

機械学習技術の紹介とその鉄鋼業への応用事例

Introduction of Machine Learning Techniques and Application to Steel Industry

上杉 徳照*

UESUGI Tokuteru

Synopsis: In the steel industry, machine learning techniques, including deep learning, are increasingly being employed to automate various processes, leading to the replacement of skilled workers with advanced computer systems. These efforts are distinct from conventional automation approaches, as they rely on implementing machine learning algorithms. Machine learning facilitates optimizing process control and inspection procedures within the steel industry by streamlining and automating key processes. However, the successful application of machine learning necessitates carefully selecting suitable technology, mainly when data availability is limited. This paper briefly introduces machine learning concepts, deliberately excluding detailed explanations of individual algorithms underlying mathematics and principles. Instead, we focus on the critical technical components that enhance accuracy utilizing the steel industry and present relevant application examples. The ongoing advancement of machine learning and associated information technologies is anticipated to boost efficiency in the steel industry further. Consequently, the judicious selection and deployment of appropriate technologies will contribute to sustained improvements in industrial competitiveness.

Keywords: machine learning; steel industry; artificial intelligence (AI); digital transformation (DX); deep learning; neural network; convolutional neural network (CNN); gradient boosting decision tree; explainable artificial intelligence (XAI); feature engineering; regularization; data augmentation; transfer learning; Mg yield rate; fractography

1. はじめに

鉄鋼業では安全性と品質を保証するために、その製造プロセスでは欠陥に対して厳しい検査が行われている。欠陥を含む不良品は検査により除外されるので、製品の歩留まりを低下させ、生産コストを上昇させる。欠陥の主な原因は、プロセス条件の不十分な管理によるものである。そのため、鉄鋼業ではプロセスを制御あるいは最適化するための工程が必要不可欠である。ただし、古くからプロセス制御には計算機が導入され¹⁾、自動化、無人化への取り組みがなされてきた。また、検査工程そのものも、生産コストを上昇させる要因の一つであるので、画像検査を中心に自動化の取り組みがなされてきた²⁾。

しかしながら、プロセス制御と検査工程が全て自動化され、無人化されることで、実質的に工程省略が達成されたわけではない。原材料の成分などを含めると、全く同じ条件の下で製造することは不可能なので、物理法則に基づいた原理原則に従って制御しても、同じ品質の製品を製造す

ることはできない。そのため、自動化が達成できたのは、物理モデルとして記述でき、自動化し易い範囲に留まっている。原材料の成分、プロセスの温度、装置のウォームアップや摩耗状態などの不確かさがある生産管理の工程は、今でも熟練作業者が担っている。同様に検査工程の自動化においても、人間が判断したとしても良品と不良品との閾値が明確でない場合は、欠陥の要件を定義して記述することができないので、自動検査は困難である。そのため、現在でも良品と不良品との閾値が明確でない検査工程では熟練検査員の目視に頼っている。

人に頼った生産管理や検査工程は熟練者と非熟練者のばらつきが大きく、熟練者による属人的な暗黙知は伝承が難しい。このような背景から、計算機により熟練者を代替する自動化への取り組みは継続的に行われてきた。これまで、熟練者の大量定年退職が問題視された時には、熟練者の暗黙知を何らかの方法で形式知化し、自動化する試みが行われた³⁾。ただし、実際には熟練者は高度な情報処理を行っており、その複雑な知的判断の全てを形式知化するこ

* 大阪公立大学大学院 情報学研究科 学際情報学専攻 准教授

とが難しかった。

人工知能 (Artificial Intelligence: AI) 分野では同様の問題が、古くからエキスパートシステムの問題点として知られている。第2次AIブームと呼ばれ、1980年代に隆盛したエキスパートシステムでは、IF-THENルールの形式で膨大な知識表現を登録する必要がある。そのため、人間ではニュアンスで表現される熟練者の知識表現も、計算機には明示的に表現する必要があるため、暗黙知を表現できないという課題を解決できず、第2次AIブームは下火になった⁴⁾。

暗黙知を形式知化することの困難さは変わらないまま、計算機技術は指数関数的に発展した。その結果、IoT (Internet of Things) と呼ばれる技術に象徴されるように、様々な製造プロセスのデータを収集するためのコストは大きく低下し、ビッグデータと呼ばれる技術のように、収集した大容量のデータを蓄積するためのコストも大きく低下した。また、大量の並列処理を可能にするGPU (Graphics Processing Unit) の性能が向上した。これらビッグデータとGPUの性能向上は、人工知能分野では、深層学習の登場をもたらし、深層学習は現在も続く第3次AIブームを牽引する技術となった⁵⁾。

以上の歴史的な背景から、鉄鋼業でも深層学習を始めとした機械学習により、計算機により熟練者を代替する自動化への取り組みが再び見直されている。これまでの、自動化との取り組みとの違いは機械学習を用いている点である。機械学習は適用できる領域が広く、既に一般向けから専門家向けまでの数多くの解説書が出版され、金属材料分野でも解説記事が存在する⁶⁻⁸⁾。本寄稿では機械学習については、入門的知識を紹介するに留め、機械学習の各アルゴリズムの数理や原理の解説は省いている。その代わりに、鉄鋼業への利用において重要となる精度向上への技術を紹介し、その適用事例について述べる。

2. 機械学習とは

2.1 ルールベースの人工知能との違い

機械学習を理解するには、ルールベースの人工知能との対比が役立つ。ルールベースでは、熟練者の知識や物理モデルを基盤に、手作業で策定されたルールが適用される。これは、IF-THENルールによって構成される。例として、「もし、温度が1500℃を超えたら、原料を1kg追加する。」というルールや、「もし、輝度値が95以下かつ面積が10ピクセル以上であれば欠陥と判断する。」といった規則を作成することが挙げられる。特定の問題に対して適切なルールを構築することができれば、ルールベースのアプローチは理解しやすく、説明が容易であり、かつ高い精度を達成することが可能である。しかし、ルールベースのアプローチには限界が存在する。問題が複雑であるほど、適切なルールを作成するのが困難になる。先に述べたように熟練

者は高度な情報処理を行っており、その複雑な知的判断をルールとして作成することは困難である。同じく良品と不良品との閾値が複雑で明確でない検査工程では、ルールを作成することは困難である。つまり、ルールベースでは熟練者の複雑で曖昧な知識表現を明示的なルールに表現する必要があるが、熟練者の暗黙知を形式知化することは、実際には困難である。その結果、歴史的に見れば、ルールベースは熟練者の知的判断の一部しか代替できなかった。

一方で、機械学習はルールベースとは全く異なるアプローチを採用している。機械学習では、アルゴリズムが大量のデータから自動的に特徴やパターンを学習し、その知識を用いて新たなデータに対して予測や判断を行う。そのため、結果的には熟練者の暗黙知を形式知化することなく、熟練者と同等に、複雑な問題や曖昧な状況に対処することができる。

アルゴリズムとは、一連の計算的な手続きであり、これに従うことで誰でもモデルを構築することができる。モデルとは、学習対象を数学的に形式化したもので、通常は数式で表現される。この数式によって、データの構造や関係性が表現され、予測や分類などのタスクが実行できる。

IF-THENルールでは、人間が直接的に操作できる明確なルールで構成されていたが、機械学習モデルでは、その性質上、人間が直接的に操作することが難しい。機械学習モデルは、大量のデータとアルゴリズムを用いて、自動的にパラメータを調整している。例えば、VGG16は、機械学習の中でも、深層学習 (Deep Learning) の一種である畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) を利用した画像認識モデルである⁹⁾。VGG16では、およそ1.4億のパラメータが含まれているが、人間が1.4億ものパラメータを直感的に理解して操作することは困難である。

VGG16のような深層学習に限らず多くの機械学習モデルでは、モデルの内部構造や意思決定プロセスを人間が直接操作や理解することは難しい。そのため、多くの機械学習モデルは、しばしばブラックボックスと呼ばれることがある。しかしながら、ブラックボックスモデルであっても、その内部は前述のように、通常は数式で記述されている。そのため、モデルがどのように機能しているかを解析することは可能である。ただし、非線形な関数や、無数のパラメータが関与することにより、人間が直感的に理解することが困難なだけである。そのため、説明可能な人工知能 (Explainable AI: XAI) という分野では、モデルの説明可能性や解釈性を向上させるための手法が開発されている。

2.2 機械学習の枠組み

機械学習は、データから有用な知識を引き出すための技術であり、その目的は予測、発見、説明の3つに大別できる。ただし、鉄鋼業においては、予測が主に利用されるこ

とが多く、製造プロセスの最適化と検査工程でも、機械学習による予測の利用が考えられる。予測が目的の場合では機械学習モデルは、過去のデータをもとに、結果を予測することができる。身近な例で言えば、天気予報、顧客の購買行動の予測などである。これには、機械学習の回帰や分類といった教師あり学習が用いられる。

熟練者によるプロセス制御と検査工程も、機械学習モデルで予測することができる。プロセス制御では、熟練者は長年の経験を通じて、さまざまなプロセス条件と、それに対応する製品品質のパターンを学習している。その結果、熟練者は、品質が規定値に収まるように、操作可能なプロセス因子の値を適切に予測することができる。検査工程では、熟練者は長年の経験を通じて、様々な性状の製品を観察しており、欠陥を素早く検出・分類する能力を有している。熟練者は、微妙な違いやパターンを認識しており、良品と不良品との閾値が明確でない場合でも、分類を行うことができる。これらの熟練者の複雑な知的判断は十分な教師データを用意できれば、教師あり学習の枠組みで解決できる。教師データとは、正解ラベルが付与されたデータのことであり、教師あり学習では教師データを用いてモデルの学習を行う。

発見が目的の場合では、機械学習では、データから潜在的なパターンや関係性を探索することが可能である。発見を目的とする場合には、主にクラスタリングや次元削減といった教師なし学習が適用される。一般的な用途では例えば、マーケットバスケット分析、異常検知などである。マーケットバスケット分析では「おむつを買った人はビールを買う傾向がある」といった顧客が同時に購入する商品の組み合わせを分析することができる。異常検知はデータセット内の異常値や外れ値を検出する手法である。鉄鋼業では、例えば、設備の故障予測に異常検知が利用される。故障予測においては、センサーデータやメンテナンス履歴データなどの時系列データを利用し、未知の異常パターンを検出（発見）することが求められる¹⁰⁾。

説明を目的とする機械学習では、データに基づいて問題や現象に関する因果関係やメカニズムを明らかにすることが目指される。この目的を達成するために、回帰分析や決定木といった比較的単純な手法が適用されることが多い。これらの手法は、モデルの構造が単純であり、その結果、人間が理解しやすい形で相関関係を示すことができる。ただし、機械学習の多くの応用では、説明よりも予測が主な目的となる。一方で、統計学は機械学習と区別が難しい分野であるが、予測よりも説明が主な目的である点で機械学習とは異なる。統計学では、データの背後にある確率分布やその構造を明らかにし、因果関係の推定やメカニズムを理解することが求められる。

機械学習では深層学習のような、非常に複雑なモデルであっても、XAIにより、モデルの説明が可能である。ただ

し、XAIの手法はモデルの予測結果と入力特徴量の相関関係に着目しているだけである。モデルが予測を行う際に、どの特徴量が重要であったかを示すことができるが、それが因果関係を示しているわけではない。因果関係を明らかにするには、自然科学では異なる状況下での実験を通じて、特徴量と実験結果の関係性を検証することが一般的である。しかし、因果関係の解明には、適切な仮説と、それを実証する精緻なデータを必要とする。また、因果関係を明らかにすることができたとしても、それがモデルの予測性能に寄与するわけでもない。機械学習で相関関係を説明できても、因果関係を説明するには注意が必要である。

機械学習の用途を大別したが、それぞれの目的に応じて、異なるアルゴリズムが適用される。機械学習アルゴリズムの観点からは、機械学習は教師あり学習、教師なし学習、強化学習の3つに大別することができる。教師あり学習は、入力データとそれに対応するラベルと呼ばれる出力データのペアが与えられた学習データセットを用いて、新しい入力データに対する出力を予測するモデルを構築する。教師なし学習は、入力データのみが与えられた学習データを用いて、データの潜在的な構造やパターンを発見する。強化学習は、エージェントが環境と相互作用し、報酬を最大化するような行動を学習するアルゴリズムである。強化学習はタスク達成や意思決定の問題に適用されることが多く、その用途としては、特定のケースに対応するものとなる。また、教師なし学習の用途も教師あり学習と比べると限定的である。そのため、本寄稿では、予測を目的とした教師あり学習についてのみ言及する。

2.3 教師あり学習

機械学習では、教師の有無は正解ラベルの有無を指し、教師あり学習は正解ラベルの種別によって分類と回帰に大別される。正解ラベルが離散値であれば、データが属するクラスを予測する分類問題となり、正解ラベルが連続値であれば値を予測する回帰分析となる。図1に分類問題と回帰分析の概略を示す。例えば、製品の画像データから、その製品が良品か不良品かを予測するタスクは分類問題である。一方、化学組成やプロセス条件のデータから素形材の歩留まりや機械的特性を予測するタスクは回帰問題とな

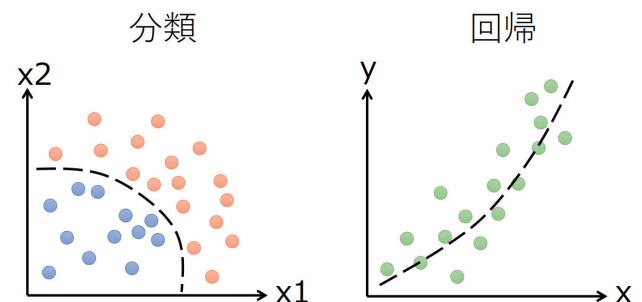


図1 分類と回帰

る。これらの問題に対して、教師あり学習アルゴリズムを適用することで、新しいデータに対して予測を行うモデルを構築することが可能である。

教師あり学習の問題の枠組みをより詳しく見ておくと、データ x の入力に対して、出力となるラベルを y とする。この時、 x は説明変数、特徴量、記述子、入力変数などと呼ばれる。また、 y は被説明変数、従属変数、目的変数、出力変数などと呼ばれる。分類問題においては、ラベルは0と1の離散値で表されることが一般的であり、これは2クラス分類問題に適用される。一方、多クラス分類問題では、ラベルはone-hotベクトルという形式で与えられる。one-hotベクトルとは、全ての要素が0であるベクトルのうち、対応するクラスのインデックスのみが1であるようなベクトルである。また、回帰問題であれば、 y は連続値となる。

教師あり学習の事例は x と y の組で与えられ、データセットは事例の集合で与えられる。教師あり学習の目標は、 x を入力とするモデル $f(x)$ の出力値が y に近づくように学習を行うことである。ここで、 $f(x)$ と y の近さ・遠さを測る指標として損失関数を用いる。そのため、教師あり学習とは損失関数の最小化・最適化を行うプロセスとも言える。

回帰問題では、平均二乗誤差 (Mean Squared Error: MSE) が一般的な損失関数として用いられるが、平均絶対誤差 (Mean Absolute Error: MAE) や分位点損失 (Quantile Loss)¹¹⁾ が選択されることがある。分類問題では、交差エントロピー誤差 (Cross-Entropy Loss) がよく用いられ、二クラス分類の場合は二値交差エントロピー誤差 (Binary Cross-Entropy Loss) が適用される。クラスの不均衡がある場合、重み付きクロスエントロピー誤差 (Weighted Cross-Entropy Loss)¹²⁾ などが選択されることもある。損失関数は問題設定に応じて適切に選択する必要がある。

損失関数は、訓練データを用いてモデルの最適化を行う際に使用されるが、性能評価では、検証データやテストデータに基づいて評価関数が適用される。回帰では、損失関数としてMSEが用いられる一方で、評価関数としては決定係数 (Coefficient of Determination: R^2) や二乗平均平方根誤差 (Root Mean Squared Error: RMSE) が選ばれることが多い。分類問題では、評価関数には正解率 (Accuracy) の他に、適合率 (Precision)、再現率 (Recall)、適合率と再現率の双方が考慮されるF値 (F-measure) などが用いられる。

損失関数と評価関数が設定された後に、モデリングが行われる。教師あり学習の目標は正確な判断や予測を行うことである。つまり、教師あり学習の目的は既知データの入出力関係を単に再現するのではなく、未知の入力に対して正確な出力を行うことである。未知のデータへの対応を汎化と呼び、未知データに対する誤差を汎化誤差という。汎化誤差は損失関数の期待値として表現でき、教師あり学習

の目標は汎化誤差の最小化と言い換えられる。

汎化誤差は未知データに対する期待値であり、データ x の全ての候補を収集することは不可能であるため、汎化誤差を正確に評価することも不可能である。仮にすべてのデータを収集できた場合、そもそも予測する必要性はなくなる。このため、機械学習による予測は常に不確実性を伴う。機械学習の不確実性は、多くの事例から一般的な規則を見出そうとする帰納法の限界と言い換えることができる。帰納法の限界はヘンペルのカラスについてのパラドックスで理解することができる。このパラドックスは、どれだけ多くの事例 (黒いカラス) を観察しても、その一般化規則 (全てのカラスは黒い) は証明できないことを示している。これは、帰納法に基づく推論が完全な証明ではなく、あくまで蓋然性をもとにした推論であることを示している。

汎化誤差とは対照的に、既知データに対する誤差は訓練誤差である。訓練誤差は訓練データと呼ばれる既知データの損失関数から計算される。訓練誤差の最小化のみを追求すると、訓練誤差と汎化誤差の乖離が大きくなる過学習という現象が発生する。図2に学習不足、汎化、過学習を例示した。過学習の状態では、複雑なモデルが採用され、訓練データには良好な一致が見られるが、未知のデータに立ってテストデータに対しては一致しないことが分かる。

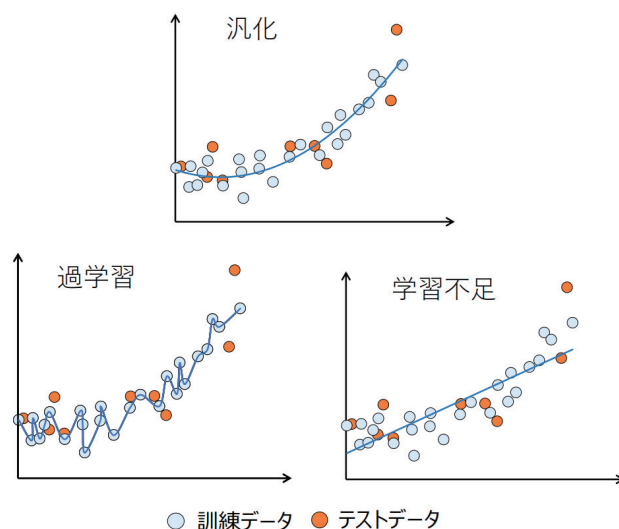


図2 回帰における汎化、過学習、学習不足

2.4 機械学習のアルゴリズム

教師あり学習のアルゴリズムは、訓練データの損失関数を最適化する手法であり、様々なアルゴリズムが存在する。以下にいくつかの代表的なアルゴリズムを紹介する。

教師あり学習は回帰と分類に分けられ、回帰は線形回帰と非線形回帰に分けられる。線形回帰にはリッジ回帰 (Ridge Regression) やラッソ回帰 (Lasso Regression)¹³⁾ が含まれ、非線形回帰には k 近傍法 (k -Nearest Neighbor)、

回帰木、ランダムフォレスト (Random Forest)¹⁴⁾、勾配ブースティング (Gradient Boosting Decision Tree)、サポートベクター回帰 (Support Vector Regression)¹⁵⁾、ニューラルネットワーク (Neural Network) などが存在する。分類も線形分類機と非線形分類機に区分され、線形分類機には単純パーセプトロン、ロジスティック回帰、線形判別分析、線形サポートベクターマシンが含まれる。一方、非線形分類機には k 近傍法、決定木、ランダムフォレスト¹⁴⁾、勾配ブースティング、サポートベクターマシン (Support Vector Machine)¹⁵⁾、ニューラルネットワークなどがある。回帰と分類で同名のアルゴリズムが見られるが、これは損失関数を変更することで両方に対応可能であるためである。

これらの機械学習のアルゴリズムごとに、モデルの複雑さは異なるが、同じアルゴリズムにおいてもハイパーパラメータによって大きく変わることがある。ハイパーパラメータとは、機械学習アルゴリズムの動作を設定するパラメータである。例として、決定木を基にした機械学習では、ハイパーパラメータとして木の深さや葉の数が挙げられる。ニューラルネットワークでは、多くのハイパーパラメータがあるが、ハイパーパラメータには、層の数、ノード (ニューロン) の数、活性化関数、学習率、エポック数などがある。

サポートベクターマシン¹⁵⁾は、データを分離する境界の中で最も離れた境界を決定する手法であり、カーネル関数を用いることで非線形に拡張される。データを分離する境界に最も近いデータはサポートベクトルと呼ばれ、境界とサポートベクトル間の距離はマージンと称される。多次元空間で境界を決定するため、この境界は超平面であり、サポートベクターマシンでは最大マージン超平面を求める。回帰への応用はサポートベクター回帰とも呼ばれる。サポートベクターマシンやサポートベクター回帰は、ハイパーパラメータが少なく扱いやすく、優れた汎化性能を持つ手法の1つである。

決定木は、木構造を利用した分類や回帰手法であり、回帰を行う際には回帰木とも呼ばれる。決定木は、根から段階的に枝分かれする木構造に従って条件が整理される。ランダムフォレスト¹⁴⁾と勾配ブースティング¹⁶⁻¹⁸⁾は、決定木を基にした手法で、どちらも決定木を用いたアンサンブル学習である。アンサンブル学習とは、単独では精度が低いモデルを多数利用することで高精度を実現する手法群であり、バギングとブースティングが代表的である。ランダムフォレストはバギングを用い、勾配ブースティングはブースティングを採用している。ランダムフォレストはよく使用される手法であるが、機械学習コンペティションであるKaggle¹⁹⁾の上位ランカーにはXGBoost¹⁶⁾、LightGBM¹⁷⁾、CatBoost¹⁸⁾と呼ばれる勾配ブースティングが頻繁に用いられている。

ニューラルネットワークは、微分可能な線形変換と非線形変換を連結した計算グラフである。計算グラフは、図3で示すようなノードとエッジで構成され、ノードでは活性化関数による非線形変換が行われ、エッジでは線形変換が行われる。ノードの集合を層と呼び、多層化したニューラルネットワークが構築される。

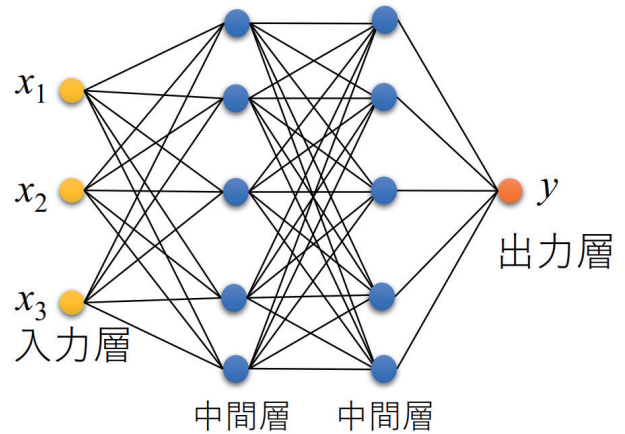


図3 ニューラルネットワークの構成

層が多いニューラルネットワークを使用した機械学習手法は、特に深層学習と呼ばれる。深層学習は、大規模なデータセットに対して高い表現力を持つため、画像認識、自然言語処理、音声認識など様々な分野で優れた性能を発揮している。調整可能なハイパーパラメータは膨大で、その構造 (アーキテクチャ) は自由に設計できる。深層学習の発展により、画像処理でよく用いられるCNN以外にも、時系列データでよく用いられる再帰型ニューラルネットワーク (Recurrent Neural Network: RNN)、データ生成が可能な敵対的生成ネットワーク (Generative Adversarial Network: GAN) など、多様なアーキテクチャが提案されている。これらのアーキテクチャは、特定のタスクやデータ形式に適應するために設計されており、深層学習の応用範囲を広げている。

深層学習が注目され始めたのは、2010年代初めである。この時期に、CNNが画像認識で顕著な性能向上を示し、従来の機械学習手法を凌駕する結果が報告された。特に2012年のImageNetコンペティションで、AlexNet²⁰⁾と呼ばれるCNNが他の手法を圧倒的に上回る成績を収めたことが、深層学習のブームのきっかけとなった。なお、ImageNet²¹⁾とは1400万枚以上の教師ラベル付き画像データベースであり、コンペティションではImageNetに対して画像分類の精度が競われた。これ以降、深層学習技術の発展と応用が急速に進み、様々なアーキテクチャが開発されることになった。AlexNetでは5層の畳み込み層と3層の全結合層で構成されていたが、VGG16⁹⁾では13層の畳み込み層と3層の全結合層で構成されていた。その後も深層化が進み、例えば2017年に発表された

InceptionResNetV2²²⁾では572層もの深層化を行っている。

現在では深層学習は、画像、自然言語、音声といった大量の非構造化データを扱うドメインで大きな成功を収めている。なお、機械学習におけるドメインとは、特定の問題領域、もしくは特定のデータセットに関連する背景情報を指す。一方で、構造化データである表形式データにおいては、深層学習はCatBoostなどの勾配ブースティングの性能を上回ることができず、データ依存性が大きいと指摘されている²³⁾。鉄鋼業においては、肉眼や顕微鏡を用いた検査工程を代替する目的での深層学習の利用が考えられる。これは、深層学習が非構造化データである画像を扱うのに適した手法であるためである。しかし、プロセス制御では、時系列の表形式データが主に用いられる。このような構造化データに対しては、深層学習は万能な解決策とは言えず、他の手法が適切に選択されることが求められる。

3. 機械学習の技術

既に述べたように、教師あり学習の目標は汎化誤差の最小化である。そのため、機械学習における技術は、ほぼ、汎化性能を向上させることに注力される。これには、適切なモデル選択、ハイパーパラメータチューニング、特徴量エンジニアリング (Feature Engineering)、データの前処理、正則化 (Regularization)、データ拡張 (Data Augmentation)、転移学習 (Transfer Learning) の導入など、さまざまなアプローチが取り入れられる。これらの手法を適切に組み合わせることで、学習データだけでなく未知のデータに対しても高い予測性能を持つモデルを構築できる。

3.1 データ量

率直に言えば、汎化性能を向上させる目的を達成するための最も確実で、最も有効な手段はデータを増やすことである。データ量が増えることで、学習モデルはより多様なデータに適応し、未知のデータに対する予測性能が向上することが期待される。

図4は分類問題における学習データ量が境界決定に与え

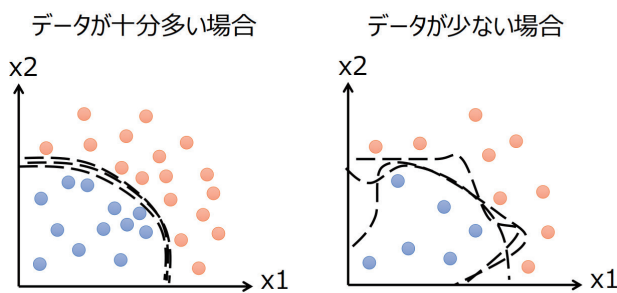


図4 データ量が分類境界に与える影響

る影響を示している。データが十分多い場合、複雑なモデルを適用することで、境界をほぼ一意に決定できることが期待される。これは、多くのデータによって、モデルがより正確に確率分布を捉えることが可能となり、汎化性能が向上するためである。一方、データが少ない場合、学習データに対して高性能な分類境界が多数存在し、それらが無数に獲得可能である。この状況では、どの境界が真に汎化性能に優れているかを判断することが困難となる。従って、データ量が境界決定において重要な役割を果たし、十分なデータが揃っていることが、高い汎化性能を実現するための前提であると言える。

データ量の適切性はタスクやモデルの複雑さに依存し、一概に定めることは困難である。しかし、いくつかの経験則が存在する。複雑なモデルは多くのパラメータを持ち、過学習を防ぐために多くのデータが要求されるとの経験則が古くは指摘されている²⁴⁾。また、タスクの難易度が高いほどデータ量の必要性が増すことも理解できる。例えば、画像認識や自然言語処理のような非構造化された高次元データでは、表形式データと比較して大量のデータが必要である。さらに、データ量だけでなく品質も重要であり、ノイズの多いデータや外れ値を含むデータはモデル性能に悪影響を及ぼすことがある。従って、データの前処理やクリーニングも重要であると言える。

適切なデータ量を把握する方法として、増分的にデータを追加しながら性能の向上を評価するアプローチが有効である。図5は、深層学習におけるデータ量と汎化誤差の関係を示している²⁵⁾。まず、小データ領域では、モデルは十分な汎化性能を発揮できず、深層学習モデル以外の選択肢を検討すべきである。データが増加すると、画像認識や自然言語処理などのタスクに関係なく、べき乗則あるいはスケールリング則と呼ばれる領域が観察される。データ量と汎化誤差のべき指数は、両対数スケールの傾きから算出でき、タスクの難しさの指標とも言える。ただし、べき指数はドメインやデータ分布に依存しており、理論的に予測す

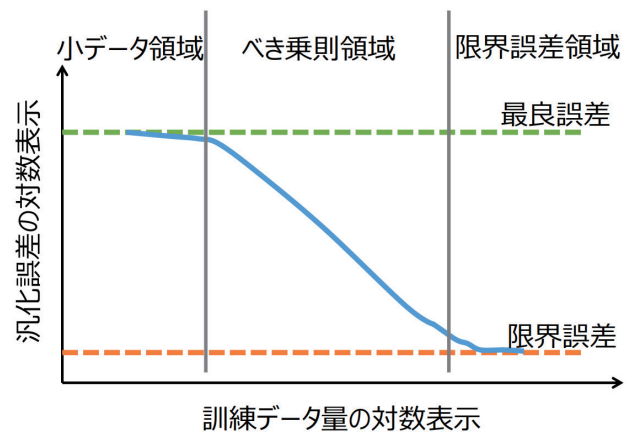


図5 深層学習におけるデータ量と汎化誤差の経験的な関係 (文献²⁵⁾より改変引用)。

ることは容易ではない。そのため、ある程度のデータを用意し、べき指数を算出してから、べき指数に基づいて目標とする汎化性能に必要なデータ量を見積もるという運用が考えられる。べき乗則の領域から、さらにデータ量を増やすと、モデルが改善できない限界誤差が存在することが予測されている²⁵⁾。この下限の原因としては、例えば学習用データセットの誤ったラベル付けが挙げられ、これによって不可避な誤差が発生する可能性がある。

べき乗則が成立する限りでは、本寄稿で紹介する汎化性能向上のための技術は、ある程度限定的な効果しか持たない。例えば、OpenAIが開発したChatGPT (Chat Generative Pre-trained Transformer) は、生成系大規模言語モデルであり、自然言語処理モデルのTransformer²⁶⁾が基盤である。Transformerにおいてもデータ量と汎化性能との間でべき乗則が成り立つと報告されている²⁷⁾。ただし、データ量とモデルのパラメータ数は相補的な関係にあり、一方が固定されると、もう一方を増加させても汎化性能の向上は鈍化する。そのため、データ量とパラメータ数の両方を増やす必要がある。さらに、学習に用いるGPUの数や計算時間といった計算資源と汎化性能との間でもべき乗則が成り立つことが示されている。そのため、データ量、モデルサイズ、計算資源を拡大すれば性能が向上し続け、投資効果が予測可能である。この予測を基に開発されたGPT-3 (Generative Pre-trained Transformer 3) は、1750億個のパラメータを持つ大規模なモデルであり²⁸⁾、1.4億パラメータを有するVGG16よりも格段に大きい。その後、GPT-3はGPT-3.5に発展し、教師あり学習と強化学習の両方の手法で転移学習することで、ChatGPTが生まれ出された。GPT-3.5は更に、学習データを増やすことで、GPT-4へと大規模なモデルに発展している。

しかし、限られたデータ量の中で高い汎化性能を達成するためには、検証データの分割や交差検証などの評価を通じて、様々な機械学習技術を活用して汎化性能の向上に努める必要がある。鉄鋼業における現実の課題であるプロセス制御や検査工程では、データが十分に利用できる状況が稀であるため、大規模モデルよりも、様々な汎化性能向上の技術が重要となる。即ち、データ量が限られた状況下でも適切に機能するモデルを開発し、実用的な解決策を提供することが求められる。次節から、そのための、具体的な機械学習の技術を紹介する。

3.2 データの分割

教師あり学習の目標は汎化誤差の最小化であり、未知データに対する期待値として捉えられる。モデルが複雑であれば訓練誤差は小さくなるが、過学習の原因となるため、訓練データだけでモデルの複雑さを決定することは困難である。そこで、汎化性能を評価し、モデルの複雑さを調整するために、検証データを利用する。一般的には、既

知データを訓練データ、検証データ、テストデータの3種類に分割する。

訓練データでモデルを学習させ、検証データで汎化性能を評価することで、モデル選択やハイパーパラメータチューニング、特徴量エンジニアリング、データの前処理、正則化などを行う。最終的に完成したモデルは、未知のデータに見立てたテストデータを用いて評価を行うのが一般的な手法である。検証データを用いてモデルの複雑さの調整を繰り返すと、ある意味では検証データに対して過学習することになる。この問題を克服するために、最終評価用のテストデータを訓練データおよび検証データとは別に用意する。テストデータを用いて最終的なモデルの汎化性能を評価することで、未知のデータに対する性能の妥当性が保証される。このアプローチにより、過学習の影響を最小限に抑えつつ、モデルの汎化性能を適切に評価することが可能となる。

交差検証法とホールドアウト法は、機械学習モデルの性能評価において一般的な検証方法である。ホールドアウト法は、データセットを訓練データと検証データに一度だけ分割し、訓練データでモデルを学習させ、検証データでモデルの性能を評価する方法である。シンプルで実装が容易であるが、データの分割方法によっては性能評価が不安定になることがある。

交差検証法の概要を図6に示す。交差検証法ではデータセットを複数の部分集合にランダムに分割し、各部分集合を検証データとして使い、残りの部分集合を訓練データとして用いる。このプロセスを全ての部分集合に対して繰り返し行い、得られた性能評価の平均値を最終的な評価指標とする。交差検証法は、データセット全体を検証データとして利用するため、性能評価が安定し、モデルの汎化性能をより正確に評価できる。ただし、計算コストが高いという欠点がある。

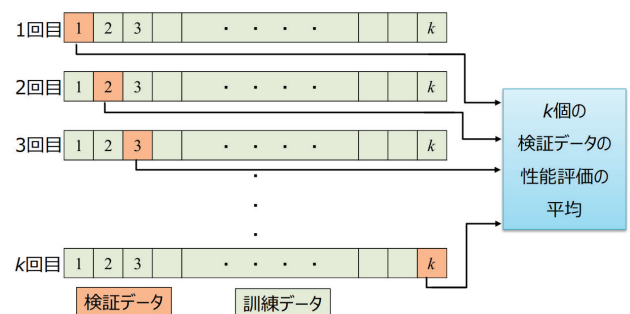


図6 交差検証法

ホールドアウト法はモデルの学習に時間がかかる場合に利用されるが、深層学習ではホールドアウト法が用いられ、それ以外の学習時間の短い機械学習手法では交差検証法が用いられることが多い。交差検証を用いてモデルの複雑さの調整が完了した場合、テストデータを除いたデータ

で再学習を行い、最終的な評価をテストデータで実施する。ホールドアウト法においても、訓練データのみ、あるいはテストデータを除いたデータ（訓練データと検証データの合計）で再学習を行い、テストデータを用いて最終評価を実施する。このテストデータを用いた最終評価のためのデータ分割方法も、ホールドアウト法と称されることがある。この手順により、モデルの未知データに対する性能を適切に評価することが可能である。

3.3 モデル選択

十分なデータが用意できれば、べき乗則に従いパラメータ数の多い深層学習モデルを学習させることが可能である。しかし、現実的にはデータの無尽蔵な供給は期待できない。結果的には、モデル選択に関して、ホールドアウト法や交差検証を用いて汎化性能の良いモデルを試行錯誤で選択することが必要となる。

この問題はノーフリーランチ定理として、機械学習では古くから知られている²⁹⁾。機械学習アルゴリズムのノーフリーランチ定理とは、ある課題に対してアルゴリズムAがアルゴリズムBより汎化性能が優れていたとしても、別の課題でアルゴリズムBがAを上回るケースが必ず存在することを示すものである。従って、どの手法も他の手法に対して絶対的な優位性を持たないため、課題やデータ量に応じたアルゴリズムの選択や開発が不可欠となる。

機械学習研究者は、常に新たなアルゴリズムの開発を目的としているが、これまで無数のアルゴリズムが開発されているにも関わらず、万能な機械学習アルゴリズムが存在しない現状が、ノーフリーランチ定理の正当性を示唆している。ノーフリーランチ定理により、課題やデータに応じて最適なアルゴリズムを選択し続けることが重要であることが示される。

主要な機械学習アルゴリズムはプログラミング言語Pythonで利用可能であり、Pythonを使うことで複数の機械学習アルゴリズムの検証コストは低くなる。標準的に使用される機械学習ライブラリはオープンソースのscikit-learn³⁰⁾であり、多くの機械学習アルゴリズムの検証はこれだけで完結する。CatBoostのような最新の手法はscikit-learnに含まれていないが、scikit-learn互換のライブラリがPythonで利用可能である。深層学習においては、Tensorflow³¹⁾、Keras³²⁾、PyTorch³³⁾がよく使用される。KerasはTensorflowのモジュールとして利用でき、Tensorflow/KerasはGoogleにより、PyTorchはFacebookにより開発されているが、いずれもPythonで書かれたオープンソースである。

3.4 ハイパーパラメータチューニング

ハイパーパラメータは機械学習アルゴリズムの挙動を設定するパラメータであり、その値に応じてモデルの汎化性

能が変化することがある。自動化されたハイパーパラメータチューニングも存在するが、一般的にはモデルの設定者が手動で調整を行う。

ハイパーパラメータの調整には3つの主要な手法が存在する。1つ目はグリッドサーチで、事前に定められたハイパーパラメータの組み合わせを全て試す総当たりの手法である。ハイパーパラメータの数が少なく、学習時間が短いモデルに適しているが、深層学習のような訓練時間が長い場合には不向きである。2つ目はランダムサーチで、ハイパーパラメータの空間内でランダムに組み合わせを選び、評価を行う手法である。3つ目はベイズ最適化で、ハイパーパラメータの空間内で、不確かさを利用して次に探索を行うべき値を探していく最適化手法である³⁴⁾。

いずれのハイパーパラメータチューニング手法においても、ハイパーパラメータの探索空間はモデルの設定者が決定する必要がある。例えば、サポートベクターマシンでは、設定者はカーネル関数の種類を探索するかどうかや、正規化パラメータやカーネルパラメータの範囲指定が求められる。勾配ブースティングや深層学習では、ハイパーパラメータの探索空間はさらに広がり、設定者が適切な範囲を指定する必要がある。実際には、モデルの設定者がタスクごとに探索空間の範囲を経験的に決定する。しかし、ハイパーパラメータの適切な探索範囲が既知であれば、それは手動調整と大差ないため、ハイパーパラメータチューニングが完全に自動化されない原因となっている。

3.5 前処理と特徴量エンジニアリング

前処理は、機械学習モデルへの入力データを生成するために、生データに対して行われる一連の変換処理である。前処理の目的は、データの品質を向上させ、機械学習アルゴリズムの性能を最大限に引き出すことである。代表的な前処理にはデータの最大値が1、最小値が0となるように行われる正規化や、データの平均が0、分散が1となるように行われる標準化がある。ただし、標準化を含めてデータのスケールを整えること全般を正規化と呼ぶこともある。決定木ベースの機械学習を除けば、多くの機械学習手法で正規化は必須の作業である。前処理では他には欠損値や異常値を取り除いたりもする。

特徴量エンジニアリングは、データの加工を行いつつ、モデルの性能向上を目的とする手法である。特徴量エンジニアリングでは、前処理と同様にデータ加工を行うが、性能向上を重視する側面が強い。価値ある特徴量の生成や、ノイズの原因となる特徴量の除去が重要であり、データを数値的に表現する多様な方法を試行錯誤することが一般的である。さらに、ドメイン知識を活用し、データの背景情報に基づいて有益な特徴量の設計を行うこともある。なお、ドメイン知識とは、特定の問題領域に関する専門知識で、そのデータセットに関する固有の知識も含んでいる。

例えば、時系列データに対する特徴量生成は特徴量エンジニアリングの一例である。周期性や自己相関などの時系列データ固有の構造を捉える特徴量の作成が求められる。また、テキストデータや画像データに対する特徴量エンジニアリングなど、データの種類や問題設定に応じた様々な特徴量設計手法が存在する。これらの特徴量エンジニアリングを適切に実施することで、機械学習モデルの性能向上が期待される。

表形式のデータでは、特徴量の候補から有用な特徴量を選び出す場合、特徴量選択あるいは特徴量削減と呼ばれる手法が用いられる。特徴量選択は、ドメイン知識を必要とせず、フィルター法、ラッパー法、埋め込み法といったアルゴリズムによって自動的に特徴量を選択することが可能である。フィルター法は、モデルの学習を行わずに特徴量の重要度を測定し、有効な特徴量を選択する手法である。ラッパー法は、複数の特徴量を組み合わせ、モデルの学習を繰り返して評価を行い、評価指標が最も良くなる特徴量の組み合わせを探索する方法である。埋め込み法は、特徴量選択が組み込まれた機械学習アルゴリズムを使用する方法であり、代表的な手法として線形回帰のラッソ回帰¹³⁾がある。これらの特徴量選択手法を適切に用いることで、汎化性能が向上することがあり、また不要な特徴量を削減することにより、計算コストが削減される。

ラッパー法は、フィルター法とは異なり、機械学習アルゴリズムに依存して特徴選択を行うため、より厳密な選択が可能である。また、埋め込み法とは対照的に、機械学習アルゴリズムを問わず利用できる。しかし、計算コストが高いという欠点が存在する。ラッパー法にはステップワイズ法によるアプローチがあり、前進法や後退法などがその例である。前進法は特徴量の空の集合から始め、機械学習アルゴリズムの評価に基づいて、評価指標が最も向上する特徴を逐次追加する方法である。全ての特徴が選択されるか、評価指標の改善が停止するまで繰り返す。一方、後退法は全ての特徴量を含む集合から開始して、機械学習アルゴリズムの評価に基づいて、評価指標を最も悪化させる特徴を逐次削除する方法である。全ての特徴が削除されるか、評価指標の改善が停止するまで繰り返す。

特徴量エンジニアリングを含めてモデルに入力する前のデータの加工全般を前処理と呼ぶこともある。画像データに対する前処理も、機械学習モデルの性能向上に寄与する。表形式データとは異なり、画像データはピクセルの集合として表現されるため、前処理手法も異なる。主な画像データの前処理手法には、画像のリサイズ、画像の正規化、グレースケール変換、ほかし、ヒストグラム平坦化、データ拡張などがある。

画像のリサイズでは画像サイズを一定の解像度に揃えることで、機械学習モデルが一定の入力サイズを扱えるようにする。画像の正規化では各ピクセルの値を一定の範囲

(通常は0～1)にスケールすることで、学習の収束速度を向上させる。グレースケール変換ではカラー画像をグレースケールに変換し、色情報を省略し、輝度情報のみに注目することで、計算量を削減する。逆にグレースケールを疑似カラーに変換して、色情報を追加する場合もある³⁵⁾。ほかしを行うことで画像の細かなテクスチャやノイズが平滑化され、機械学習モデルが学習すべき本質的な特徴に焦点を合わせることが可能となる。ヒストグラム平坦化は、画像のコントラストを改善することで、特徴量抽出が容易となる。データ拡張は、既存の画像データに回転、反転、シフトなどの変換を加えることで、データセットの多様性を向上させ、モデルの汎化性能を向上させる手法であるが³⁶⁾、詳細は後述する。これらの前処理手法は、特定のタスクやデータセットに対してドメイン知識を活用しつつ、検証データでの評価を通じて試行錯誤的に適用されることで、画像データに対する機械学習モデルの性能向上が期待できる。

3.6 正則化

正則化は、過学習を防ぎ、汎化性能を高めるためにモデルに事前知識を追加する手法である。モデルの複雑さに罰則を科すなど、事前知識から何らかの方法であり得なさそうなモデルを学習の候補から除外することで、適切なモデルが学習される可能性が高まる。代表的な正則化手法には、L1正則化やL2正則化がある。L1正則化はラッソ回帰、L2正則化はリッジ回帰とも呼ばれ、ラッソ回帰は前出したように特徴量選択の効果を持つ。

深層学習の正則化手法としては、早期打ち切り (Early Stopping)³⁷⁾、ドロップアウト (Dropout)³⁸⁾、バッチ正規化 (Batch Normalization)³⁹⁾が良く用いられる。なお、深層学習でも、モデルのパラメータに対して、L1やL2正則化項を追加することで過学習を防ぐこともできるが、ドロップアウトやバッチ正規化がより一般的に用いられている。

これらの正則化手法を説明する前に、深層学習で使われるミニバッチ学習について説明しておく。ミニバッチ学習は、訓練データをランダムに小さなグループ (ミニバッチ) に分割し、ミニバッチごとに重みの更新を行う手法である。これは、訓練データ全体を一度に使用するバッチ学習と異なり、計算コストが低く、学習の収束が速いことが特徴であり、深層学習では、GPUメモリが許す範囲でより大きなバッチサイズのミニバッチ学習を用いることが一般的である。ミニバッチ学習が一巡し、全ての訓練データが1回学習されるサイクルをエポックと呼び、1エポック終了後、訓練データは再度シャッフルされ、次のエポックが開始される。

早期打ち切りは、図7に示すようにエポックが進むと、訓練データに対する損失が減少し続ける一方で、検証デー

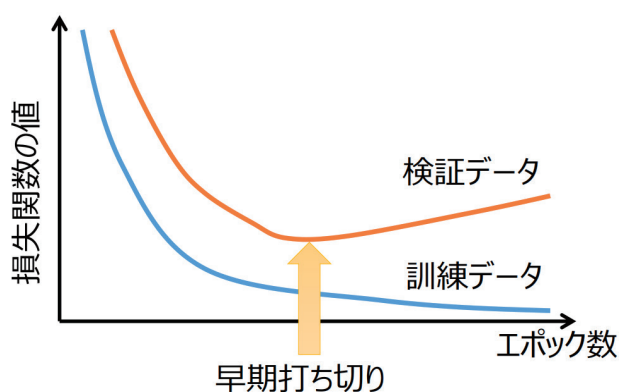


図7 早期打ち切り

タに対する損失が上昇する時点で訓練を停止する手法である³⁷⁾。これにより、過学習の兆候が見られる前にモデルの訓練を終了させ、汎化性能を向上させることができる。実際には、エポックごとに、検証データで評価を行い、最も良い検証データに対する損失を持つエポックでのモデルを採用する。

ドロップアウトは、深層学習において、訓練時にランダムに選択されたノードを無効化する手法である³⁸⁾。これにより、ニューラルネットワークが特定のノードに依存せず、より堅牢なモデルが構築される。ドロップアウトは、推論時には全てのノードを使用するので、アンサンブル学習の一種とも捉えられる。

バッチ正規化は、各層の出力をミニバッチごとに平均0、標準偏差1となるよう正規化し、学習の安定化と正則化効果をもたらす手法である³⁹⁾。これは、モデル内の特徴表現において平均と分散が寄与しないという事前知識に基づいている。バッチ正規化は強力に広く利用される手法であるが、バッチサイズが非常に小さい場合や時系列データ（例えば動画フレーム）を扱う際には、平均・分散の推定が安定せず注意が必要である。

3.7 データ拡張

正則化はモデルに関する事前知識を用いて、あり得なさそうなモデルを学習候補から除外する手法であるが、データ拡張はデータに関する事前知識を利用して、有り得そうなデータを増やす手法である³⁶⁾。データ拡張は、表形式データを含む様々なデータ形式で利用可能であるが、画像データの深層学習において特に重要な役割を果たす手法である。画像データに対するデータ拡張は、元のデータセットに対して様々な画像操作を行い、新たな訓練データを生成することができる。その結果、過学習を抑制し、汎化性能の向上が期待できる。主な画像データ拡張手法には、画像の回転、平行移動、拡大縮小、反転、明るさや色調の調整などがある³⁶⁾。これらの操作をどの程度行うかについては、特定のタスクやデータセットに依存するので、ドメイン知識を活用しつつ、検証データでの評価をもとに試行錯

誤的に適用される。例えば、物体の位置や向きが重要でない場合、データ拡張として画像の回転や平行移動を適用することが適切である。

データ拡張は学習前にデータを操作することから、前処理の一環としても捉えられる。しかし、データ拡張の画像処理は、通常、学習前に大規模な学習データセットを作成せずに実行される。これは、ストレージやメモリの使用量を削減するためである。実際の運用では、エポックごとにランダムに選択された画像処理によってデータ拡張が行われ、学習が繰り返される。その結果、データ拡張は、入力層の前に存在し、特定の処理を確率的に実行する層と見なすことも可能である。

データ拡張の手法として、基本的な画像処理の他にもCutoutやMixupなどが存在する⁴⁰⁾。Cutoutは、画像の一部をランダムにマスク（黒く塗りつぶすなど）し、モデルが画像全体に依存せず特徴を捉えることを促す効果を持つ。一方、Mixupは複数の画像とそのラベルを線形に混合し、新たなデータを生成する手法である。検査工程での製品の外観画像や顕微鏡画像を用いる場合、これらのデータ拡張手法が有効であるかは検証が必要である。さらに、GANから派生した画像生成モデルによるデータ拡張の報告も存在するが、その方法論は十分に確立されていない⁴⁰⁾。

データ拡張のポリシー（処理方法や強度）を設計することは困難であり、ハイパーパラメータチューニングと同様に、データ拡張の探索範囲も広大である。そのため、データ拡張のポリシーを自動的に最適化する手法も研究されているが、一般化されてはいない³⁶⁾。鉄鋼業における検査工程などの実用的なタスクにおいては、ドメイン知識を活用しつつ、適切なデータ拡張手法を選択し、検証データを用いた評価を行うことが求められる。

実用的な面においては、推論時データ拡張（Test Time Augmentation）は広く有効な手法である⁴⁰⁾。TTAは、学習時に用いたデータ拡張手法を推論時にも適用することで、モデルそのものは変えずに、モデルの汎化性能を底上げする手法である。具体的には、テストデータに対してデータ拡張を行い、複数の入力画像を生成する。次に、これらの入力画像に対してモデルで推論を実行し、得られた予測結果を平均や多数決を行い最終的な予測とする。ただし、TTAの効果は訓練時のデータ拡張の効果に依存し、また推論時の計算コストが増加するという実用的には大きなデメリットも考慮する必要がある。

3.8 転移学習

転移学習は、広範な概念や技術を指す言葉であり、研究者によって異なる方法で解釈されることがある⁴¹⁾。しかし、「新規タスクの効果的な仮説を効率的に見つけ出すために、一つ以上の別のタスクで学習された知識を得て、それを適用する問題」という定義は、広く受け入れられてい

るものの一つである⁴²⁾。この定義に基づくと、転移学習では既存のタスクで獲得された知識を新しいタスクに適用する手法を提案し、汎化性能の向上を目指す。特徴空間におけるデータの分布が与えられた場合、ドメインが定義されるが、知識の転移元は元ドメインと呼ばれ、転移先は目標ドメインと呼ばれる。

転移学習では、元ドメインと目標ドメインの類似性を仮定し、その仮定に基づいて適切な転移学習手法を選択する必要がある。元ドメインと目標ドメインが大きく異なる場合、元ドメインのデータを利用することで予測精度が低下する負の転移が発生することがある。このため、転移学習を適用する際には、元ドメインと目標ドメインの類似性を考慮し、適切な手法を選択することが重要である。

転移学習手法には、大きくは、事例転移、特徴転移、パラメータ転移の3つのアプローチが考えられる⁴¹⁾。事例転移は、元ドメインのデータを目標ドメインのデータとして扱う方法である。特徴転移は、元ドメインと目標ドメインのデータ間で不変な特徴表現を学習することを目指す。パラメータ転移では、元ドメインで訓練された学習器のパラメータを目標ドメインで利用するアプローチである。

深層学習の画像タスクにおいて、パラメータ転移は一般的に用いられる手法である。特に、ImageNetのような大規模なデータセットで事前に訓練されたモデル（例えば、VGG16など）が、新しいタスクやデータセットへの適用に利用される。このアプローチでは、事前学習済みの畳み込み層のパラメータをそのまま使用するか、一部の層について目標ドメインのデータでファインチューニング（微調整）を行う。この手法により、データ量が少ない場合でも高い性能が達成されることが期待される。パラメータ転移は、画像分類に限らず、深層学習のアーキテクチャを変更することで、物体検出やセグメンテーションなどのタスクにも幅広く適用されている。

このようなファインチューニングでは、一般的に Model Zooと呼ばれる、事前学習済みモデルのリポジトリが利用される。Model Zooでは、様々なデータセットで訓練された、有名なアーキテクチャのモデルが提供されており、事前学習済みモデルを利用することで、短期間で高性能なモデルを構築することができる。例えば、KerasはImageNetで学習済みのVGG16⁹⁾、VGG19⁹⁾、ResNet50V2⁴³⁾、ResNet101V2⁴³⁾、ResNet152V2⁴³⁾、InceptionV3⁴⁴⁾、InceptionResNetV2²²⁾、MobileNetV2⁴⁵⁾、Xception⁴⁶⁾、DenseNet121⁴⁷⁾、DenseNet169⁴⁷⁾、DenseNet201⁴⁷⁾など有名アーキテクチャのモデルを提供している。

これらのアーキテクチャについて簡単に紹介しておく。VGG16およびVGG19は、16または19層の畳み込み層と全結合層を持つ浅いアーキテクチャである⁹⁾。ResNet50V2、ResNet101V2、ResNet152V2は、残差接続を使用し、50層から152層の深い畳み込み層で

の学習が可能である⁴³⁾。InceptionV3は、Inceptionモジュールによりネットワークが分岐し、異なるサイズの畳み込みを行う⁴⁴⁾。InceptionResNetV2は、Inceptionモジュールに残差接続を適用し、深層化が実現されている²²⁾。MobileNetV2は、モバイル端末での使用に適した、パラメータ数が削減されたアーキテクチャである⁴⁵⁾。Xceptionは、Inceptionの改良版であり、高精度化とパラメータ数の削減が図られている⁴⁶⁾。DenseNet121、DenseNet169、DenseNet201は、ResNetの改良版であり、ショートカット接続を導入することで、121層から201層までの深層化が可能となっている⁴⁷⁾。これらのアーキテクチャは異なる特性を持つことから、転移学習を行うにもタスクやデータセットに応じて適切なモデルを選択することが求められる。

3.9 不均衡データ

ここまで、少ないデータで汎化性能を向上させるための技術を解説してきたが、データ量だけでなく、不均衡データも機械学習において重要な問題である。不均衡データとは、クラス間のデータ数に偏りがあるデータセットを指し、多くの実問題ではこのような不均衡が存在する。不均衡データへの対処方法も幾つか提案されている⁴⁸⁾。

不均衡データへの対処として一般的な方法は、オーバーサンプリングとアンダーサンプリングである。オーバーサンプリングは、少数クラスのデータ数を増加させることでクラス間のバランスを調整する手法である。一方、アンダーサンプリングは、多数クラスのデータ数を減少させることでクラス間のバランスを調整する手法である。具体的には、表形式のデータにおいては、データセットから少数クラスの事例の行を複製する。画像データにおいては、少数クラスの事例の画像を複製し、データ数を増やす。表形式データのオーバーサンプリングではSMOTE⁴⁹⁾などの少数クラスの事例を合成する手法も開発されている。SMOTEでは少数クラスの事例に対して、そのk近傍サンプルをランダムに選び、選ばれた近傍サンプルとの内挿により新たな事例を生成する。SMOTEは新たなデータを合成しており、データ拡張の手法とも言える。

そもそも、不均衡データではテストデータや評価指標の選択も重要である。データセットが不均衡のままテストデータに分割して評価すると、多数クラスに対して過剰に適合された評価を行うことになる。そのため、正解率の他に、適合率、再現率、F値などを用いることがある。

データに不均衡の対策を施すだけでなく、機械学習アルゴリズムに不均衡データへの対策を組み込むことも一つのアプローチである。具体的には、重み付きクロスエントロピー誤差のように、損失関数にクラスの重み付けを導入することで、少数クラスの誤分類に対するペナルティを増加させることが可能となる¹²⁾。この方法により、モデルは少

数クラスに対してより敏感になり、全体の分類精度が向上することが期待される。

画像のセグメンテーションタスクでは、損失関数に重み付けを導入するアプローチが特に重要である¹²⁾。セグメンテーションとは、画像内の各ピクセルに意味的なクラスラベルを割り当て、画像を複数の領域に分割する処理である。例として、金属組織画像から粒度を測定する際、粒界をセグメント化することが挙げられる。このタスクでは、粒界のピクセルは他のピクセルに比べて少数クラスとなる。粒界のピクセルを画像中で人為的に増加させることは困難であるため、損失関数に重み付けを行うことで対処が可能である。

3.10 半教師あり学習

特定の状況では、少量のデータに対応する手法として、半教師あり学習が効果的である⁵⁰⁾。半教師あり学習は、少量のラベル付きデータと多数のラベルなしデータを活用し、機械学習モデルの性能向上を目指す手法である⁵⁰⁾。画像データのタスク、特にセグメンテーションにおいて、データの取得は容易だが、アノテーションコストが高いという状況があり得る。アノテーションは、データにラベルを付与する作業であり、専門知識が求められる場合や手作業が必要な場合には、コストが大きくなる。ラベルなしデータの活用は、この問題を解決する重要なアプローチである。

半教師あり学習には様々なアプローチが存在するが、代表的な擬似ラベリング手法では、初めに与えられたラベル付きデータを用いて予測モデルを学習し、そのモデルでラベルなしデータにラベルを付与する⁵⁰⁾。次に、元のラベル付きデータセットに新たなラベル付きデータを加え、学習を再度実行する。このプロセスを繰り返すことで性能向上が図られる。多くの半教師あり学習手法では、信頼度で選抜するなどのアルゴリズムを用いて自動的にラベルなしデータのラベルを推定する。一方で、半教師あり能動学習では、重要と判断されるラベルなしデータを選択し、人間の専門家がラベル付けを行うことで性能向上が期待される。これらの手法を適切に適用することで、限られたラベル付きデータを活用しながら、モデルの汎化性能を向上させることが可能となる。

3.11 機械学習を解釈する技術

機械学習の解釈可能性を追求する技術であるXAIの目的は、機械学習モデルの理解を促進することである⁵¹⁾。ただし、解釈性の向上は性能向上にも寄与する場合がある。解釈性が高いモデルは、特徴量と予測結果の関係性を明確に示すことができるため、モデルが適切な予測を行っているかどうかの判断が容易になる。このような理解しやすいモデルとドメイン知識を利用することで、不適切な特徴量の

削除や、新たな特徴量の追加を実施できる。

機械学習を解釈する技術として、線形回帰は最も理解しやすい手法である。線形回帰モデルの各説明変数に対応する回帰係数は、その変数が目的変数に与える影響の大きさと方向を表す。標準化回帰係数は、各説明変数を標準化した後に計算され、異なる尺度の説明変数間で影響力を比較しやすくなる。ラッソ回帰では、特徴選択の効果によりモデルの解釈性が更に向上する。ただし、線形回帰では、非線形性や交互作用を捉える能力に限界があることに注意が必要である。また、汎化性能に優れた複雑な機械学習モデルとは別に、解釈のために線形回帰モデルを用意する必要がある。

機械学習の解釈技術には、大局的な手法から局所的な手法まで多数が提案されているが、ここで紹介する手法は機械学習モデルを問わず適用することができる⁵¹⁾。大局的な手法では、モデルの精度の観点から特徴量の重要度を評価する。例えば、PFI (Permutation Feature Importance) では、特徴量をランダムに並べ替えたデータセットでモデルの予測精度を測定する。並べ替えによる予測精度の低下が大きいほど、その特徴量は重要であると判断される。また、LOCOFI (Leave-One-Covariate-Out Feature Importance) では、モデルの評価指標に対して、特定の特徴量を除いた場合と含んだ場合の差を計算することで、その特徴量の重要度を評価する。ステップワイズ法による特徴量選択は、LOCOFIの結果を利用しているとも言える。

より局所的な手法では、感度分析により特徴量の影響を評価する。例えば、PD (Partial Dependence) では、他の特徴量の影響を平均化し、特定の特徴量が予測結果に与える影響を可視化する。PDプロットは、特徴量のある範囲における予測結果の変化を示すため、特徴量と予測結果の関係を把握できる。一方、ICE (Individual Conditional Expectation) は、PDと同様に特徴量と予測結果の関係を可視化するものの、他の特徴量の影響を平均化せず、個々のデータ点に対して計算を行い、可視化を行う。これにより、ICEプロットでは、モデルの挙動を個々の事例における特徴量と予測結果の関係として把握することが可能である。ただし、ICEプロットは局所的な個々のデータ点に対して計算されるため、複数の曲線が重なり合い、解釈しづらいことがある。

機械学習モデルを局所的にも大局的にも可視化できるのはSHAP (SHapley Additive exPlanations) である⁵²⁾。SHAPはゲーム理論に基づく手法で、各特徴量が予測結果に対してどれだけ貢献しているかをSHAP値と呼ばれる数値で評価する。SHAP値は各特徴量の寄与度を表しており、SHAPプロットでは、事例ごとにSHAP値と各特徴量をプロットすることで、特徴量の影響力と方向を可視化することができる。SHAPは非線形性や交互作用を考慮しながら、機械学習モデルを問わず、局所的な解釈だけでなく、

大局的な解釈も提供できるので、解釈性を向上させる強力な手法である。

4. 機械学習の適用事例

既に機械学習は鉄鋼業でのプロセス制御と検査工程で適用が進んでいる。ここでは、著者が関わった研究の中から抜粋し、プロセス制御の事例としてダクタイル鋳鉄でのMg歩留まり⁵³⁾、検査工程の事例として合金鋼とステンレス鋼での破面解析⁵⁴⁾を紹介する。

4.1 ダクタイル鋳鉄でのMg歩留まり⁵³⁾

Mg処理は、鋳鉄における代表的な黒鉛球状化処理方法であり、多くの鋳物工場で導入されている。Mg濃度が不足すると、黒鉛の球状化が進まず、機械的性質が不十分となる。そのため、Mg処理の歩留まりを予測し、Mg濃度が下限値以上になるように投入Mg量を算出するという管理方法が行われている。Mg歩留まり(Mg yield rate)は、熟練技術者が溶鉄湯中の硫黄濃度などのデータをもとに、Mg濃度が下限値以上になるように過剰にMgを投入するために、その実測値よりも低くなるように経験的に予測している。このような熟練技術者の暗黙知を伝承することは困難である。

そこで、熟練技術者の代わりに人工知能を用いたプロセス制御が求められていた。既にMg歩留まりを予測するための機械学習モデルは報告されていた。しかし、実測歩留まりを下回るように歩留まりを予測するという熟練技術者が行っているタスクを代替することはできていなかった。そこで、本研究⁵³⁾では、損失関数として分位損失を用いることで、実測値よりも低くMg歩留まりを予測する機械学習モデルを提案した。なお、分位損失を歩留まりの管理に適用したのは本研究が初めてであった。

データセットにはダクタイル鋳鉄管の製造における、44,969杯の取鍋のMg歩留まりとMg処理に関するプロセスデータを用いた。プロセスデータには、溶鉄湯の炭素、アルミニウム、硫黄、チタンなどの濃度や温度などが含まれていた。

まず、データの関係性を理解するために、機械学習を解釈する技術であるラッソ回帰を行った結果が図8aである。ラッソ回帰では、正則化パラメータが大きいほど重要度の低い係数が削除されるため、正則化パラメータが大きくても標準化偏回帰係数が大きい変数ほど、影響が大きい説明変数であると解釈できる。また、標準化偏回帰係数が正に大きければMg歩留まりは向上し、負に大きければ低下することを意味する。図から、硫黄濃度(歩留まり低下)と炭素濃度(歩留まり向上)の影響が大きく、Mg処理の冶金学的な知識と矛盾していないことが分かった。

しかし、このデータセットでは十分な性能が得られな

かったため、熟練技術者からMg歩留まり予測において考慮している因子について聞き取りを行った。その結果、熟練技術者は硫黄濃度などの他に、過去直近のMg処理のデータを参考にしているとの回答が得られた。ただし、具体的にどの因子がどのように影響するのかは形式知化できておらず、暗黙知の状態であった。そこで、熟練技術者からの聞き取りに基づき、前回杯及び前々回杯のMg処理に関するプロセスデータを熟練技術者の暗黙知としてデータセットに加えた。

熟練技術者の暗黙知を加えたデータセットでのラッソ回帰の結果を図8bに示すが、前回杯のMg歩留まりと前々回杯のMg歩留まりの影響が大きい。熟練技術者の暗黙知を含めることで予測精度が向上し、熟練技術者の暗黙知であったMg歩留まりに及ぼす因子を形式知化することができた。なお、前回杯・前々回杯のMg歩留まりの影響が大きいという分析結果は、用いた説明変数には観測できていない因子が含まれていることを示唆している。ここからは機械学習の範疇ではないが、仮説として、データに含まれていない酸素濃度などの影響が大きいと考察した。

熟練技術者の暗黙知を加えたデータセットと損失関数に

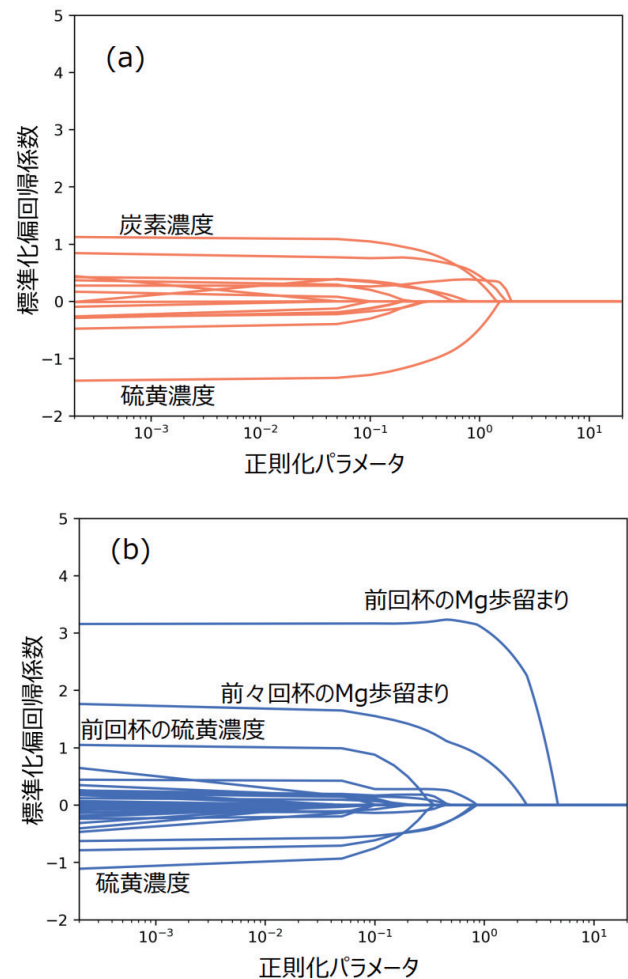


図8 熟練技術者による暗黙知を (a) 含まない場合と (b) 含む場合のラッソ回帰の標準化偏回帰係数(文献⁵³⁾より改変引用)。

分位損失を用いた提案手法により、Mg歩留まりの実測値よりも低く予測する機械学習モデルが構築できた。その際、分位レベルを決定するハイパーパラメータは、精度評価指標のコストに対してグリッドサーチを行うことで決定した。訓練データとテストデータでの精度評価結果を表1に示す。なお、分位レベル以外のハイパーパラメータの調整には訓練データを5分割交差検証で検証を行い、グリッドサーチを用いた。

表1 訓練データとテストデータでの精度評価結果。精度評価指標のコストは平均絶対誤差と不良率より算出(文献⁵³⁾より改変引用)。

損失関数	手法	データ	平均絶対誤差 (%)	不良率	コスト (%)
二乗誤差	線形回帰	訓練	2.46	0.5106	104.57
		テスト	2.46	0.5112	104.71
分位損失	線形回帰	訓練	7.77	0.0090	9.57
		テスト	7.84	0.0091	9.66
分位損失	勾配ブースティング	訓練	7.70	0.0080	9.31
		テスト	7.77	0.0088	9.53
分位損失	ニューラルネットワーク	訓練	7.64	0.0084	9.33
		テスト	7.77	0.0092	9.62

表1では、Mg歩留まりを百分率で表記するため、精度評価指標となる平均絶対誤差とコストも百分率で表記している。二乗誤差を用いた線形回帰では、平均絶対誤差は最小となるが、ほぼ半数が不良杯となるため、コストは最も高かった。一方で、分位損失を用いた場合には、コストを大幅に低下でき、Mg歩留まりの実測値を下回り、かつMg歩留まりの実測値とよく一致して予測する機械学習モデルの構築ができた。構築した機械学習モデルにおいて、テストデータでのコストが示すように、勾配ブースティング、ニューラルネットワーク、線形回帰の順に汎化性能に優れた機械学習アルゴリズムであった。

本研究で構築した勾配ブースティングの機械学習モデルを使用して、鑄鉄管工場で概念実証を行ったところ、機械学習モデルの方が熟練技術者よりも精度評価指標であるコストを低くできた。今後の課題として、構築した機械学習モデルをプロセス制御の自動システムとして実装することを進めている。

4.2 合金鋼とステンレス鋼での破面解析⁵⁴⁾

鉄鋼材料は、多くの製品に使用されており、その寿命は、使用状況や周囲環境によって影響を受け、やがて破損することが避けられない。破壊原因を特定するためには、破断面の破面解析(fractography)が必要である。破面解析は、検査工程と同様に、製品の品質を評価し、維持・向上させるための重要なプロセスである。破面解析において、解析者は破断面の模様を解釈し、破壊機構の観点から破面分類を行う。これは検査工程と同様に、熟練者による観察と経験に基づく複雑な知的判断を必要とする作業である。解析

者の経験は、言語化できない暗黙知として蓄積され、伝承が容易ではない。一方で、熟練した解析者は徐々に引退している。これらの理由から、破面解析の初心者を支援し、熟練した解析者の経験を伝承するシステムが求められる。

CNNを用いた破面分類が報告されているが、先行研究では、合金鋼とステンレス鋼の違いのような材料の違いを考慮しておらず、すべての材料を同一の破断面のクラスで分類していた。熟練した解析者は、材料の違いを暗黙知として考慮していると考えられる。そこで、本研究⁵⁴⁾では、材料の種類を考慮した転移学習により、CNNを用いた破面分類の高精度化を実現した。そのため、本研究では、他の材料からの学習結果を利用するために、新しい転移学習手法としてSSSC(Stepwise Selection of Source Classes)を提案した。

既存手法の破断面の分類では、材料を区別しないため、元ドメインのみを用いた学習になる。ディンプル、ストライエーション、ぜい性疲労破面、へき開破面、擬へき開破面、粒界破面、ファンシェーブトパターンの7クラス分類で、元ドメインのみを用いた分類モデルを図9aに示す。一方で、材料の種類ごとにそのまま用いた学習を図9bに示すが、これは目標ドメインのみを用いた学習となる。提案手法のSSSCでは、元クラスと目標クラスを混合した事例で複数のモデルを学習し、目標ドメインのデータで評価を行う。そして、ステップワイズ法により元クラスと目標クラスの組み合わせを探索する。

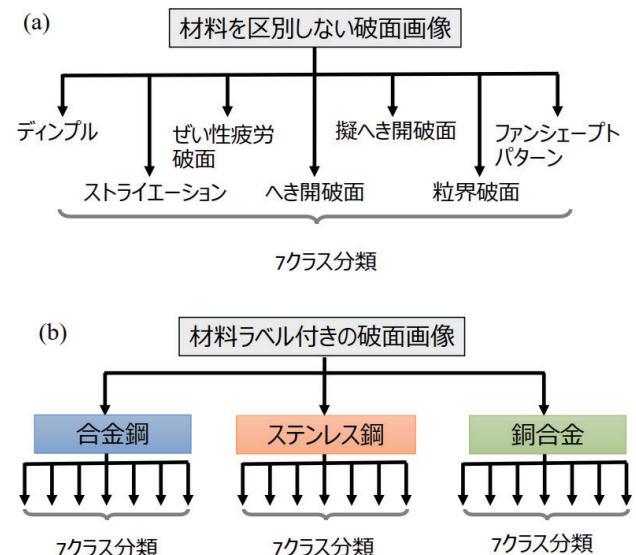


図9 (a) 元ドメインでの学習と (b) 目標ドメインでの学習(文献⁵⁴⁾より改変引用)。

データセットには、破面解析の経験豊富な専門解析者がラベル付けした走査型電子顕微鏡(SEM)画像の1033枚の破断面画像を用いた。このデータセットにはドメイン別(ここでは鋼種別)では、合金鋼は670枚、ステンレス鋼は230枚、銅合金は133枚の画像が含まれていた。クラス

別(ここでは破面形態の違い)では、ディンプルは231枚、ストライエーションは212枚、ぜい性疲労破面は164枚、へき開破面は127枚、擬へき開破面は105枚、粒界破面は161枚、ファンシェーブトパターンは33枚の画像が含まれていた。

用いたデータ量は十分とは言えず、また不均衡データである。そのため、本寄稿で紹介した、様々な技術を用いた。具体的には訓練データについては3.9節で紹介したオーバーサンプリングにより少数クラスの画像を複製し、データ数を増やしてクラス不均衡を解消した。検証のための検証データとテストデータの事例数はそれぞれ各クラスで5とした。前処理では画像のリサイズを行った。図10には検証データの正解率と画像サイズ(解像度)の関係を示すが、画像サイズは360×270が最適であることが分かる。正則化手法としては、早期打ち切り、ドロップアウト、バッチ正則化を使用し、活性化関数の種類やノード数などのハイパーパラメータのチューニングも検証データを用いて行った。データ拡張は訓練データに対してのみ行い、検証データを用いて検証を行い、画像に対してランダムに移動、拡大・縮小、回転、せん断、反転の操作を行った。CNNのアーキテクチャについては、VGG16、VGG19、ResNet50V2、ResNet101V2、ResNet152V2、InceptionV3、InceptionResNetV2、MobileNetV2、Xception、DenseNet121、DenseNet169、DenseNet201の候補の中から、検証データでの結果に基づき、VGG16でのファインチューニングを採用した。

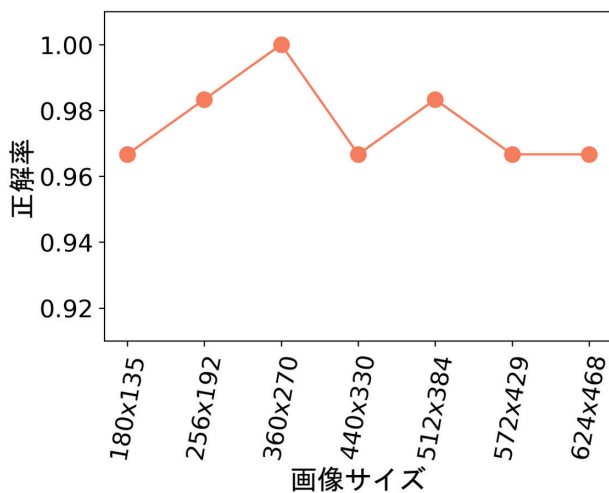


図10 正解率と画像サイズの関係(文献⁵⁴⁾より改変引用)。

前処理を行ったデータとチューニングされたVGG16モデルでSCCCによる元クラス選択を行った。図11に前進法と後退法での、ステンレス鋼での正解率と元クラス数の関係を示す。元クラス数が0と7の時は転移学習ではな

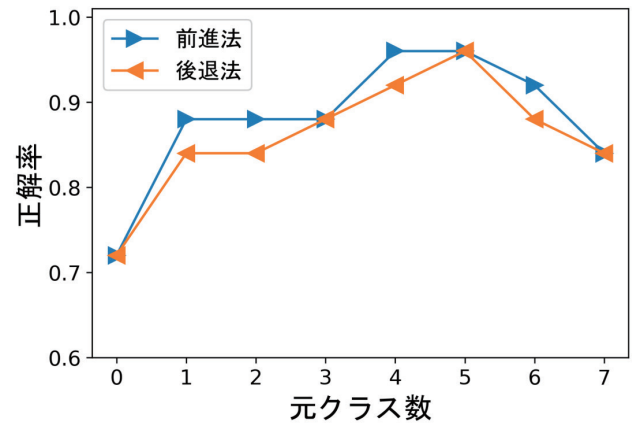


図11 ステンレス鋼での正解率と元クラス数の関係(文献⁵⁴⁾より改変引用)。

く、それぞれ、目標ドメインでの学習と元ドメインでの学習である。元クラス数が4もしくは5の時に最も高い正解率を示していた。元クラスと目標クラスの選択には、両者のデータ数の比率が関係しており、目標クラスの訓練データが十分に多ければ、目標クラスが選択されるが、目標クラスの訓練データが少なければ元クラスが選択されていた。

表2に目標ドメインでの学習、元ドメインでの学習、そしてSSSCを用いた時のテストデータの正解率を示す。元ドメインでの学習は、材料を区別しない既存手法であり、合金鋼、ステンレス鋼、銅合金での平均正解率は92.4%であった。目標ドメインでの学習では、材料の種類を考慮し、材料ごとにそのまま学習させたが、平均正解率は81.5%であり、元ドメインでの学習よりも悪化していた。一方で、提案手法のSSSCでは平均正解率は98.7%であり、既存手法から6.3%向上していた。材料の種類を考慮した転移学習手法であるSSSCが、破面分類の精度向上に寄与することが明らかとなった。

表2 テストデータの正解率(文献⁵⁴⁾より改変引用)

ドメイン	目標ドメインでの学習	元ドメインでの学習	提案手法 SSSC
合金鋼	0.933	0.933	1.000
ステンレス鋼	0.720	0.840	0.960
銅合金	0.792	1.000	1.000
平均	0.815	0.924	0.987

本研究の目標ドメインは、合金鋼、ステンレス鋼、銅合金の3種類に限定されていたが、今後、他の材料でのデータが収集できれば、SSSCによる転移学習により、新たな材料でも高い予測精度が期待できる。また、本研究で開発したモデルは解析者の育成に既に試験的に利用されており、今後は、Webサービスとして利用公開を検討している。

5. おわりに

DX (Digital Transformation) とは、企業がAI、IoT、ビッグデータなどの技術を活用し、業務の改善や新規ビジネスモデルの創出を行うことである。鉄鋼業においても、DXが注目されているが、DXに類似した言葉にはインダストリー 4.0 というのもあった。このようなもっともらしいが実際には意味が曖昧な用語のことをバズワードという。バズワードはいずれ、その具体的な事例が増えて、抽象概念が共有されれば、個別の用語に置き換わり、消える宿命にある。しかし、バズワードには機械学習のような特定技術の普及期において、その知見を多くの人々が共有するための役割がある。そのため、本寄稿を総括するにあたり、DXという言葉を使用することは適切と思われる。

本寄稿では、既に機械学習は、鉄鋼業においてプロセス制御や検査工程の自動化や効率化を通じて、実質的には工程省略に貢献していることを解説した。ただし、機械学習の活用は、データ数が少ない場合など限定的な状況においては、適切な技術の選択が求められるので、本寄稿では、その技術について詳細に述べた。また、データ量が技術以上に重要であることを指摘した。

機械学習は、多くのデータから一般的な規則を見出す人間の行う帰納法を置き換えたものである。従って、無差別にデータを収集しても、業務改善につながる機械学習モデルを構築することはできない。業務改善に貢献する課題の特定と、その課題に対応できるデータが必要である。適切な課題設定とデータ蓄積を実行できる人材が、DXを成功させる上で欠かせない要素である。

そのようなDX人材は機械学習だけでなく情報処理技術全般を広く網羅しておく必要がある。機械学習モデルによる予測は、人間の帰納推論と同様に必ず不確実性が存在するため、完全な精度と説明が要求される場合や問題が単純な場合には、ルールベースのアプローチの適用を検討すべきである。また、機械学習が対処する課題の要求精度に達しないと予想される場合には、課題設定の見直しが必要となる。特に、データ蓄積に関するコストと課題解決による効果の費用対効果を適切に見積るには、情報処理技術以外にも含む幅広い知識が求められる。

今後も鉄鋼業において、機械学習とその関連する情報処理技術の発展が続くことで、さらなる効率化が期待される。そのためには、適切な技術を選択し、活用することが、鉄鋼業における持続的な競争力向上につながる。バズワードに惑わされず、具体的な課題の整理と、必要なデータを蓄積する体制を継続的に整えることが重要である。

参考文献

- 1) 野坂康雄: 鉄と鋼, 59 (1973), 557-569.
- 2) 大重貴彦: 計測と制御, 55 (2016), 228-233.
- 3) 玉置久, 小林敬和: 鉄と鋼, 97 (2011), 315-315.
- 4) 榊剛史, 井上昂治, 山野泰子, 岩澤駿, 松原仁, 杉本舞, 谷口忠大: 人工知能, 38 (2023), 298-301.
- 5) 榊剛史, 井上昂治, 山野泰子, 岩澤駿, 杉本舞, 松原仁, 谷口忠大: 人工知能, 38 (2023), 92-95.
- 6) 足立吉隆: まてりあ, 58 (2019), 6.
- 7) 中島智晴: 材料, 69 (2020), 633-639.
- 8) 上杉徳照: 軽金属, 73 (2023), 104-111.
- 9) K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition" The 3rd International Conference on Learning Representations (ICLR2015) (2015).
- 10) 平田丈英, 松下昌史, 飯塚幸理, 鈴木宣嗣: 鉄と鋼, 107 (2021), 897-905.
- 11) R. Koenker and G. Bassett: *Econometrica*, 46 (1978), 33-50.
- 12) O. Ronneberger, P. Fischer and T. Brox: *Lect. Notes Comput. Sci.*, 9351 (2015), 234-241.
- 13) R. Tibshirani: *J. R. Statist. Soc. B*, 58 (1996), 267-288.
- 14) L. Breiman: *Mach. Learn.*, 45 (2001), 5-32.
- 15) H. Drucker, C. J. C. Surges, L. Kaufman, A. Smola and V. Vapnik: *Advances in Neural Information Processing Systems*, NIPS 1996, (1997), 155-161.
- 16) T. Chen and C. Guestrin: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (2016), 785-794.
- 17) G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T. Y. Liu: *Advances in Neural Information Processing Systems*, NIPS 2017, (2017), 3147-3155.
- 18) L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Drogush and A. Gulin: *Advances in Neural Information Processing Systems*, NeurIPS 2018, (2018), 6638-6648.
- 19) Kaggle: Your Machine Learning and Data Science Community, <https://www.kaggle.com/>.
- 20) A. Krizhevsky, I. Sutskever and G. E. Hinton: *Advances in Neural Information Processing Systems*, NIPS 2012, 2 (2012), 1097-1105.
- 21) O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg and L. Fei-Fei:

- Int. J. Comput. Vis., 115 (2015) , 211-252.
- 22) C. Szegedy, S. Ioffe, V. Vanhoucke and A. Alemi: 31st AAAI Conference on Artificial Intelligence, AAAI 2017, (2017) , 4278-4284.
- 23) Y. Gorishniy, I. Rubachev, V. Khulkov and A. Babenko: Advances in Neural Information Processing Systems, NeurIPS 2021, 23 (2021) , 18932-18943.
- 24) N. Morgan and H. Bourlard: Advances in Neural Information Processing Systems, NIPS 1989, (1989) , 630-637.
- 25) J. Hestness, S. Narang, N. Ardalani, G. Diamos, H. Jun, H. Kianinejad, M. M. A. Patwary, Y. Yang and Y. Zhou: Deep Learning Scaling is Predictable, Empirically, *arXiv preprint arXiv:1712.00409* (2017) .
- 26) A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin: Advances in Neural Information Processing Systems, NIPS 2017, (2017) , 5999-6009.
- 27) J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu and D. Amodei: Scaling Laws for Neural Language Models, *arXiv:2001.08361* (2020) .
- 28) L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike and R. Lowe: Training language models to follow instructions with human feedback, *arXiv:2203.02155* (2022) .
- 29) D. H. Wolpert: Neural Comp., 8 (1996) , 1341-1390.
- 30) scikit-learn: machine learning in Python, <https://scikit-learn.org/>.
- 31) M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu and X. Zheng: Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016, (2016) , 265-283.
- 32) F. Chollet: Keras, <https://github.com/fchollet/keras>.
- 33) A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai and S. Chintala: Advances in Neural Information Processing Systems, NeurIPS 2019, 32 (2019) .
- 34) S. Falkner, A. Klein and F. Hutter: 35th International Conference on Machine Learning, ICML 2018, 4 (2018) , 2323-2341.
- 35) A. Nazri and O. Agbolade: Int. J. Eng. Tech., 7 (2018) , 322-325.
- 36) N. E. Khalifa, M. Loey and S. Mirjalili: Artif. Intell. Rev., 55 (2022) , 2351-2377.
- 37) L. Prechelt: Lect. Notes Comput. Sci., 7700 (2012) , 53-67.
- 38) N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov: J. Mac. Learn. Res., 15 (2014) , 1929-1958.
- 39) S. Ioffe and C. Szegedy: 32nd International Conference on Machine Learning, ICML 2015, 1 (2015) , 448-456.
- 40) C. Shorten and T. M. Khoshgoftaar: J. Big Data, 6 (2019) .
- 41) K. Weiss, T. M. Khoshgoftaar and D. D. Wang: J. Big Data, 3 (2016) , 9.
- 42) 神島敏弘: 人工知能学会第二種研究会資料, SIG-DMSM-A803-06 (2009) , 36-43.
- 43) K. He, X. Zhang, S. Ren and J. Sun: Lect. Notes Comput. Sci., 9908 (2016) , 630-645.
- 44) C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (2016) , 2818-2826.
- 45) M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. C. Chen: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (2018) , 4510-4520.
- 46) F. Chollet: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, (2017) , 1800-1807.
- 47) G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, (2017) , 2261-2269.
- 48) H. He and E. A. Garcia: IEEE Trans. Knowl. Data Eng., 21 (2009) , 1263-1284.
- 49) N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer: J Artif. Intell. Res., 16 (2002) , 321-357.
- 50) Z. H. Zhou: Natl. Sci. Rev., 5 (2018) , 44-53.

- 51) 森下光之助：機械学習を解釈する技術，技術評論社，(2021)。
- 52) S. M. Lundberg and S. I. Lee: Advances in Neural Information Processing Systems, NeurIPS 2017, (2017), 4766-4775.
- 53) 福島瑞貴，上杉徳照，辻川正人，堤親平，小川耕平，澤田健二，中本光二：鑄造工学，94 (2022)，69-75.
- 54) 上杉徳照，近藤充晟，濱田真行，喜多俊輔，平田智丈：材料，72 (2023)，376-383.