

特集：認知行動療法研究の新時代を切り開く研究法  
〈展望〉

## 機械学習アプローチの臨床応用 — デジタル革新がもたらすメンタルヘルスケアの新たな形 —

山本 哲也<sup>1</sup>      吉本 潤一郎<sup>2</sup>

### 要 約

機械学習は、近年注目を集めている人工知能技術の一分野であり、問題に対して最適な解決策に到達するための方法やパラメータを自動的に決定する計算戦略である。このアプローチでは、多次元データセットに内在する規則性を発見することによって、個人の状態に焦点をあてた予測モデルを構築することができる。そのため、認知行動療法をはじめとした臨床実践において、アセスメントの効率化・精緻化や、最適な介入方法の選定に寄与する可能性がある。そこで本論文では、まず機械学習アプローチの枠組みや、統計学との違い、そしてその長短を概観する。加えて、これまでのメンタルヘルス領域において、機械学習アプローチが適用されている主な研究テーマを整理した上で、臨床心理学および認知行動療法研究に寄与しうる活用例を紹介する。最後に、機械学習アプローチの限界に触れながら、今後の応用可能性について論じる。

キーワード：機械学習、人工知能、認知行動療法、メンタルヘルス、臨床心理学

### はじめに

効果的な認知行動療法の実践には、問題となる行動（認知や生理的反応も含む）の同定と、それに影響を及ぼす変数間における関数関係の分析（以下、機能分析）を行うことが必要不可欠である。機能分析に基づいて変数間の制御可能な要因を見出すことによって、問題解決に有効な介入計画を策定することが可能になる（O'Brien, Haynes, & Kaholokula, 2016）。

機能分析に代表されるこのような個別化されたアセスメントは、一般的に熟練したセラピストによる綿密な情報収集と仮説化、そして検証が繰り返されることで行われる。一方で、クライアントの意識に顕在化しない変数（たとえば、本人が気づいていない行動習慣や生理的反応など）や、セラピストによる観察が困難な変数（たとえば、クライアントの内的体験など）に対しては、関数関係の把握に時間がかかることがしばしば起こりうる。また、アセスメントの結果、有効性

---

<sup>1</sup> 徳島大学大学院社会産業理工学研究部

<sup>2</sup> 奈良先端科学技術大学院大学先端科学技術研究科

が想定された介入方法についても、効果的な実践やその効果の評価には試行錯誤の期間が必要であり、クライアントに適した介入方法の策定・調整には一定の期間を要することが少なくない。そのため、クライアントの困難状況を維持・増悪する問題構造を迅速に理解し、クライアントに最適な介入方法を早期に提案するための手法の開発は重要であると考えられる。本稿では、このような個別化されたアセスメントと介入方法の選択・拡充に寄与するアプローチとして、「機械学習 (machine learning)」の利活用を提案する。

機械学習は、近年注目を集めている人工知能技術の一分野であり、問題に対して最適な解決策に到達するための方法やパラメータを自動的に決定する計算戦略である。このアプローチでは、多次元データセットから関数関係を学習することによって、個人の状態に焦点をあてた予測モデルを適用することが可能である。そのため、個性記述的な認知行動療法、ならびに臨床心理学研究にとって、機械学習が有する潜在的な応用可能性は極めて大きいと考えられる。しかしながら、これまで国内外において、臨床心理学への応用可能性について論じた研究は少なく、認知行動療法への機械学習の応用方法を展望した論考は見当たらない。

そこで、本論文ではまず機械学習アプローチの手法や特徴、意義を概観する。加えて、これまでのメンタルヘルス領域において、機械学習アプローチが適用されている主な研究テーマを整理した上で、臨床心理学および認知行動療法研究に寄与しうる活用例を紹介する。最後に、機械学習アプローチの留意点や限界に触れながら、今後の応用可能性について論じる。

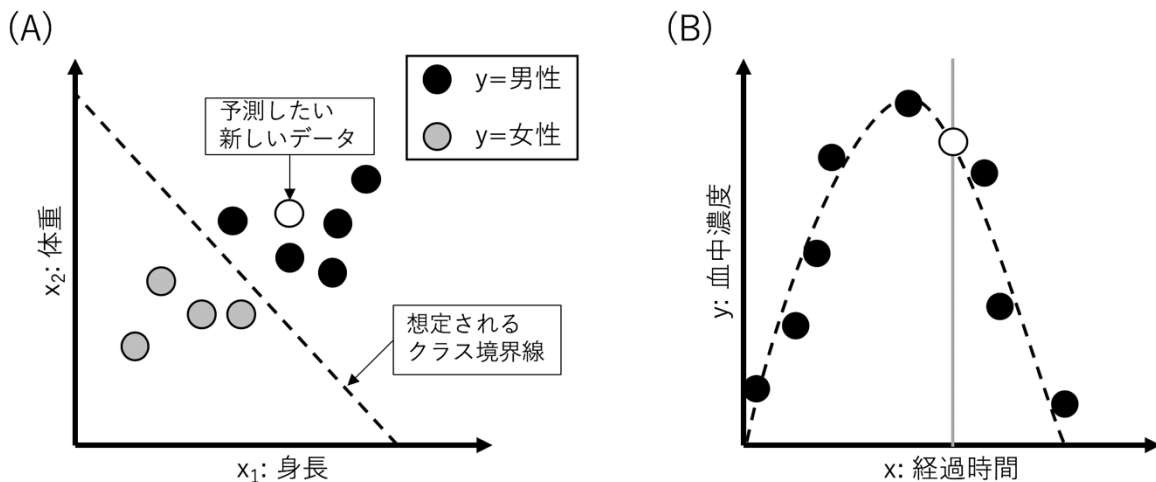
## 機械学習の概観

### 機械学習とは？

機械学習は、1960年頃から人工知能技術の一つとして始まった研究分野であり、その提唱者である Arthur Samuel によると、「*Field of study that gives computers the ability to learn without being explicitly programmed* (明示的にプログラムをすることなくコンピュータに学習能力を与える研究分野)」と定義されている (Samuel, 1959)。ここでいう、学習能力とは、過去の経験に基づいてプログラムの入出力関係を変化させることができる能力のことを意味する。データ分析という応用面を考えると、「これまでに得られたデータ集合からその背後に潜む規則性を発見し、それを基に将来の予測や意思決定を行うためのコンピュータプログラム」と解釈しても良いかと思われる。

### 機械学習の3つの枠組み

機械学習は、対象とする問題の性質によって、「教師あり学習 (supervised learning)」, 「教師なし学習 (unsupervised learning)」, および、「強化学習 (reinforcement learning)」の3つに分類ができる。教師あり学習は、入力とそれに対する理想的な出力 (正解出力) が組となったデータが多数与えられた状況で、入力と出力の間の規則性を発見し、新たな入力に対して適切な出力を予測するためのものである。例えば、健康診断などによって、身長、体重および性別に関する Figure 1(A) のような9人分のデータが得られたとする。そして、新たな人から白丸で示される身長と体重が計測されたときに、その人を男性・女性のどちらの性別と予測すべきかという問題を考えてみる。この場合、点線で示される境界線を想定し、その線より上のデータは男性、下のデータは女性であるという規則性で与えられたデータが説明できる。その規則性に従えば、新たなデータである白丸は男性であろうという予測結果が導ける。



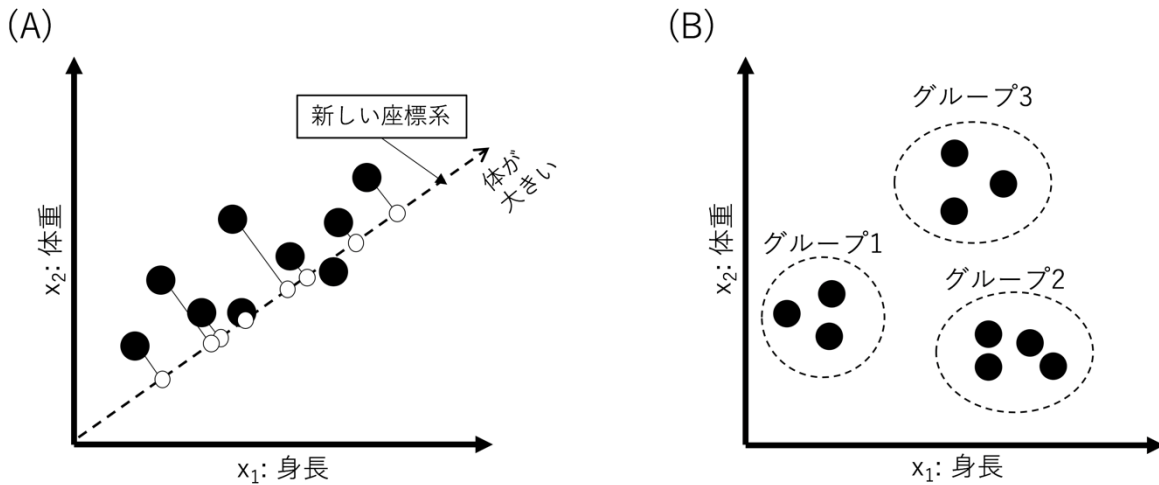
**Figure 1.** 教師あり学習の例. (A) 分類問題の一例. 身長と体重のデータから性別を予測することを想定している。黒色および灰色の点は、それぞれ、男性および女性から得られたデータを表す。白丸は、性別を予測したい人から計測されたデータ点を表す。(B) 回帰問題の一例. 投薬からの経過時間とその薬剤の血中濃度の記録からある時点での血中濃度を予測することを想定している。黒丸は与えられた計測データを表し、灰色の線は血中濃度を予測したい時点を表す。点線は与えられたデータ点の規則性をうまく説明する関数の例を示しており、白丸はこの規則性に基づく予測結果を示している。

また、ある薬剤の投薬試験を行ったとして、投薬後の経過時間とその薬剤の血中濃度の関係が Figure 1(B)のように得られた状況を考える。この場合、これらのデータは、点線で示される放物線にほぼ従っているという規則性を見出すことができる。この規則性を利用すれば、灰色の線で示される経過時間には白丸で表される血中濃度になっているだろうという予測結果が導ける。まさにこれらの例で挙げた推論過程をプログラム化したものが教師あり学習法や教師あり学習アルゴリズムと呼ばれているものである。なお、教師あり学習のうち、Figure 1(A)のように正解出力が離散値（質的データ）として与えられるような場合はクラス分類（classification）や識別（discrimination）と、Figure 1(B)のように正解出力が連続値（量的データ）として与えられるような場合は回帰（regression）と呼び、お互いを区別する場合もある。

教師なし学習では、入力データのみが多数与えられる一方でそれぞれに対する正解出力が一切与えられない状況で、何らかの有意味な情報を抽出したい問題を対象とする。例えば、Figure 1(A)から性別情報を削除した Figure 2(A)のデータが与えられたことを想定する。この時、各データを点線で示される直線に射影すると「体が大きい／小さい」という1次元の視点で各データを比較することができる。また、Figure 2(B)のデータの例では、点線で囲んだ3つのデータの塊を見つけることができ、それに対応したグループ化ができる。いずれも教師なし学習の具体例を示しているが、ここで注意してほしいのは、正解出力がないからといって、でたらめな出力をする訳でない。Figure 2(A)の場合は、「分散（情報量）が最大となるような1次元部分への写像」、Figure 2(B)の場合は、「グループ内分散が最小となるようなデータの分割」という、それぞれ数学的に明確に定義できる基準で最適な出力が決定されることに注意する。なお、教師なし学習法も以下のようにさらに細分化することができる。例えば、Figure 2(A)のように多次元のデータを解釈しやすいように変換するものは「特徴抽出(feature extraction)」と、その中でも元の入力より低次元に変換するものは「次元削減(dimension reduction)」と呼ばれている。また、Figure 2(B)のように近いもの同士が同じグループに属するようにデータ集合を分割するものは「クラスタリング(clustering)」と呼ばれている。

強化学習も明示的な正解出力は与えられないという意味で教師なし学習と分類されることがある。ただし、その出力が理想的なものに近かったか否かを表す指標として報酬（reward）と呼ばれる実数値が与えられることから、教師あり学習と教師なし学習の中間に位置づけられる。ただし、他の2つの学習と比べて強化学習がより特徴的なのは、学習のためのデータそのものもプログラムの出力を

介してサンプリングすること（試行錯誤性）を想定していることであり，そのため，「損して得取れ」という言葉で表されるように，ある時点での出力が将来得られる報酬に影響を与える遅延報酬性を考慮する必要がある。



**Figure 2.** 教師なし学習の例。(A) 次元削減の一例。データは図 1(A)と同じものを想定している。点線で示される部分空間に射影したものを白丸で表している。1次元の白丸に変換された後でも，2次元で表現された元の黒丸のデータの多様性を保持できている。(B) クラスタリングの一例。黒丸で示される 10 個のデータをデータの塊（クラスタ）という観点から点線で囲まれる 3 つのグループに分割でき，それぞれ，身長が低い群（グループ 1），身長が高く体重は軽い群（グループ 2），身長が高く体重も重い群（グループ 3）と解釈できる。

ここで紹介した機械学習の 3 つの分類は，あくまで大分類であり，それぞれに対して多様な手法・アルゴリズムが提案されている。Table 1 は，教師あり学習と教師なし学習でよく用いられる代表的なアルゴリズムをそれらの特徴とともにまとめたものである。強化学習についても，Q 学習法(Watkins & Dayan, 1992)や SARSA 学習法(Rummery & Niranjan, 1994)といった代表的なアルゴリズムが存在するが，問題設定によっては様々な近似解法が必要となるため，詳細な原理や応用例の紹介は(Sutton & Barto, 2018 三上・皆川訳 2000)に任せるものとする。

**Table 1.** 代表的な機械学習手法とその特徴

表内の参考文献と他の推奨される教科書の書誌情報は別表にしてアップロードしている。

学習の枠組み	問題設定	代表的な手法・アルゴリズム	特徴など	参考文献
教師あり学習	分類	判別分析法	同一クラス内の入力データが同じ正規分布に従うことが想定した分類手法。	(Lachenbruch & Goldstein, 1979)
		ロジスティック回帰	クラス間が線形分離可能である（入力空間におけるクラス間の境界面が線形（超平面）となる）ことを想定した分類手法であり、クラス境界は最尤推定法を基に決定される。	(Cox, 1958)
		単純ベイズ分類器	入力が高次元の場合に、各次元が統計的に独立であることを想定した分類手法。各次元の分布が正規分布に従うことを想定した場合は、判別分析法の特別な場合とみなすことができる。	(Jiang, Wang, Cai, & Yan, 2007)
		最小二乗回帰	入力の線形変換で予測される出力と与えられた正解出力の間の二乗誤差を最小にする回帰手法。多項式回帰に代表されるように、事前に設計された非線形関数で入力を変換したものを線形変換する場合も最小二乗回帰の一般化として位置づけられる。	(Fomby, Johnson, & Hill, 1984)
教師あり学習	回帰	部分最小二乗回帰	入力空間・出力空間をそれぞれ低次元の部分空間に射影した後に最小二乗回帰を行う回帰方法。入力が特に多次元の場合に有効となる場合が多い。	(Rosipal & Krämer, 2006)
		多層パーセプトロン（フィードフォワードニューラルネットワーク）	神経細胞の振舞いにヒントを得た形式ニューロン素子を層状かつ順伝線的に構造化した人工ニューラルネットワークモデルを用いた回帰手法。特に、多層パーセプトロンのうち4層以上の構造を持つものは深層ネットワーク・深層学習とも呼ばれている。	(Schmidhuber, 2015)
		K近傍法	入力空間で近傍にあるデータ同士は類似した出力を持つべきという原理に基づいて、近傍K個（Kは事前に設定された整数値）のデータの多数決や平均で出力を決定する。	(Peterson, 2009)
		サポートベクターマシン (SVM; support vector machine)	マージン最大化原理（クラス境界面とデータ点との最短距離を最大化すべきという基準）に基づいて、線形（超平面）のクラス境界を定める。もともとは、分類問題のために開発された手法であり、それはサポートベクター分類（support vector classification; SVC）とも呼ばれる。それを回帰問題に拡張した手法はサポートベクター回帰（support vector regression; SVR）と呼ばれる。	(Cortes & Vapnik, 1995) (Drucker, Surges, Kaufman, Smola, & Vapnik, 1997) (Cervantes, Garcia-Lamont, Rodríguez-Mazahua, & Lopez, 2020)
教師なし学習	分類&回帰	決定木	入力に対するIf-thenルールに基づいて出力を決定する手法。If-thenルールに当たる部分をノード（節点）とすると木構造で表現することができるためこのような名前と呼ばれる。	(Safavian & Landgrebe, 1991)
		ランダムフォレスト	バギングというサンプリング手法を用いて表現能力の低い（弱学習器）である決定木を複数統合させて汎化能力の向上を目指したアンサンブル学習法の一つ。	(Breiman, 2001) (Biau, 2012)
		勾配ブースティング	ブースティングというサンプリングを用いて弱学習器（決定木の場合が多い）を複数統合させて汎化能力の向上を目指したアンサンブル学習法の一つ。	(Friedman, 2001) (Natekin & Knoll, 2013)
		階層クラスタリング	各データ点をそれぞれ1つのグループ（クラスタ）とする初期状態からスタートして、近傍のクラスタを順に統合することでクラスタ構造を構成する手法。	(Murtagh & Contreras, 2012)
教師なし学習	クラスタリング	K平均法	初期状態としてK個クラスタの代表点を定めた後、1) 各データ点を最近傍の代表点を持つクラスタに割り当てる、2) 割当てられたデータの平均値でクラスタの代表点を更新するという2つの手続きを収束するまで繰り返すことによってクラスタを決定する手法。	(MacQueen, 1967) (Blömer, Lammensen, Schmidt, & Sohler, 2016)
		混合ガウス分布モデル	各データが中心と平均が異なる複数の正規分布のうちのいずれかから生成されたことを仮定して、ベイズ事後確率最大化基準に基づいてクラスタを決定する手法。ある特別な条件下では、K平均法と等価になる。クラスタ数も確率変数とみなすことによって、ベイズ最適化原理の意味でクラスタ数の自動決定することも可能と提案されている。	(Melnykov & Maitra, 2010)
		主成分分析	データを分散が最大となるような部分空間へ線形変換する次元削減手法。正規分布を仮定した場合はシャノンの情報量最大となる部分空間への線形変換とみなすこともできる。	(Shlens, 2014)
教師なし学習	次元削減 特徴抽出	多次元尺度構成法	データ間の距離（類似度、位相関係）を可能な限り保存しながら、低次元空間へ変換する方法。	(Saeed, Nam, Haq, & Muhammad Saqib, 2018)
		独立成分分析	データが統計的に独立で非ガウス性を持つ潜在変数の合成変数によって生成することを想定し、その逆問題を解くことによってデータを独立な成分に分解する手法。	(Bell & Sejnowski, 1995) (Hyvarinen, 1999)
		非負値行列因子分解	要素がすべて非負という制約を持つ行列の低ランク近似を用いて、データを低次元空間へ変換する方法。	(Lee & Seung, 1999) (Wang & Zhang, 2013)
		オートエンコーダ	構造に制約を入れた多層パーセプトロンを用いて入力と同じ出力に変換（恒等写像）するように学習させた手法であり、各データに対する中間層の出力が特徴抽出後のデータとして用いられる。	(Hinton, 2006) (Bank, Koenigstein, & Giryès, 2020)

## 機械学習と統計学との違い

Table 1 を見ると、最小二乗回帰やロジスティック回帰、クラスタリングなど統計解析の教科書でよく目にする用語が多い。実際、機械学習のアルゴリズム開発や理論解析の際には、多変量解析を中心とした統計学の知見を利用することが多く、応用統計学の一つという側面もある。Bzdok, Altman, & Krzywinski (2018) でも議論されているように、あえて細かい違いを挙げるとすると、統計学は、記述統計のように与えられたデータの要約の仕方や、仮説検定のようにデータを解釈する上での根拠を与えるための理論体系を扱う分野である。一方で、機械学習は将来の意思決定に役立てられるように新規データに対してできる限り正確な予測をすることを目指した分野である。それに加えて、機械学習では、適用場面に応じた制約時間内に有益な予測結果を得ることができるコンピュータプログラムでなければならないというのも重要な要件とされる。そのため、数学から派生した統計学の厳格性に比べると、機械学習の方が近似やヒューリスティック (heuristics; 厳密な理論保証がない、試行錯誤から得られた経験則) も許容される傾向にある。

## なぜ今機械学習に注目するのか？

最初に述べた通り、機械学習に関する研究は人工知能に端を発しているが、近年ではデータ解析の主要ツールとして様々な研究領域で利用されるようになった。その背景にはいくつかの要因がある。

一つは、機械学習アルゴリズムのパッケージ化 (ライブラリ化) が進み、ユーザは詳細な原理や実装法を知らなくても、それらをブラックボックス的にしかも無償で利用できるようになったことである。例えば、Python では `scikit-learn`<sup>3</sup>, R では `caret`<sup>4</sup> や `mlr`<sup>5</sup> といったライブラリがオープンソースとして利用することができる。また、Java により開発された `Weka`<sup>6</sup> というフリーソフトウェアも公開されている。このソフトウェアは GUI も提供されているため、プログラムを書かなくても、機械学習を用いた簡単なデータ解析を実行することができる。

また、科学のパラダイムシフトも大きな要因の一つである。これまでの主流は、過去の研究や事例をサーベイし、そこから類推されるある仮説を事前に設けて、それを実験や調査で実証するという、仮説駆動型科学 (hypothetical-driven

<sup>3</sup> <https://scikit-learn.org>

<sup>4</sup> <https://cran.r-project.org/web/packages/caret/index.html>

<sup>5</sup> Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., ... Jones, Z. M. (2016). mlr: Machine Learning in R. *Journal of Machine Learning Research*, 17(170), 1-5.

<sup>6</sup> <https://www.cs.waikato.ac.nz/ml/weka/>

science)であった。一方で、計測技術の発展や情報インフラの整備に伴い、大量かつ多種多様なデータを取得・集積することや、それらをインターネットを介して共有したり、メタ解析のような別の用途に再利用したりすることが容易になってきた。いわゆるビッグデータ時代の到来である。それに伴い、そのビッグデータから有益な情報を抽出し、新たな知見や仮説を得ようとするデータ駆動型科学(data-driven science)への取り組みも急速に増えた。規則性を見つけ、予測・意思決定に役立てるといふ機械学習の特徴は、まさにデータ駆動型科学で目指しているものであり、そのニーズとうまく合致した点も機械学習が浸透した大きな理由であると考えられる。

技術面から見ると、「 $p \gg n$ 問題」に対する有望な解決策が長年研究されてきた点も大きい(Johnstone & Titterton, 2009)。ここで、 $n$ はサンプル数を、 $p$ は変数の数(各サンプルの属性ベクトルの次元数)をそれぞれ意味しており、変数が膨大であるにも関わらず、解析に利用できるサンプル数が非常に限られた状況で生じる問題を指している。この問題に対して、統計解析でよく用いられる2標本 $t$ 検定や回帰分析を単純に適用しても、それぞれ、多重比較補正による検出力の低下や解の不定性のために有意な結果を得ることができない。一方で、機械学習の分野では、L1正則化の導入やARD(Automatic Relevant Determination)事前分布を仮定したベイズスパースモデリングによる自動変数削減法が存在する(Neal, 1996; Tibshirani, 1996, 2011; O. Yamashita, Sato, Yoshioka, Tong, & Kamitani, 2008)。もちろん、これらを用いても分類や回帰に寄与する変数を完全に同定することは不可能であるが、入出力間に何らかの規則性があれば、ランダムよりは有意な予測を実現できることが多い。この特性は、ヒトを対象とし、多数のデータサンプルを収集することが困難な状況が多いメンタルヘルス領域にとっては恩恵の多いものであると考えられる。

### メンタルヘルス領域における機械学習アプローチの臨床応用

メンタルヘルス領域において、機械学習アプローチを活用した主要な研究テーマとしては、(1)精神疾患の診断、(2)症状・リスクの検出、(3)予後や治療効果の予測が挙げられる。研究によって扱われる指標はさまざまだが、時には数百から数千にも及ぶ変数群に機械学習を適用することで、研究知見が見出されてきた(Hilbert et al., 2020; A. Yamashita et al., 2020)。

#### 精神疾患の診断

臨床心理学と精神医学の領域において、精神疾患の診断が早期に確定すること



は、効果的な治療・介入方法の策定や、経過・予後の予測につながる。こうした診断への寄与を目的とした機械学習研究では、特に構造的・機能的脳画像を用いることによって、診断の再現性に焦点をあてた検討が重ねられてきた。

これまでの主な研究対象として、うつ病(Fu et al., 2008)や、統合失調症(Skåtun et al., 2017)、アルツハイマー病(Dimitriadis, Liparas, & Tsolaki, 2018)に関して検討が行われ、近年では不安症(Lueken et al., 2015)や食欲不振症(Lavagnino et al., 2015)など、対象とされる病態が広がりつつある。たとえば、未投薬のうつ病患者 19 名を対象とし、悲しみ表情の呈示下における機能的脳活動データに対して、教師あり学習法（サポートベクターマシン）を適用することで、86%の精度で急性期のうつ病と健常者を分類可能であった(Fu et al., 2008)。同様に、3 拠点において収集された 182 名の統合失調症スペクトラム患者の安静時脳活動データに対して、教師あり学習法（正規化判別分析）を適用し、異なるスキャナーやサンプルにおいても、高い精度（最大 80%）で分類されていた(Skåtun et al., 2017)。こうした神経画像を活用した機械学習の精度は、撮像データの種類によって異なるものの、近年のメタ分析の結果から、うつ病患者では全体的に 77%の感度と 78%の特異度であり(Kambeitz et al., 2017)、統合失調症患者ではいずれも 80.3%であったことが示されている(Kambeitz et al., 2015)。

神経画像データ以外にも、統合失調症患者の面接時における表情表出を記録した動画データ(Tron, Peled, Grinsphoon, & Weinshall, 2016)や、アルツハイマー病患者の認知課題時における発声データ(König et al., 2015)、自殺念慮のある若者の面接時における会話記録(Pestian et al., 2016)に対して機械学習法を適用することで、一定の精度の診断分類が達成されている。

以上のような診断補助を目的としたものにくわえ、アクチグラフから測定された活動指標に対して、教師あり学習法（サポートベクターマシン）を用いることで、注意欠如・多動症と双極性障害の子どもを分類可能であることが示されており(Faedda et al., 2016)、症状の鑑別にも寄与する研究が行われている。

### 症状・リスクの検出

症状の早期検出やリスクの特定は、疾患の発症や重症化を防ぐための予防教育や早期介入を行うために重要である。

これまで、脳神経画像に対して機械学習法を適用することで、アルツハイマー病を示す特異的な脳活動パターンの同定(Doan et al., 2017)や、精神病発症の早期予測(Koutsouleris et al., 2012)、機能的結合性の観点からのうつ病再発をもたらす脆弱性の検出(山本, 2018)がなされてきた。他にも、カウンセリングの対話記録を用いた自殺念慮の検出(Oseguera, Rinaldi, Tuazon, & Cruz, 2017)や、患者

が提出した文章からの統合失調症症状の検出(Strous et al., 2009)、音声データの活用によるうつ病や統合失調症と関連した精神状態の検出(Kliper, Portuguese, & Weinshall, 2016)といった、さまざまな情報を活用した試みがなされている。

日常生活で収集可能なデータに対して機械学習法を適用した研究として、個人の日々の認知・行動・生理的情報（以下、ライフログ）を用いて頭痛の発生要因の検出が試みられている(Yamamoto & Yoshimoto, 2018)。この研究では、スマートリストバンドとスマートフォンがデータ収集に用いられ、約80日間測定された個人のライフログデータに対して、教師なし学習法（クラスタリング）を適用することで、頭痛と関連する要因の相互作用構造を網羅的に可視化している(Figure 3)。この研究は、ライフログの有用性を示唆することに加えて、個人のデータのみから本人に特化した予測モデルを構築することが可能であることを示している。

また、個人における症状やリスクの検出以外にも、集団データにおける変数の複雑な相互作用構造を網羅的に抽出し、集団内におけるさまざまな状態像やサブタイプを可視化することも可能である。たとえば、緊急事態宣言下において実施された、日本国民のメンタルヘルスに関する大規模オンライン調査データに対して、同様の教師なし学習法が適用されている(Yamamoto, Uchiumi, Suzuki, Yoshimoto, & Murillo-Rodriguez, 2020)。これによって、重度の心理的ストレスを感じている層の背景には、孤独感の高さや、身近な人物との対人関係の悪化、新型コロナウイルス感染症に関連する不眠や不安、家計の悪化、仕事・学業といった、多様な要因の相互作用があることが示され、さまざまな機関による困難に応じたサポートの必要性が示唆されている。

### 予後や治療効果の予測

臨床心理学や精神医学において、個人の予後や治療方法の効果を予測できることは、症状の経過の判断や、心理教育などの予防的関わり、最適な心理的・薬理的介入の選択などのために非常に重要である。

これまで、患者の予後や治療効果の予測を目的として、慢性的なうつ病患者における2年間の自然経過(Schmaal et al., 2015)、統合失調症患者における抗精神病薬の6週間後の治療効果(Bak et al., 2017)、アルツハイマー病患者らにおける24ヶ月間にわたる認知機能の低下(Zhu et al., 2016)、退院から3ヶ月後における子どもの心的外傷後ストレス障害の発症(Saxe, Ma, Ren, & Aliferis, 2017)など、さまざまな病態の予測に機械学習が適用されてきた。他にも、精神状態についての質問に対する患者から得られた自由記述の回答を活用して、自殺念慮と精神症状の悪化を一定の精度で予測できる結果が得られている(Cook et al., 2016)

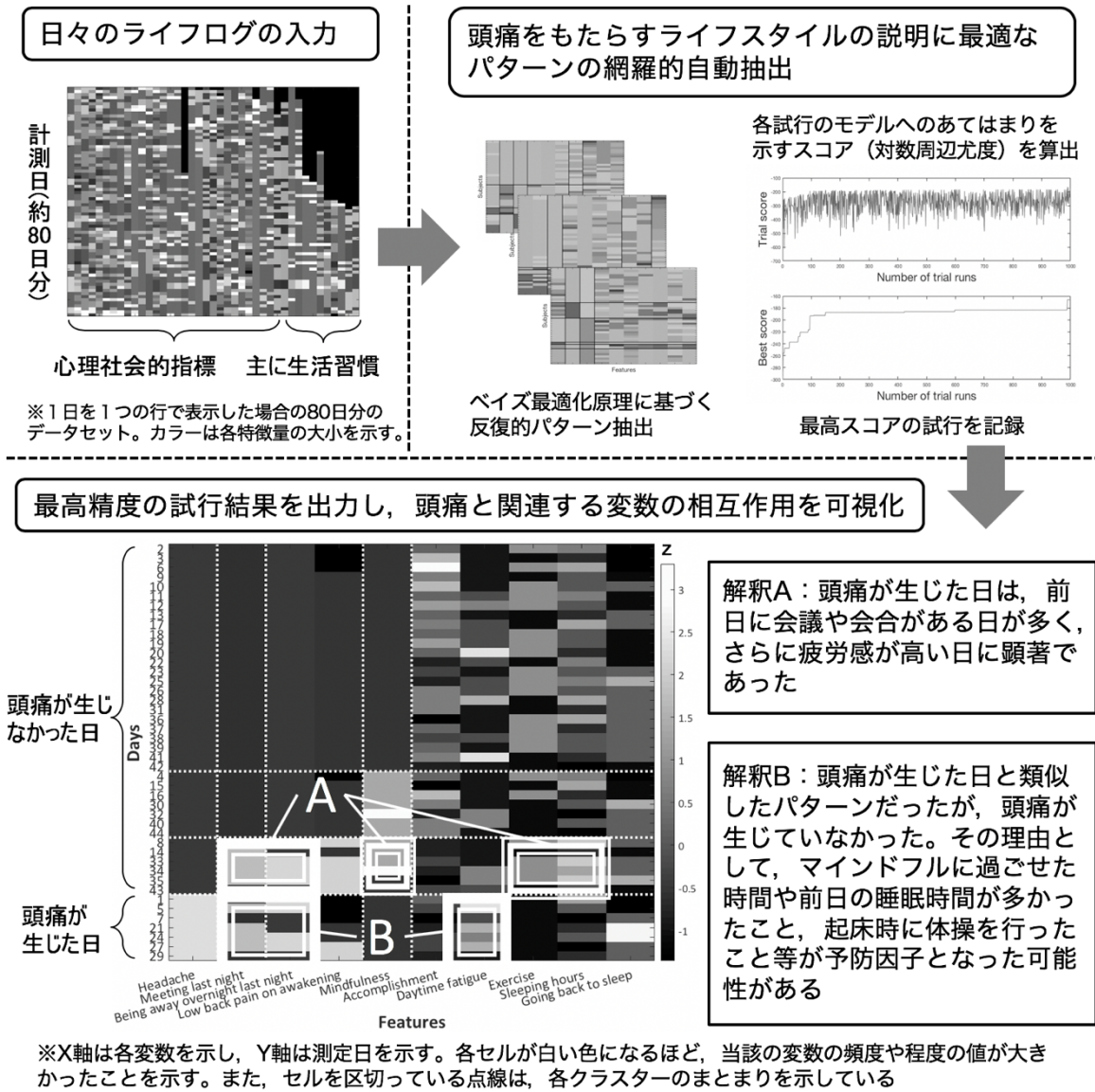


Figure 3. 頭痛に関わる変数の相互作用を可視化するノンパラメトリック・ベイズ共クラスタリングの概要と解釈の例

うつ病に対する認知行動療法の治療効果の予測を試みた研究も報告されている (Tymofiyeva et al., 2019)。この研究では、30名の思春期うつ病の患者に対して3ヶ月間の認知行動療法を実施しており、実施前の構造的コネクトームデータ (拡散テンソル画像から推測される脳領域間の神経繊維の接続状態を、全脳レベ

ルで定量化したデータ)と抑うつ得点に教師あり学習法(C4.5アルゴリズムに基づく決定木)を適用することで、83%の精度で認知行動療法実施後の抑うつ症状の軽減が予測された。これらの結果は、神経画像を活用した機械学習アプローチが、思春期のうつ病に対する認知行動療法の効果を高精度に予測しうることを示している。そのため、こうした手法を活用することによって、特定の患者に対する認知行動療法の有効性の予測が可能となり、治療計画の策定に役立つ可能性がある。

一方で、日常臨床において取得可能な社会人口統計学的データ(たとえば、年齢や性別など)や臨床的データ(たとえば、症状の重症度など)に対して教師あり学習法(ランダムフォレスト他)を用いた研究では、認知行動療法による治療効果を有意に予測できたものの、その予測精度は59%と不十分であり、臨床的有用性には限界が残ることが指摘されている(Hilbert et al., 2020)。そのため、機械学習を日常臨床で広く活用するためには、臨床現場で容易に収集可能な予測に寄与しうるデータの同定と、こうしたデータに対して高精度の予測性能を発揮する学習アルゴリズムを開発することが重要であると考えられる。

#### 機械学習アプローチの限界

以上のように機械学習アプローチには、多くの利点と可能性がある一方で、欠点と限界も存在する。おそらく最大の欠点は、与えられたデータの中から規則性を見つけることが前提となっているため、探索的研究や事後解析研究に留まりがちな点である。新たな仮説生成という意味では、それでも十分に意義はあるかもしれない。しかし、臨床現場において機械学習による予測が汎用的に役に立つものなのかを検証するためには、独立コホートによる前向き試験の実施が不可欠である。例えば、試験の前に学習アルゴリズムやそれに付随するパラメータは完全に固定しておいて、対象群(既存のガイドラインに従う治療が施される患者群)よりも機械学習の予測に基づいて治療方法を選択したほうが統計的に有意に治療効率が增大したという検証まで完了して、初めて科学的な意味で機械学習による予測が治療選択に貢献できたと結論づけることができる。その意味では、機械学習はあくまで方法論の一つを与えるものであって、研究全体を全て自動化するものではないことを我々は留意しなければならない。

また、予測に有用な指標や因子(変数)がデータに含まれていなければ、どんな機械学習アルゴリズムを使ってもチャンスレベルの予測しかできない。したがって、データ取得時には、少しでも予測に寄与する可能性のある指標はできる限りすべてを正確に計測するとともに、数値化する際にもその情報が欠落したり、バイアスが入ったりしないよう注意が必要である。

パッケージやライブラリのオープンソース化は、気軽に機械学習アルゴリズムを利用できるようになった反面、その中身のブラックボックス化が進んでいる。予測精度の向上だけを考えるのであれば、コンピュータビジョンや機械翻訳、音声認識など様々な分野で成功を収めている深層学習の活用は非常に有望である。しかしながら、予測を導くための関数（入力から予測出力結果へと変換する数式）は非常に複雑であり、その規則性を我々人間が解釈することは一般に困難である。メカニズム研究のように一般原理を探索することが目的であれば、変数選択付きロジスティック回帰や線形回帰のような簡単なアルゴリズムの方が適している場合もある。

また、機械学習の分野では「ノーフリーランチ定理 (No Free Lunch theorem)」(Wolpert & Macready, 1997)という言葉がある。これは、あらゆる任意の問題で性能の良いアルゴリズムを構築することは、そもそも不可能であり、本当に性能の良いアルゴリズムを導出したければその問題に特化した事前知識が必要となることを示した理論である。したがって、何を目的とし、それを達成するためにはどの手法が良いかを機械に任せることは、(少なくとも現在のところは)実現可能性に乏しい。よって、その分野の専門家である精神科医や臨床心理士と各アルゴリズムの特質をよく知っている機械学習の専門家のタイトな共同研究が、本稿で述べている新たな研究アプローチの成功の鍵であると著者らは考えている。

### 機械学習アプローチの活用に関する今後の展望

#### 臨床場面におけるアクセシビリティの向上と有用性の検証

これまでの機械学習アプローチにおいては、特にうつ病や自殺リスク、認知機能についての診断と予測の精度に焦点が当てられており、概ね良好な精度をあげている。そのため、今後こうしたアプローチが十分に確立されることによって、精神疾患の発見、診断、予測などに大きく寄与することが期待される。

一方、用いられる解析技術の精度やデータセットは研究間で一貫性が乏しく、臨床実践において活用可能な技術にするためには、多くの機関で適用可能な標準化された手法を提起する研究が必要である。また、一定の精度をあげている研究のほとんどは、神経画像データを用いた教師あり学習による分類技術を用いている。しかしながら、日常臨床の現場では、問診や質問紙などによる診断・評価が一般的であり、これらのデータを用いた高精度な診断補助技術の開発が望まれる。

さらに、上述のように先行研究の多くは、すでに得られたデータのみを活用して、レトロスペクティブ（後方視的）に予測精度の検討を行なっている。そのた

め、臨床現場において新たに得られるデータへのプロスペクティブ（前方視的）な精度については、いまだ不明確である。よって、前向き試験の活用による臨床的な有効性におけるエビデンスの蓄積が必要であるといえる。臨床現場では、研究のためにスクリーニングされた患者（たとえば、ある程度症状が均質と考えられる患者など）ではなく、さまざまな併存疾患を呈したり、難治化している外来患者も少なくなく、精度を十分に確保するためには新たな検討課題が生じる可能性がある。また、研究知見があまり蓄積されていない他の精神疾患（不安症や、神経発達障害など）に対する予測精度や、機械学習アプローチが最適となりうる疾患については不明であるなど、未解決な課題が多く残されている。

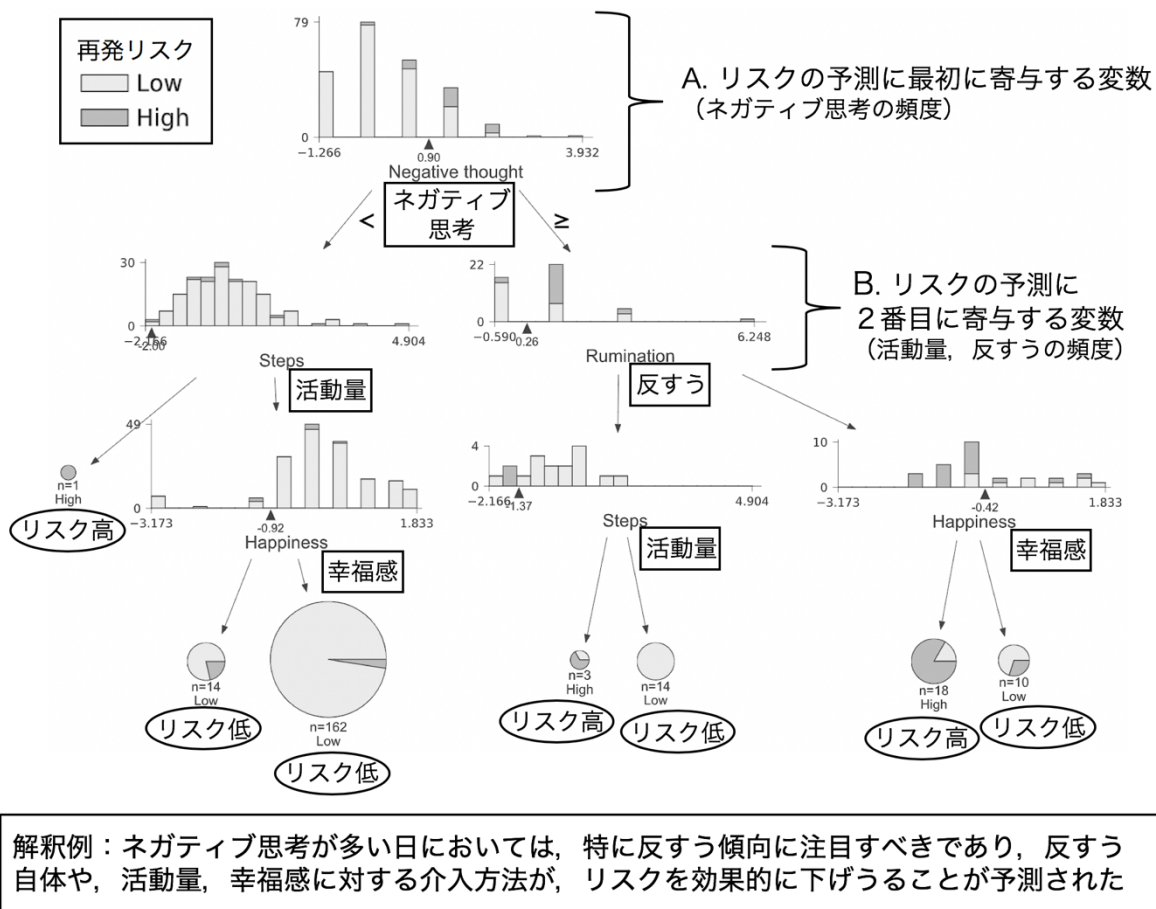
他にも、ライフログに対して教師あり学習法（サポートベクターマシン）を適用し、幸福感の予測を行なった研究は一部あるものの(Yamamoto, Yoshimoto, Murillo-rodriguez, & Machado, 2019)、こうしたポジティブな心理的状态（レジリエンス、個人の成長など）に機械学習法を適用した研究は数少ない(Shatte, Hutchinson, & Teague, 2019)。そのため、症状などの不適応的側面のみならず、クライアントの適応的側面の促進に焦点をあてた機械学習の応用研究も、今後は必要であると考えられる。

#### 機械学習アプローチに基づく個別化された介入手続きの開発

これまで広く用いられてきた「古典的推論アプローチ」(グループの分散や誤差に対して、 $p$  値に基づく検定や効果量の観点から、研究をデザインし結果を評価する手法)は、特定の集団の比較や理解には有用であり、臨床心理学や精神医学研究を大きく推進してきた。一方で、変数間の複雑な相互作用や個人差が除外される傾向にあり、結果的に個人への研究知見の適用が困難であった(Dwyer, Falkai, & Koutsouleris, 2018)。これに対して、機械学習アプローチでは、個人に対する一般化可能な予測を行うことを大きな特長としている。このことは、クライアントにとって最も効果的・効率的な介入方法の選択に際して、機械学習アプローチが寄与する可能性を示唆している。

このような個別化された機械学習アプローチは、心理療法や薬理的治療法を選択する際に有用であると考えられ、今後の応用に寄与する提案や基礎研究がなされてきた。たとえば、うつ病の予測と対処を目的とする、利用者に個別化された評価・介入システムの提案(Yang, Zhou, Duan, Hossain, & Alhamid, 2018)、携帯電話のセンサーデータを活用したハイリスクな飲酒の検出(Bae ら, 2018a)、損失金額に関する個別化されたフィードバックによる賭博行動の変容プロセスの検討(Auer & Griffiths, 2018)などの研究は、個人に最適化されたタイミングでの介入を提供するための示唆を与えている。

また、特定のうつ病寛解患者から収集された約8ヶ月間のライフログデータに対して、教師あり機械学習法（ランダムフォレスト）を活用することで、うつ病の再発リスクを最も下げうる変数を同定する試みがなされている(山本 & 吉本, 2019)。以下に、この研究において出力された代表的な予測結果（Figure 4）を用いて、学習によって得られた分類規則を個別化治療に役立てるための応用方法を提案する。



**Figure 4.** うつ病の再発リスクを予測するランダムフォレストの結果と解釈の例。なお、この Figure 4 自体は複数の予測結果のうちの1つであり、必ずしも最高精度の予測結果を指すものではない。

解析結果から、この患者においては、ネガティブな思考が最も再発兆候を規定しうる特徴量（変数）であり（Figure 4.A）、その次には活動量と反すうが、再発に大きな影響を有することが示されている（Figure 4.B）。すなわち、ネガティブな自動思考や活動量、反すうといった変数を優先的にターゲットにするべきであることを示唆しており、この予測結果は認知行動療法による抑うつ治療モデルや治療手続きと合致している。そのため、この患者にとってはうつ病の再発予防に対して認知行動療法を活用したアプローチが有効であることが推定される。

機械学習法を活用したこうした臨床的アプローチは、いまだ端緒についたばかりであり、現状では臨床現場での実用可能性を改善するためのさらなる基礎的知見の蓄積が必要である。しかしながら、昨今注目を集めている精密医療（precision medicine）の流れと同様に、クライアントに個別化された治療をよりよく実践するための大きな可能性を秘めているといえる。

#### 専門家間・研究チーム間の協力体制

機械学習で構築されるモデルは、実際に学習されるデータの質によって、その精度は大きく影響を受ける。そのため、よりよい学習データの収集と共有を可能とする環境整備が必要となる。上述のノーフリーランチ定理が事前知識の重要性を示しているように、データ収集時には、臨床家と研究者の協働的な取り組みが求められるだろう。同様に、開発されたモデルの臨床的な有用性を最大限に高めるためには、臨床家や患者からのフィードバックと、それに対する研究者によるモデルの改善が必要である。また、倫理的な面を配慮した機械学習アプローチの活用には、専門家、意思決定者、利用者を含む学際的なチームによって行われることの重要性が指摘されている（Wiens et al., 2019）。以上を踏まえると、機械学習アプローチの性能を十分に発揮させるためには、これまで以上に研究者・臨床家・患者の間の協力関係は重要だと言える。

さらに、さまざまな個人に対して適用可能な手法（一般化可能なモデル）を構築するには、性差や、年齢層、人種や居住地といった人口統計学的な属性や計測機器の違いによるバイアスの効果をしっかり検証し、それらを補正する手法開発も重要である（A. Yamashita et al., 2019）。したがって、異なる研究機関や異なる研究プロジェクトによるデータ収集に加え、それを共有するための研究グループ間の協力や、それを可能とする仕組み作りが今後は重要になるといえる。



## おわりに

機械学習アプローチを臨床現場で活用するには、容易に実施可能な手法の開発や技術的な改善が必要であり、いまだ研究知見との隔たりが大きい。しかしながら、メンタルヘルス領域における機械学習アプローチの急速な広がりを鑑みると、今後、認知行動療法や臨床心理学におけるさらなる応用研究の進展が期待される。認知行動療法の基本原則の一つである「協働的実証主義」(クライアントと協働的に問題解決に取り組むセラピストの姿勢)と同様に、クライアントとセラピストが、機械学習をはじめとした人工知能技術と「協働的」に問題に取り組むアプローチが一般的になる日も、そう遠くはないかもしれない。

また、近年では、情報通信技術の進展により、携帯情報端末や PC を用いた認知行動療法の実践(山本 & 竹林, 2020)や、仮想現実 (VR) を活用したセルフカウンセリング(Slater et al., 2019; 山下・山本, 2021)、ゲームをベースとしたアプリケーションによる認知機能の改善手法(Kollins et al., 2020)をはじめ、さまざまな情報通信技術を活用した介入方法が開発されている。こうした技術を、心の理解・予測・調整に役立てる「臨床心理情報学」といったアプローチのさらなる広がりがますます期待される。

多くのアプローチが認知行動療法の名のもとに統合され、発展してきたように、機械学習アプローチをはじめとした新たな展開は認知行動療法をさらに拡充する可能性を秘めている。一方で、認知行動療法はこれまで、その効果や手続きを繰り返し検証することによって、有効性を実証してきた。そのため、認知行動療法が大事にしてきた実証性 (エビデンス) に十分留意しながら、今後も新たなアプローチの有効性を注視し続ける必要があると考えられる。

## 引用文献

- Auer, M., & Griffiths, M. D. (2018). Cognitive Dissonance, Personalized Feedback, and Online Gambling Behavior: An Exploratory Study Using Objective Tracking Data and Subjective Self-Report. *International Journal of Mental Health and Addiction*, 16, 631–641.
- Bak, N., Ebdrup, B. H., Oranje, B., Fagerlund, B., Jensen, M. H., Düring, S. W., … Hansen, L. K. (2017). Two subgroups of antipsychotic-naïve, first-episode schizophrenia patients identified with a Gaussian mixture model on cognition and electrophysiology. *Translational Psychiatry*, 7(4), e1087.

- Bzdok, D., Altman, N., & Krzywinski, M. (2018). Points of Significance: Statistics versus machine learning. *Nature Methods*, *15*, 233–234.
- Cook, B. L., Progovac, A. M., Chen, P., Mullin, B., Hou, S., & Baca-Garcia, E. (2016). Novel Use of Natural Language Processing (NLP) to Predict Suicidal Ideation and Psychiatric Symptoms in a Text-Based Mental Health Intervention in Madrid. *Computational and Mathematical Methods in Medicine*, *2016*, 8708434.
- Dimitriadis, S. I., Liparas, D., & Tsolaki, M. N. (2018). Random forest feature selection, fusion and ensemble strategy: Combining multiple morphological MRI measures to discriminate among healthy elderly, MCI, cMCI and alzheimer's disease patients: From the alzheimer's disease neuroimaging initiative (ADNI) data. *Journal of Neuroscience Methods*, *302*, 14–23.
- Doan, N. T., Engvig, A., Zaske, K., Persson, K., Lund, M. J., Kaufmann, T., ... Westlye, L. T. (2017). Distinguishing early and late brain aging from the Alzheimer's disease spectrum: consistent morphological patterns across independent samples. *NeuroImage*, *158*, 282–295.
- Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology*, *14*, 91–118.
- Faedda, G. L., Ohashi, K., Hernandez, M., McGreenery, C. E., Grant, M. C., Baroni, A., ... Teicher, M. H. (2016). Actigraph measures discriminate pediatric bipolar disorder from attention-deficit/hyperactivity disorder and typically developing controls. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *57*(6), 706–716.
- Fu, C. H. Y., Mourao-Miranda, J., Costafreda, S. G., Khanna, A., Marquand, A. F., Williams, S. C. R., & Brammer, M. J. (2008). Pattern Classification of Sad Facial Processing: Toward the Development of Neurobiological Markers in Depression. *Biological Psychiatry*, *63*(7), 656–662.
- Hilbert, K., Kunas, S. L., Lueken, U., Kathmann, N., Fydrich, T., & Fehm, L. (2020). Predicting cognitive behavioral therapy outcome in the outpatient sector based on clinical routine data: A machine learning approach. *Behaviour Research and Therapy*, *124*, 103530.
- Johnstone, I. M., & Titterton, D. M. (2009). Statistical challenges of high-dimensional data. *Philosophical Transactions of the Royal Society A:*

- Mathematical, Physical and Engineering Sciences*, 367(1906), 4237–4253.
- Kambeitz, J., Cabral, C., Sacchet, M. D., Gotlib, I. H., Zahn, R., Serpa, M. H., … Koutsouleris, N. (2017). Detecting Neuroimaging Biomarkers for Depression: A Meta-analysis of Multivariate Pattern Recognition Studies. *Biological Psychiatry*, 82(5), 330–338.
- Kambeitz, J., Kambeitz-Ilankovic, L., Leucht, S., Wood, S., Davatzikos, C., Malchow, B., … Koutsouleris, N. (2015). Detecting neuroimaging biomarkers for schizophrenia: A meta-analysis of multivariate pattern recognition studies. *Neuropsychopharmacology*, 40, 1742–1751.
- Kliper, R., Portuguese, S., & Weinshall, D. (2016). Prosodic analysis of speech and the underlying mental state. In S. Serino, A. Matic, D. Giakoumis, G. Lopez, & P. Cipersso (Eds.), *Pervasive Computing Paradigms for Mental Health. MindCare 2015. Communications in Computer and Information Science* (vol 604, pp. 52–62). Springer, Cham.
- Kollins, S. H., DeLoss, D. J., Cañadas, E., Lutz, J., Findling, R. L., Keefe, R. S. E., … Faraone, S. V. (2020). A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): a randomised controlled trial. *The Lancet Digital Health*, 2(4), e168–e178.
- König, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., … David, R. (2015). Automatic speech analysis for the assessment of patients with predementia and Alzheimer’s disease. *Alzheimer’s and Dementia: Diagnosis, Assessment and Disease Monitoring*, 1(1), 112–124.
- Koutsouleris, N., Borgwardt, S., Meisenzahl, E. M., Bottlender, R., Möller, H. J., & Riecher-Rössler, A. (2012). Disease prediction in the at-risk mental state for psychosis using neuroanatomical biomarkers: Results from the fepsy study. *Schizophrenia Bulletin*, 38(6), 1234–1246.
- Lavagnino, L., Amianto, F., Mwangi, B., D’Agata, F., Spalatro, A., Zunta-Soares, G. B., … Soares, J. C. (2015). Identifying neuroanatomical signatures of anorexia nervosa: A multivariate machine learning approach. *Psychological Medicine*, 45(13), 2805–2812.
- Lueken, U., Straube, B., Yang, Y., Hahn, T., Beesdo-Baum, K., Wittchen, H. U., … Kircher, T. (2015). Separating depressive comorbidity from panic disorder: A combined functional magnetic resonance imaging and machine learning approach. *Journal of Affective Disorders*, 184, 182–192.

- Neal, R. (1996). *Bayesian Learning for Neural Networks. Ph.D. Thesis.* University of Toronto.
- O'Brien, W. H., Haynes, S. N., & Kaholokula, J. K. (2016). Behavioral assessment and the functional analysis. In C. M. [Ed] Nezu & A. M. [Ed] Nezu (Eds.), *The Oxford handbook of cognitive and behavioral therapies.* (pp. 44–61). New York, NY: Oxford University Press. Retrieved from
- Oseguera, O., Rinaldi, A., Tuazon, J., & Cruz, A. C. (2017). Automatic quantification of the veracity of suicidal ideation in counseling transcripts. In C. Stephanidis (Ed.), *HCI International 2017 – Posters' Extended Abstracts. HCI 2017. Communications in Computer and Information Science* (vol 713, pp. 473–479). Springer, Cham.
- Pestian, J. P., Grupp-Phelan, J., Bretonnel Cohen, K., Meyers, G., Richey, L. A., Matykiewicz, P., & Sorter, M. T. (2016). A Controlled Trial Using Natural Language Processing to Examine the Language of Suicidal Adolescents in the Emergency Department. *Suicide and Life-Threatening Behavior, 46*(2).
- Rummery, G. A., & Niranjani, M. (1994). *On-line q-learning using connectionist systems (tr 166). Technical Report.* Engineering Department, Cambridge University.
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development, 3*(3), 210–229.
- Saxe, G. N., Ma, S., Ren, J., & Aliferis, C. (2017). Machine learning methods to predict child posttraumatic stress: A proof of concept study. *BMC Psychiatry, 17*, 223.
- Schmaal, L., Marquand, A. F., Rhebergen, D., Van Tol, M. J., Ruhé, H. G., Van Der Wee, N. J. A., ... Penninx, B. W. J. H. (2015). Predicting the Naturalistic Course of Major Depressive Disorder Using Clinical and Multimodal Neuroimaging Information: A Multivariate Pattern Recognition Study. *Biological Psychiatry, 78*(4), 278–286.
- Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine, 49*(9), 1426–1448.
- Skåtun, K. C., Kaufmann, T., Doan, N. T., Alnæs, D., Córdova-Palomera, A., Jönsson, E. G., ... Westlye, L. T. (2017). Consistent Functional Connectivity Alterations in Schizophrenia Spectrum Disorder: A Multisite

- Study. *Schizophrenia Bulletin*, 43(4), 914–924.
- Slater, M., Neyret, S., Johnston, T., Iruretagoyena, G., Crespo, M. Á. de la C., Alabèrnia-Segura, M., … Feixas, G. (2019). An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Scientific Reports*, 9, 10903.
- Strous, R. D., Koppel, M., Fine, J., Nachliel, S., Shaked, G., & Zivotofsky, A. Z. (2009). Automated characterization and identification of schizophrenia in writing. *Journal of Nervous and Mental Disease*, 197(8), 585–588.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. The MIT Press (2nd ed.).  
(サットン, R. S., & バート, A. G. 三上 貞芳・皆川 雅章訳(2000). 強化学習 森北出版)
- Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.
- Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: A retrospective. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 73(3), 273–282.
- Tron, T., Peled, A., Grinsphoon, A., & Weinshall, D. (2016). Automated facial expressions analysis in schizophrenia: A continuous dynamic approach. In S. Serino, A. Matic, D. Giakoumis, G. Lopez, & P. Cipersso (Eds.), *Pervasive Computing Paradigms for Mental Health. MindCare 2015. Communications in Computer and Information Science* (vol 604, pp. 72–81). Springer, Cham.
- Tymofiyeva, O., Yuan, J. P., Huang, C. Y., Connolly, C. G., Henje Blom, E., Xu, D., & Yang, T. T. (2019). Application of machine learning to structural connectome to predict symptom reduction in depressed adolescents with cognitive behavioral therapy (CBT). *NeuroImage: Clinical*, 23, 101914.
- Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, 8, 279–292. <https://doi.org/10.1007/bf00992698>
- Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., … Goldenberg, A. (2019). Do no harm: a roadmap for responsible machine learning for health care. *Nature Medicine*, 25, 1337–1340.
- Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for

- optimization. *IEEE Transactions on Evolutionary Computation*, 1(1), 67–82.
- Yamamoto, T., Uchiumi, C., Suzuki, N., Yoshimoto, J., & Murillo-Rodriguez, E. (2020). The psychological impact of “mild lockdown” in Japan during the COVID-19 pandemic: a nationwide survey under a declared state of emergency. *International Journal of Environmental Research and Public Health*, 17(24), 9382.
- Yamamoto, T., & Yoshimoto, J. (2018). Artificial Intelligence-Based Approaches for Health Behavior Change. In *Association for Behavioral and Cognitive Therapies 52nd Annual Convention*. Washington, DC.
- Yamamoto, T., Yoshimoto, J., Murillo-rodriguez, E., & Machado, S. (2019). Prediction of Daily Happiness Using Supervised Learning of Multimodal Lifelog Data. *Revista Psicologia e Saúde*, 11(2), 145–152.
- Yamashita, A., Sakai, Y., Yamada, T., Yahata, N., Kunimatsu, A., Okada, N., … Imamizu, H. (2020). Generalizable brain network markers of major depressive disorder across multiple imaging sites. *PLoS Biology*, 18(12), e3000966.
- Yamashita, A., Yahata, N., Itahashi, T., Lisi, G., Yamada, T., Ichikawa, N., … Imamizu, H. (2019). Harmonization of resting-state functional MRI data across multiple imaging sites via the separation of site differences into sampling bias and measurement bias. *PLoS Biology*, 17(4), e3000042.
- Yamashita, O., Sato, M. A., Yoshioka, T., Tong, F., & Kamitani, Y. (2008). Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *NeuroImage*, 42(4), 1414–1429.
- Yang, S., Zhou, P., Duan, K., Hossain, M. S., & Alhamid, M. F. (2018). emHealth: Towards Emotion Health Through Depression Prediction and Intelligent Health Recommender System. *Mobile Networks and Applications*, 23(2), 216–226.
- Zhu, F., Panwar, B., Dodge, H. H., Li, H., Hampstead, B. M., Albin, R. L., … Guan, Y. (2016). COMPASS: A computational model to predict changes in MMSE scores 24-months after initial assessment of Alzheimer’s disease. *Scientific Reports*, 6, 34567.
- 山本哲也 (2018). 神経認知療法－神経科学は認知行動療法を増強する－. 認知療法研究, 11(1), 13–22.

- 山本哲也・吉本潤一郎 (2019). 個人に最適化されたうつ病再発兆候の早期発見技術の開発ー心理・社会・生物学的データに対する機械学習法の適用ー. 第16回うつ病学会総会, 173.
- 山本哲也・竹林由武 (2020). ウェブベースドな支援. In 竹林由武 & 前田正治 (Eds.), 遠隔心理支援ー物理的距離を超えてケアを継続するヒント (pp. 126-145). 誠信書房.
- 山下裕子・山本哲也 (2021). 親密他者の視点取得を活用したVRセルフカウンセリングの効果. 日本心理学会第85回大会発表論文集, 印刷中.

# **Clinical Applications of Machine Learning: A Novel Approach to Mental Health Care Driven by Digital Innovation**

Tetsuya YAMAMOTO<sup>1</sup> Junichiro YOSHIMOTO<sup>2</sup>

<sup>1</sup> Graduate School of Technology, Industrial and Social Sciences, Tokushima University

<sup>2</sup> Graduate School of Science and Technology, Nara Institute of Science and Technology

## **Abstract**

Machine learning is a branch of artificial intelligence technology that has received considerable attention in recent years. It is a computational strategy to discover the regularities inherent in multidimensional data sets, allowing us to build predictive models focused on individual states. Therefore, it may help increase the efficiency and sophistication of assessment and aid the selection of optimal intervention methods in clinical practice, including cognitive behavioral therapy. In this paper, we first review the framework of the machine learning approach, its differences from statistics, and its features. Subsequently, we summarize the main research topics where machine learning approaches have been applied in the field of mental health and introduce some examples of their applications that may contribute to research in clinical psychology and cognitive behavioral therapy. Finally, the limitations of the machine learning approach are discussed, as well as its potential for future applications.

**Keywords:** machine learning, artificial intelligence, cognitive behavioral therapy, mental health, clinical psychology