に対する外挿能力が得られることを示している。

これらの手法は、データの不足や計算モデルの解釈性への対処だけでなく、データの背後にある法則を科学的知識として発見するための方法論としても研究が進んでいる<sup>68),69)</sup>。防災上、必ずしも十分な量のデータが得ることが難しい複雑な自然現象の予測が必要となる場合においては、現象のより本質的な理解に基づいて計算モデルを構築しようとするこうしたアプローチも今後重要になるものと考えられる。

#### 5. おわりに

本稿は、防災上の課題に対して機械学習手法を適用する際に共通の課題となり得る、データの不均衡性と計算過程の解釈性という2点の問題への対処という観点から、近年の機械学習分野における研究動向をその基本的な考え方と具体的な手法、研究事例について調査、整理した (表-1)。本稿では主に手法的な観点に比重を置いた調査を行ったが、応用事例についてはここで挙げた以外にも災害時の人流解析など多岐に渡る

課題	データの不足・不均衡性	モデルの説明性
対処法	サンプルバランシング	説明可能 AI
	コスト考慮型学習	ベイズ学習
	異常検知	
	数理モデルとデータ駆動モデルの統合	

表-1 防災上の問題に対する機械学習活用時の課題と対処法のまとめ

先駆的な研究も見られるようになっており、それらについて今後整理することも重要である。

本稿で整理した考え方は防災分野以外にも既に多くの事例で用いられている基本的なものも多いが、今後は防災上の課題により特化した各手法の改良などを通じて、実務への適用や研究分野の拡大が進むことが期待される。

#### 参考文献

- 1) 情報処理推進機構 AI 白書編集委員会：AI 白書 2020, KADOKAWA, 2020.
- 2) 一言正之, 櫻庭雅明, 清雄一: 深層学習を用いた河川水位予測手法の開発, 土木学会論文集 B1 (水工学), Vol.72, No.4, pp.1.187-1.192, 2016.
- 3) Miyamoto, T. and Yamamoto, Y.: Using multimodal learning model for earthquake damage detection based on optical satellite imagery and structural attributes, Proceedings of 2020 IEEE International Geoscience and Remote Sensing Symposium, accepted, 2020.
- 4) Rifai, S., Dauphin, Y. N., Vincent, P., Bengio, Y. and Muller, X.: The Manifold Tangent Classifier, *Advances in Neural Information Processing Systems 24*, pp.2294-2302, 2011.
- 5) Bengio, Y., Courville, A. and Vincent P.: Representation Learning: A Review and New Perspectives, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No.8, pp.1798-1828, 2013.
- 6) 宮本崇: 人工衛星撮影データの時系列性と不均衡性を考慮した機械学習モデルによる地震被害判別, 建造物の安全性・信頼性 Vol. 9 JCOSAR2019 論文集 B, pp.298-299, 2019.
- 7) Buhmester, V., Münch, D. and Arens, M.: Analysis of explainers of black box deep neural networks for computer vision: a survey, arXiv preprint, arXiv:1911.12116, 2019.
- 8) Thabtah, F., Hammoud, S., Kamalov, F. and Gonsalves, A.: Data imbalance in classification: Experimental evaluation, *Information Sciences*, Vol. 513, pp. 429-441, 2020.
- 9) He, H. and Garcia, E. A.: Learning from Imbalanced Data, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 21, No.9, pp.1263-1284, 2009.
- 10) Barandela, R., Rangel, E., Sánchez, J. and Ferri, F.: Restricted Decontamination for the Imbalanced Training Sample Problem, *Progress in Pattern Recognition Research*, Vol. 2905, pp.424-431. 2003.
- 11) Sun, L., Shang, Z., Xia, Y., Bhowmick, S. and Nagarajaiah, S.: Review of Bridge Structural Health Monitoring Aided by Big Data and Artificial Intelligence: From Condition Assessment to Damage Detection, *Journal of Structural Engineering*, Vol.146, Issue 5, online, 2020.
- 12) Pozzolo, A. D., Caelen, O. and Bontempi, G.: When is Undersampling Effective in Unbalanced Classification Tasks?,

- Machine Learning and Knowledge Discovery in Databases*, pp.200-215, 2015.
- 13) Sietsma, J. and Dow, R. J. F.: Creating artificial neural networks that generalize, *Neural Networks*, Vol.4, Issue 1, pp.67-79, 1991.
  - 14) Solovyev, R. A., Vakhrushev, M., Radionov, A., Romanova, I. I., Amerikanov, A. A., Aliev, V. and Shvets, A. A.: Deep learning approaches for understanding simple speech commands. In *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, pp.688-693, 2020.
  - 15) 宮本崇, 古屋貴彦, 盛川仁: 深層学習を用いた常時微動記録からの解析対象区間の自動抽出, 土木学会論文集 A2 (応用力学), Vol. 73, No. 2, p.1.321-1.331, 2017.
  - 16) Francois Chollet(著), クイープ(訳), 巢籠悠輔(監訳): Pyton と Kares によるディープラーニング, マイナビ出版, 2018.
  - 17) Goodfellow, I., Bengio, Y. and Courville, A. (著), 岩澤有祐, 鈴木雅大, 中山浩太郎, 松尾豊(監訳): 深層学習, ドワンゴ, 2018.
  - 18) Chawla, N. V., Bowyer, K. W., Hall, L. O. and Kegelmeyer, W. P.: SMOTE: Synthetic Minority Over-sampling Technique, *Journal of Artificial Intelligence Research*, Vol.16, pp.321-357, 2002.
  - 19) Fernandez, A., Garcia, S., Herrera, F. and Chawla, N. V.: SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary, *Journal of Artificial Intelligence Research*, Vol.61, pp.863-905, 2018.
  - 20) 古川昭太, 丸山喜久: 共分散構造分析と機械学習に基づく地すべりと土砂崩壊地点の地形的特徴の評価, 土木学会論文集 A1 (構造・地震工学), Vol. 74, No. 4, pp. 1.369-1.380, 2018.
  - 21) 能島暢陽, 大西克茂: 地中埋設管の地震被害予測への機械学習の適用に関する基礎的検討, 東濃地震科学研究所防災研究委員会 2018 年度報告, pp.23-31, 2018.
  - 22) Zhou, Z.-H. and Liu, X.-Y.: Training cost-sensitive neural networks with methods addressing the class imbalance problem, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 18, No.1, pp.63-77, 2006.
  - 23) Chung, Y.-A., Lin, H.-T. and Yang, S.-W.: Cost-aware Pre-training for Multiclass Cost-sensitive Deep Learning, *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)*, pp.1411-1417, 2016.
  - 24) Shrivastava, A., Gupta, A. and Girshick, R.: Training Region-based Object Detectors with Online Hard Example Mining, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.761-769, 2016.
  - 25) Koga, Y., Miyazaki, H. and Shibasaki, H.: A CNN-Based Method of Vehicle Detection from Aerial Images Using Hard Example Mining, *Remote Sensing*, Vol. 10, No. 1, online, 2018.
  - 26) 中山 英樹: 深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習, 電子情報通信学会技術研究報告, Vol. 115, No. 146, SP2015-45, pp.55-59, 2015.
  - 27) 植田大介, 間普真吾, 呉本亮: 災害前後 SAR 画像と DEM データを用いた CNN による土砂災害検出, 人工知能学会全国大会論文集, JSAI2018 巻, セッション ID 1D1-02, p. 1D102, 2018.
  - 28) 小松佳弘: 「線路設備モニタリング装置」による軌道設備不良の検知—転移学習によるレール表面傷の判定—, 土木学会誌, Vol. 103, No. 2, pp.26-27, 2018.
  - 29) フロイド ヨアブ, シャピリ ロバート, 安倍直樹(訳): ブースティング入門 (<特集> 計算学習理論の進展と応用可能性), 人工知能学会誌, Vol. 14, No. 5, pp.771-780, 1999.
  - 30) 服部洋, 古田均, 野村泰稔, 中津功一朗, 石橋健: AdaBoost



- による振動予測を用いた橋梁ヘルスマニタリングシステムの構築に関する実験的研究, 土木学会論文集 A2 (応用力学), Vol. 67, No. 2, I.825-I.832, 2011.
- 31) Ying, Y., Garrett Jr, J. H., Oppenheim, I. J., Soibelman, L., Harley, J. B., Shi, J. and Jin, Y.: Toward data-driven structural health monitoring: Application of machine learning and signal processing to damage detection, *Journal of Computer Civil Engineering*, Vol.27, No. 6, pp.667-680, 2013.
  - 32) Chalapathy, R. and Chawla, S.: Deep learning for anomaly detection: A survey, arXiv preprint arXiv:1901.03407v2, 2019.
  - 33) Hochreiter, S. and Schmidhuber, J.: Long short-term memory, *Neural Computation*, Vol. 9, No. 8, pp.1735-1780, 1997.
  - 34) Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: Generative adversarial Nets, *Advances in Neural information processing systems 27 (NIPS 2014)*, pp.2672-2680, 2014.
  - 35) Kingma, D.P. and Welling, M.: Auto-encoding variational bayes, arXiv preprint arXiv:1312.6114v10, 2014.
  - 36) Schlegl, T., Seebock, P., Waldstein, S. M., Schmidt-Erfurth, U. and Langs, G.: Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, *Proc. of International Conference on Information Processing in Medical Imaging*, pp.146-157, 2017.
  - 37) Zenati, H., Foo, C. S., Lecouat, B., Manek, G. and Chandrasekhar, V. R.: Efficient GAN-based anomaly detection, arXiv preprint arXiv:1802.06222v2, 2019.
  - 38) 立花亮介, 松原崇, 上原邦昭: 深層生成モデルによる非正規化異常度を用いた工業製品の異常検知, 2018 年度人工知能学会全国大会 (第 32 回) 講演概要集, 2A1-03, 2018.
  - 39) Chalapathy, R., Menon, A. K. and Chawla, S.: Anomaly detection using one-class neural networks, arXiv preprint arXiv:1802.06360v2, 2019.
  - 40) 井手剛, 杉山将: 異常検知と変化検知, 講談社, 2017.
  - 41) Defense Advanced Research Projects Agency (DARPA): Explainable Artificial Intelligence (XAI), <https://www.darpa.mil/program/explainable-artificial-intelligence>
  - 42) 総務省 AI ネットワーク社会推進会議: AI 利活用ガイドライン～AI 利活用のためのプラクティカルリファレンス～  
[https://www.soumu.go.jp/main\\_content/000637097.pdf](https://www.soumu.go.jp/main_content/000637097.pdf)
  - 43) 内閣府 統合イノベーション戦略推進会議: AI 戦略 2019 ～人・産業・地域・政府全てに AI～  
[https://www.kantei.go.jp/jp/singi/ai\\_senryaku/pdf/aistrategy2019.pdf](https://www.kantei.go.jp/jp/singi/ai_senryaku/pdf/aistrategy2019.pdf)
  - 44) 原聡: 機械学習における解釈性, 人工知能, Vol. 33, No. 3, pp.366-369, 2018.
  - 45) 原聡: 説明可能 AI, 人工知能, Vol. 34, No. 4, pp.577-582, 2019.
  - 46) Molnar, C.: Interpretable Machine Learning <https://christophm.github.io/interpretable-ml-book/index.html>
  - 47) Breiman, L.: Random Forests, *Machine Learning*, Vol. 45, No. 1, pp.5-32, 2001.
  - 48) Fisher, A., Cynthia R. and Francesca D: Model Class Reliance: Variable Importance Measures for Any Machine Learning Model Class, from the "Rashomon" Perspective, arXiv preprint arXiv:180101489, 2018.
  - 49) Friedman, J. H.: Greedy Function Approximation: A Gradient Boosting Machine, *Annals of Statistics*, Vol. 29, No. 5, pp.1189-1232, 2001.
  - 50) Ribeiro, M. T., Sameer S. and Carlos G.: "Why Should I Trust You?": Explaining the Predictions of Any Classifier, *Proc. of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016.
  - 51) Lundberg, S. M. and Lee, S.-I.: A Unified Approach to Interpreting Model Predictions, *Proc. of Neural Information Processing Systems 2017*, 2017.
  - 52) Ramprasaath, R. S., Michael, C., Abhishek, D., Ramakrishna, V., Devi, P. and Dhruv, B.: Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, *Proc. of IEEE International Conference on Computer Vision 2017*, pp.618-626, 2017.
  - 53) 久保久彦, 刀卓, 鈴木亘, 木村武志, 青井真: 説明可能な地震動指標のランダムフォレスト予測器の構築に向けた取り組み, 2020 年度人工知能学会全国大会, 4Rin1-94, 2020.
  - 54) 青島亘佐, 河村伸哉, 中野聡, 中村秀明: 深層学習による画像認識を用いたコンクリート構造物の変状検出に関する研究, 土木学会論文集 E2 (材料・コンクリート構造), Vol. 74, No. 4, pp.293-305, 2018.
  - 55) 須山敦志, 杉山将: ベイズ推論による機械学習入門, 講談社, 2017.
  - 56) 須山敦志: ベイズ深層学習, 講談社, 2019.
  - 57) 松平京介, 永田賢二, 本武陽一, 岡田真人: レプリカ交換モンテカルロ法を用いた Mixture of Experts モデルにおけるベイズ推論, 情報処理学会論文誌数理モデル化と応用, Vol. 12, No.3, pp.37-45, 2019.
  - 58) 下坂正倫, 早川裕太, 坪内孝太: 携帯端末位置履歴を用いた階層ディリクレ混合回帰モデルに基づく活動人口予測, 研究報告ユビキタスコンピューティングシステム, Vol.15, pp.1-7, 2018.
  - 59) 五井良直, 金哲佑: ベイズ推論を用いた橋梁の振動特性推定に伴う不確実性の定量化手法の提案, 土木学会論文集 A2 (応用力学), Vol. 75, No. 2, pp.1.647-657, 2019.
  - 60) 能島暢呂, 杉戸真太, 金澤伸治: 被害情報の逐次処理による地震時緊急対応の意思決定支援モデル, 土木学会論文集, No.682, pp.129-142, 2001.
  - 61) Matsuoka, D., Nakano, M., Sugiyama, D. and Uchida, S.: Detecting the Precursor of Tropical Cyclone using Deep Neural Networks, *Proceedings of the 7th International Workshop on Climate Informatics*, pp.73-76, 2017.
  - 62) Araya-Polo, M., Jennings, J., Adler, A. and Dahike, T.: Deep-learning Tomography, *The Leading Edge*, Vol. 37, No. 1, pp.58-66, 2018.
  - 63) Kutz, J. N.: Deep Learning in Fluid Dynamics, *Journal of Fluid Mechanics*, No. 814, pp.1-4, 2017.
  - 64) 西尾真由子, 三浦正樹, 珠玖隆行: スパースモデリングによる既存橋梁の構造信頼性計算のための代替モデル構築, 土木学会論文集 A2 (応用力学) Vol. 74, No.2, pp.I.125-I.136, 2018.
  - 65) 一言正之, 桜庭雅明: 深層ニューラルネットワークと分布型モデルを組み合わせたハイブリッド河川水位予測手法, 土木学会論文集 B1 (水工学), Vol. 73, No. 1, pp.22-33, 2017.
  - 66) Greydanus, S., Dzamba, M. and Yosinski, J.: Hamiltonian Neural Networks, *Advances in Neural Information Processing Systems 32*, pp.15379-15389, 2019.
  - 67) Sanchez-Gonzalez, A., Bapst, V. and Cranmer, K.: Hamiltonian graph networks with ODE integrators, arXiv preprint arXiv:1909.12790, 2019.
  - 68) Takeishi, N., Kawahara, Y. and Yairi, T.: Learning koopman invariant subspaces for dynamic mode decomposition, *Advances in Neural Information Processing Systems 30*, pp.1130-1140, 2017.
  - 69) Berg, J. and Nyström, K.: Data-driven discovery of PDEs in

## RESEARCH TRENDS IN MACHINE LEARNING FROM THE PERSPECTIVE OF DISASTER PREVENTION APPLICATIONS

Takashi MIYAMOTO, Tadasu ASAKAWA, Hisahiko KUBO, Yasutoshi NOMURA and  
Yasunori MIYAMORI

In recent years, machine learning methods such as deep learning have evolved greatly in terms of performance, and they have been used for various purposes in disaster prevention. On the other hand, the intrinsic shortage of the number of data, the improvement of explanatory and interpretive nature of the task processing process are important issues that need to be addressed by computational models for decision making in disaster management. In this paper, we will discuss the concept, methods, and applications for addressing these two points based on research trends in the field of machine learning, and also introduce approaches that integrate mathematical and data-driven models to address these two issues.