

[機械学習工学]

5 機械学習のための ヒューマンインタフェース



五十嵐健夫 | 東京大学

機械学習におけるヒューマンファクタ

機械学習を利用したシステムというと、多くの人は自動運転のように、まったく人間がかかわることなく全自動で動くシステムを思い浮かべるかもしれない。しかし、よく考えてみると、たとえ全自動で動くシステムであったとしても、あらゆる人工物がそうであるように、そのシステムを「作った人間」と、そのシステムを「利用する人間」がいるはずである。そして、人間がかかわるあらゆる側面において、ヒューマンインタフェースを考慮したシステムの設計や運用が必要不可欠である。本稿では、機械学習応用システムの構築と利用におけるヒューマンインタフェースに関する技術について、最近の研究動向を中心に紹介する。

ヒューマンインタフェースの観点から、従来型の明示的なプログラミングによるシステムと機械学習に基づくシステムの構築を比較した場合、大きな特徴として訓練データの存在と動作の帰納的定義性がある。機械学習応用システムの動作は、アルゴリズムだけでなく訓練データによって定義されるので、訓練データを適切に準備することが問題解決にとってクリティカルである。よって、機械学習応用システムの構築にあたっては訓練データを準備するプロセスを支援するヒューマンインタフェースが求められる。また、機械学習の動作は、事前に演繹的に動作を定義するといったことができないので、その構築にあたっては動作を確認しながら修正・調整するというプロセスが必要となる。よって、そのような機械学習応用システムの動作を制御するようなヒューマンインタフェースが求められる。

また、機械学習応用システムの利用という観点から考えると、演繹的なロジックに基づいて動作するシステムと比較した場合に、結果が100%正しいことは保証されないという点が重要である。また、フィードバックという形で訓練データを増やすことによって、システムの動作を改善していくことができるという点も、機械学習応用システムの特徴といえる。よって、機械学習応用システムのエンドユーザのためのヒューマンインタフェースのデザインにあたっては、このような不確実性や修正可能性について考慮することが必要となる。

以下では、上記のような観点から、機械学習応用システムの構築を支援するものとして、訓練データを準備するプロセスと、訓練プロセスを制御するプロセスを支援するもの、そして、機械学習応用システムの最終的な利用を支援するものについて具体的な研究例を挙げながら紹介する。なお、機械学習応用システムの構築にかかわる「人間」としては、機械学習の専門家やプログラマも存在するが、ここでは主にそういった技術者が作った基盤技術を利用して、実際の具体的な問題解決に取り組むドメインエキスパート(実務者)をサポートする技術について論じる^{☆1}。これらのドメインエキスパートはシステムの構築にかかわることもあれば、純粹にシステムの利用者としてかかわることもあり得る。本稿では、機械学習を利用したシステムの構築にかかわるドメインエキスパートとシステムの最終的な利用者を合わせてユーザと呼ぶこととし、後者のみの場合をエン

^{☆1} ハイパーパラメータのチューニングなど、機械学習のエキスパートによるインタラクションという観点もあるが、本稿では議論の対象外とする。

ドユーザと呼ぶこととする。

訓練データの作成に関する インタラクション

多くの機械学習応用システムに関する議論では、訓練に用いる訓練データは与えられたものとして扱うことが多い。たとえば、画像認識では、ImageNetのような既存データセットを訓練データとして扱い、その生成プロセスについては議論しないことが多い。しかし、実際にユーザ（ドメインエキスパート）が自分自身の固有の問題を解くために機械学習を利用するためには、訓練データを自分で用意することが必要である。ここでは、このような「ユーザが自分自身でデータを作成する、あるいはラベル付けする」プロセスを支援する技術についていくつか紹介する。

データ前処理のための可視化

情報システムからは膨大な量の素データが得られるが、素データにはノイズや欠損が含まれることが通常であり、解析を行う前に適切な処理が必要である。また、解析にかける前に、データを適切に編集する必要があることも多い。この作業は定型化することが難しく、データの意味を理解した上での判断が必要である。Kandelらは、このような前処理のことを data wrangling と呼び、具体的な事例を列挙しつつ、そのプロセスを支援する可視化手法を紹介している^{☆2}。具体的には、素データの問題点を発見する手法（図-1）、問題のあるデータを許容しつつ解析を行う手法、素データを機械学習に適した形へ編集する手法、さらに得られた編集手続きをシェアしたり再利用したりする手法などについて論じている。本論文で扱っている主な対象は、データマイニングやデータ分析といったものであるが、機械学習を用いる場合にもそのまま当てはまるものであるといえる。なお、機械学習のための情報可視化技術全般につい

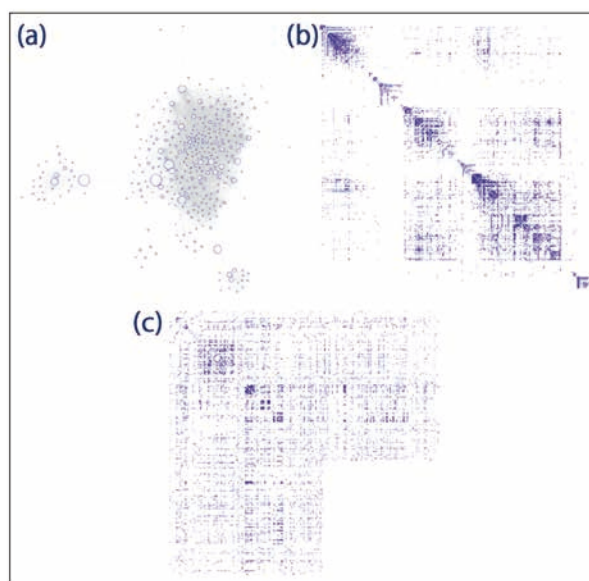
^{☆2} <http://journals.sagepub.com/doi/abs/10.1177/1473871611415994>

ては、Sacha らの論文によくまとめられている¹⁾。

対話的機械学習による画像分類

Forgarty らは、機械学習による画像分類のための訓練データ作成（ラベル付け）を対話的に行うシステムを提案している（図-2）^{☆3}。ユーザが訓練データとして画像にラベル付けを行うと、システムがそこからルールを学習し、それを残りの画像に対して適用して自動的にラベル付けを行う。システムからのフィードバックを受けずにラベル付けを行った場合と、ラベル付けのスコアの最も高いものと最も低いものをフィードバックとして提示してラベル付けを支援した場合とを比較し、後者の場合により精度の高いモデルを学習できることを示している。同じグループからの後続研究では、単純にスコアの高い/低い画像だけでなく、画像データセットの分布の概要をつかめるようにうまくサンプリングした画像を提示するとより良いラベル付けが行える、ということや、対話的なラベル付けプロセスにおいて undo や redo といった履歴管理機能をつけることでラベル付

^{☆3} <https://dl.acm.org/citation.cfm?id=1357061>



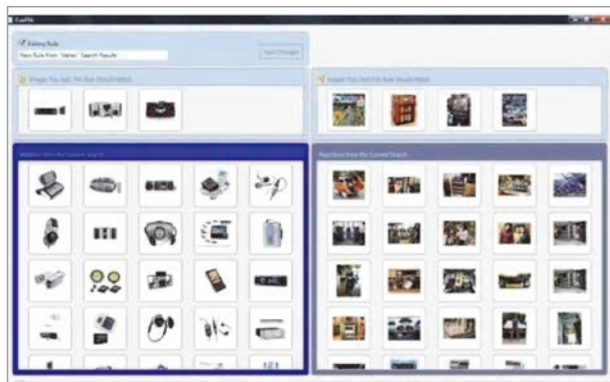
■ 図-1 素データの可視化の例。一般的な可視化 (a,b) では見つけられないデータ欠損が適切な可視化で見えるようになっている (c)。

けがより効率的に行えることなどが示されている。

なお、このようなアプローチは、一方的にデータやラベルをユーザがシステムに与えるのではなく、システムからのフィードバックを受けつつユーザがデータやラベルを与えていくという意味で、「対話的機械学習」²⁾と呼ばれている。関連した概念として、能動的学習というものがある。能動的学習では、システムの側で次々どのような訓練データを与えられれば最も効率良く訓練が進むか、を計算する。ロボットが自律的に訓練データを獲得する場合などはこのような能動的学習の考え方が有効である。しかし、対話的学習ではデータを与えるのは人間なので、単純にシステム側にとって最も都合の良いデータを次々提示されることはストレスになる可能性もある。したがって、対話的学習では、一般的な能動的学習のロジックをそのまま適用するのではなく、人間の特性に配慮した設計を行うことが必要である。

対話的強化学習による行動獲得

エージェントに行動ルールを自動獲得させる手法として強化学習がある。これは、エージェントの行動に対して適切に報酬を与えると、その報酬を最大化するように行動ルールを調整していくというものである。通常の強化学習では、事前にどのような状況になったらどのような報酬を与えるか、をルールとして記述しておく。Thomaz らは、事前に決めたルールに加えて、ユーザがエージェントの学習中の行動を観察し適宜報酬を与



■ 図-2 CueFlik. 上に表示されているのが与えた訓練データ。下に表示されているのが現在の分類結果

えることで、より効率的に学習できることを実験によって示している(図-3)^{☆4}。さらに、人間の報酬の与え方の特徴として、a) ユーザは報酬を本来あるべき「これまでの動作」に対する評価としてだけでなく、「これから何をしてほしいか」を伝える方法として使っている、b) ユーザは否定的な報酬(懲罰)よりも肯定的な報酬を与えがちである、c) ユーザは、エージェントの学習過程の観察に基づいて報酬の与え方を変化させていく、といった点を報告している。これらは、人間を学習系に組み込む際には、人間固有の特性を考慮してアルゴリズムを設計していかなくてはならないことを示す好例といえる。

スライダ入力によるベイズ的最適化

大量の訓練データを人手で作成する手法として、マイクロタスク型クラウドソーシングが広く活用されている。マイクロタスク型クラウドソーシングでは、クラウドワーカーが単純な仕事を短時間で行うことを前提としているので、それに適したタスクデザイン(= インタクションデザイン)が必要である。たとえば、画像データの良し悪しを判断するといった場合に、データに対して数値的スコアを直接与えることはばらつきが大きく現実的

☆4 <https://dl.acm.org/citation.cfm?id=1342790>



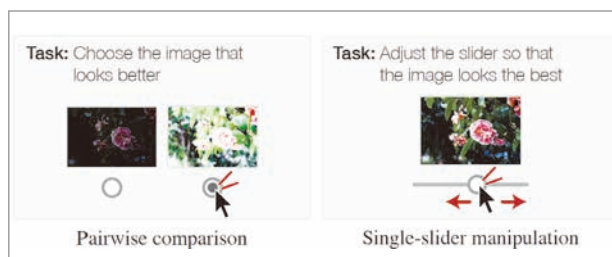
■ 図-3 Teachable robots. ロボットの行動に対して、ユーザが緑色のスライダで報酬を与える

でない。そのため、実用的な手法として、2つのデータ(画像)を見せて、どちらかを選ばせるという一対比較による方法がよく使われている(図-4左)。しかし一対比較では、1応答から得られる情報の量が限られている。Koyamaらは、一対比較に代わる方法として、2つの画像の中間をスライダーで自由に選べるようにすることで、より詳細な情報を1応答から得るという手法を提案している(図-4右)^{☆5}。提案手法を、バイズ的最適化による画像調整に適用し、一対比較よりも効率的に解が得られることが示されている。本研究は、データ作成プロセスにおけるインタラクションデザインの重要性を示す好例であるといえる。

訓練プロセスの制御に関するインタラクション

一般的な機械学習応用システムにおいては、訓練プロセスはブラックボックスとして動作する。ユーザがデータを与えると、与えられたロジックに従って訓練を行い、モデルが生成される。もしモデルの動作が期待通りでない場合には、データや訓練ロジックを変更して再び訓練を行う。しかし、このようなプロセスでは訓練結果を確認するまでに時間がかかりすぎるために、特に、訓練に基づくシステムのデザインにおける初期段階、どのようなデータやどのような訓練ロジックを使えばよいのかははっきりしていない場合には問題が多い。ここでは、訓練プロセスに対してインタラクションを加えること

☆5 <https://dl.acm.org/citation.cfm?id=3073598>



■ 図-4 Sequential line search. 従来法(左)が選択式だったのに対して、提案手法(右)ではスライダーを利用している

で、「訓練プロセスのブラックボックス性」という問題を解決しようとする試みについていくつか紹介する。

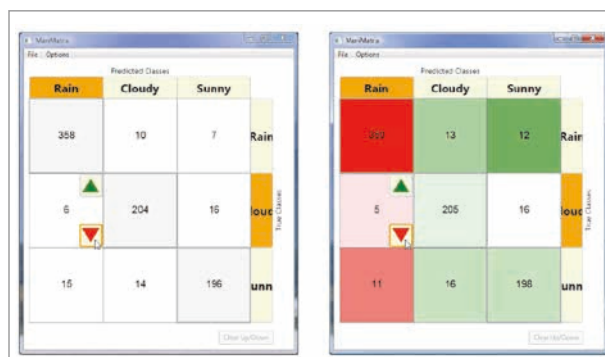
混同行列の直接操作

機械学習においては、訓練データとテストデータの間には常に差があり、常に完全な答えを返すことはできない。よって、どのような間違いをどの程度許容するか、についてユーザが判断し、期待する結果が得られるように訓練プロセスを制御することが必要となる。しかし、この訓練プロセスの制御は複雑なパラメータの調整を必要とすることが多く、ユーザには困難である。Kapoorらは、このような分類器の内部パラメータの調整を、混同行列を直接操作することで実現する手法 ManiMatrix を提案している(図-5)^{☆6}。混同行列とは、どのラベルがどのラベルとして認識されたか、を行列として表現したものである。ユーザがこの値を変更すると、システムは認識結果がその変更にあうように、内部のパラメータの自動調整を行う。このようなユーザインタフェースを利用することで、訓練アルゴリズムについて深い知見のないユーザでも適切に訓練プロセスを制御することが可能となる。

対話的アンサンブル学習

機械学習によるデータの分類を行う場合に、単一の分類器を使うのではなく複数の分類器を組み合わせ

☆6 <https://dl.acm.org/citation.cfm?id=1753529>



■ 図-5 ManiMatrix. 混同行列の値を直接大きくしたり小さくしたりすることができる

使うこと（アンサンブル学習）が有効であることが多い。しかし、複数の分類器をどのように組み合わせるかを定めることは専門家にとっても困難な作業である。Talbotらは、ユーザが複数の分類器をどのように組み合わせるかを検討するためのインタラクティブな手法 EnsembleMatrix を提案している(図-6)^{☆7}。ユーザはグラフィカルなウィジェット(図-6 右上)を使って各分類器の重みづけを変更しつつ、それが分類結果にどのような影響を及ぼすかを混同行列上で確認することができる。また、自由に分類プロセスを分割することによって、それぞれの分類器の強みをいかした分類を行うようにすることができる。このようにして、ユーザの直感や判断を制御プロセスに取り込むことによって、全自動での調整と遜色ない結果が得られることが示されている。ユーザの判断を入れられることは、説明責任の点からも重要であると考えられる。

可視化による距離の対話的定義

クラスタリングや分類といったさまざまな機械学習タスクを適用するためには、「データ要素間の距離をどのように計算するか」をユーザがあらかじめ定義してシステムに与える必要がある。しかし、適切に距離を定義することは機械学習の専門家でないユーザにとっては困難である。Brownらは、このような問題を解決する手法として、可視化されたデータを直接操

作することで距離を定義する手法 Dis-Function を提案している^{☆8}。まず、システムは、デフォルトの距離定義を用いてデータを散布図上にプロットする。次に、ユーザは、データ間の距離関係が望ましい距離関係になるよう、データ要素を散布図上で動かす(図-7左)。最後に、システムはユーザの動かした結果になるべく近くなるように距離定義を更新して提示する(図-7右)。ワインデータセットにノイズを加えたものをk近傍法で分類するタスクを用いてユーザテストを行い、本手法を用いて作成した距離定義によって、デフォルトの距離定義による分類精度を改善できることが示されている。

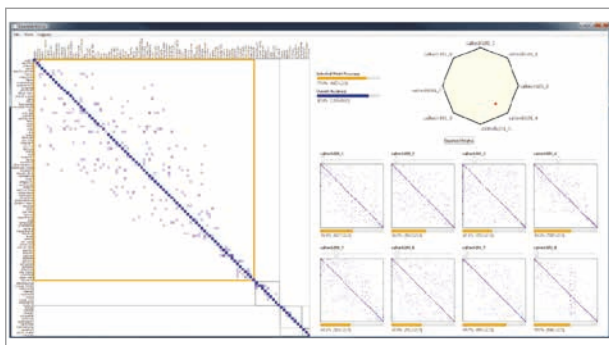
画像認識アプリケーションの開発支援

Maynes-Aminzadeらは、ユーザが機械学習に基づく画像認識を利用したアプリケーションを簡単に作成できるようなツール EyePatch を提案している(図-8)^{☆9}。ユーザがビデオカメラからの入力に対してアノテーションを与えると、即座に計算機が訓練を行い、その結果をカメラ入力に適用して提示する。さらに、システムには複数の機械学習アルゴリズムが実装されており、それぞれの現在の認識結果、およびその内部状態が常に可視化されている。このような対話的な環境を与えることで、機械学習についての専門的な知識のないユーザでも、機械学習を利用したアプリケーションを簡単に作

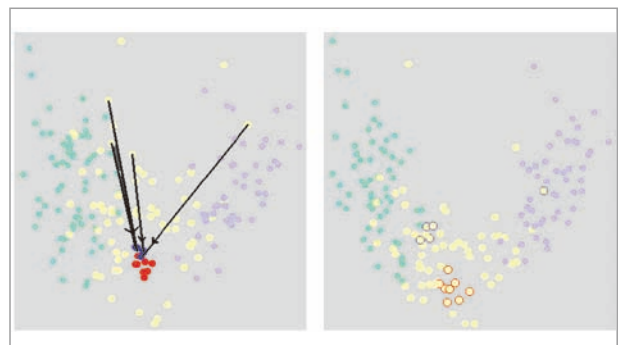
☆7 <https://dl.acm.org/citation.cfm?id=1518895>

☆8 <https://ieeexplore.ieee.org/abstract/document/6400486/>

☆9 <https://dl.acm.org/citation.cfm?id=1294219>



■ 図-6 EnsembleMatrix. 複数の分類器の組合せの重みを対話的に調整することができる



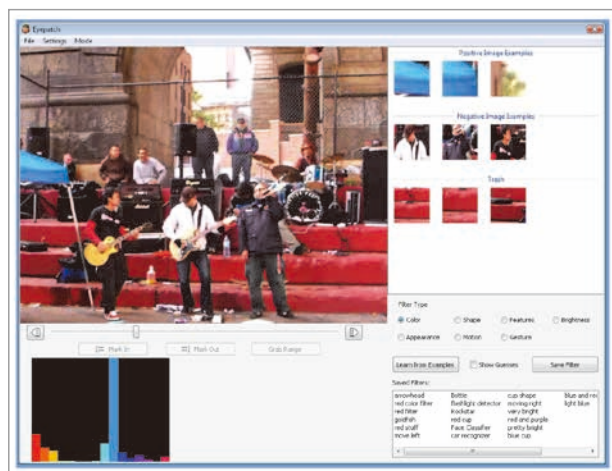
■ 図-7 Dis-Function. ユーザがサンプルの位置を移動すると(左), その距離関係を反映するように距離の定義を更新する(右)

成することが可能となっている。実際に、大学の演習で提案システムを利用し、さまざまなアプリケーションを構築できることを示している。さらに、ユーザが最初に訓練データをまとめて選んで与えた場合と、学習結果を確認しながら徐々に訓練データを選んで与えていく場合を比較し、同じデータ数であっても後者の方が質の高い認識器を得られることが報告されている。これは、学習結果を観察することで、ユーザが「どのような訓練データを与えるのが最も効果的か」を学習していった結果、より良い訓練データを選ぶことができたためと考えられる。

対話的トピックモデリング

トピックモデリングとは、文書中の単語の出現頻度をもとに、文書の主題（トピック）を自動的に抽出する機械学習手法である。訓練の結果、各主題と単語の関連や、文書と主題の関連などが確率分布として与えられる。Smithらは、このような訓練プロセスにユーザが介入し、主題と単語の関係や、主題と文書の間を直接操作して編集できるようなシステムを提案している¹⁰。図-9に画面例を示す。左側に主題の一覧が、右側に現在選択されている主題に関連した単語の

¹⁰ <https://dl.acm.org/citation.cfm?id=3172965>



■ 図-8 EyePatch. 左上が入力画像。右上が与えた訓練データ。左下が識別器の内部状態を可視化したもの

リストと文書のリストが示されている。×アイコンをクリックすることで単語や文書を主題から取り除く、といったように GUI で簡単かつ直感的にモデルの中身を編集することができる。編集システムの側では、ユーザの編集結果に合うように内部の確率モデルを適切に更新する。システムの設計にあたっては、予測可能性、信頼性、制御可能性に対する配慮が重要であるとしており、また履歴機能(undo/redo)の重要性を指摘している。

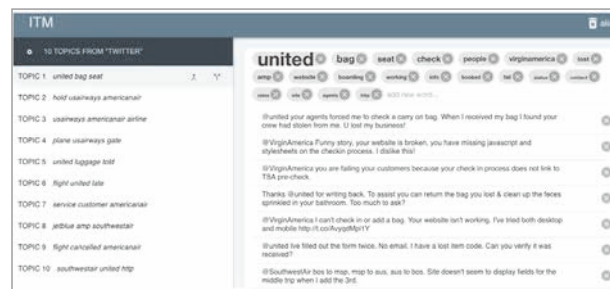
訓練結果の利用時におけるインタラクション

ここでは、訓練が終わった後で、実際にエンドユーザが機械学習に基づくアプリケーションを利用する場面でのインタラクションについて議論する。

不確実性への対処

機械学習を利用したシステムは、人間が人手で構築したロジックに従って動くのではなく、あくまで訓練データから帰納的に「推論」されたロジックで動く。そのため、訓練データにないデータが与えられたときの動作が人間にとって想定通りであるということは保証されない。したがって、機械学習を用いたアプリケーションにおいては、そのような本質的な不確実性を考慮したインタフェースデザインとする必要がある。

たとえば、現状の自動運転車においては、完全にシステムに運転を任せるといったことはまだ実現されてお



■ 図-9 Human-in-the-Loop Topic Modeling. 左がトピック。右が選択されたトピックに対応するキーワードと文書

らず、システムが対応できない状態になったときには運転者が適切に対応することが求められている。しかし、いきなり対応を求められても反応が難しい可能性がある。そのため、「運転者の対応が必要になりそう」な状況になってきたらあらかじめ運転者の注意を喚起し、心の準備をするように促す手法が提案されている^{☆11}。

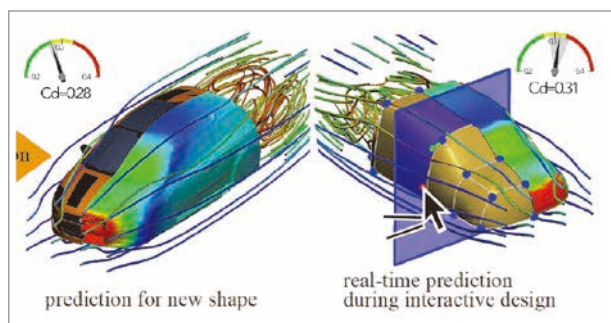
別の例として、自動車のデザインを機械学習によって支援するシステムにおいて、機械学習の推論結果の確信度を表示するといったことが行われている(図-10)^{☆12}。これは、流体シミュレーションを大量に走らせた結果をあらかじめ学習しておいて、それをもとにして、与えられた3次元形状の周囲の流体の流れを即座に計算して提示する、といったものである。しかし、与えられた3次元形状が学習時の形状と大幅に異なる場合には、計算結果が不正確である可能性が高くなるので、その点をユーザに伝えることは、適切な意思決定のために重要である。

利用時における訓練

通常の明示的なプログラミングによるシステムでは、エンドユーザが利用中に動作ロジックを更新することは困難である。しかし、機械学習を利用したシステムでは、ユーザからのフィードバックに基づいて機械がさらに訓練を行い、動作を改善していくことが可能である。すでに説明した対話的機械学習は、システム構築と利用が混在しているものであり、利用時における訓練を行っ

☆11 <https://dl.acm.org/citation.cfm?id=2516554>

☆12 <https://dl.acm.org/citation.cfm?id=3201325>



■ 図-10 Aerodynamic Design. 学習に基づき Cd 値(空気抵抗係数)を推定する。メータの針の左右のグレーが確信度を示す

ているものと見ることができる。

このように「利用時に訓練して動作を更新する」場合に注意しなければならない点として、ユーザ側の慣れや学習といった側面がある。ユーザはすでにこれまでシステムを利用してきた経験から、「このようなときにはこういう結果を返す」というメンタルモデルができてきている。ところが、利用中にシステムの動作が変わってしまうと、このメンタルモデルからはずれてしまってエラーを起こしたり余計な時間がかかったりする危険がある。例として、仮名漢字変換の候補の順が急に変わると使いにくくなるといったようなものが挙げられる。システムが複雑になると、ユーザ動作とシステム動作の更新の因果関係が見えにくくなると考えられ、より一層の配慮が必要と考えられる。

生成的システムの対話的な制御

画像生成や文章生成といった生成的なアプリケーションにおいては、機械学習応用システムにより出力された結果が想定通りでなかった場合にはユーザが対話的に結果を修正できることが望ましい。そのようなことを実現する方法として「ユーザからの対話的な入力に対してどのような修正を行って応答を返すか」についての訓練をあらかじめ行っておく、といった方法がある。

例としてスケッチによる画像の対話的修正を挙げる。深層学習を利用した画像生成は、敵対的学習(GAN)の導入により大きく発展してきており、自然な写真と区別のつかないような画像の生成が可能となってきている。しかし、元々の敵対的学習による画像生成のプロセスはノイズを与えると何らかの画像ができるというもので、出てきた画像に問題があった場合に訂正するような処理はサポートされていなかった。Portnievらは、出力された画像に対して希望する修正内容をユーザがスケッチによって与えることによって、画像を対話的に編集するシステムを提案している(図-11)^{☆13}。訓練時には、画像から自動的に手書きスケッチに相当する線

☆13 <https://dl.acm.org/citation.cfm?id=3201393>

画を抽出し、それをもとに訓練を行っている。スケッチだけでなく、色を指定することも可能となっている。

同じようなアプローチで、対話的に線画に色付けをしたり^{☆14}、対話的にラフスケッチを清書したりする手法が提案されてきている^{☆15}。

アルゴリズムからインタラクションへ

本稿では、機械学習のためのヒューマンインタフェースとして、訓練データの作成に関するもの、および訓練プロセスの制御に関するもの、および訓練結果の利用時に関するものについて紹介した。機械学習応用システムの研究においては、学習アルゴリズムの改良に重点が置かれることが多く、訓練データはすでに与えられたものとして扱い、また訓練プロセスおよび推論プロセスは完全自動のバッチ処理として扱うことが通例である。しかし、実際に機械学習技術を具体的な問題に適用して利用するのは現場のユーザであり、これらの現場のユーザが、機械学習応用システムを用いて訓練を行ったり、訓練結果を利用したりするプロセスを

支援することは非常に重要である。今後機械学習技術が研究から実用へと移ってくるにつれて、ヒューマンインタラクションに関する技術の重要性が増してくるものと考えている。

なお、近年急速に発展している深層学習については、膨大な訓練データを必要とすること、訓練に非常に時間がかかること、および内部の動作が複雑で直感的でないことから、機械学習技術の中でも特にユーザの理解や制御が難しいものとなっている。深層学習の内部動作の理解を補助しようとする試みとして、ノードの活性化状態を可視化するものや^{☆16}、画像中のどこにシステムが注目しているかを可視化する手法^{☆17}、得られた結果の根拠となっている訓練データを特定する手法^{☆18}などが提案されてきている。しかし、ユーザが訓練済みモデルが望む動作をするように対話的に修正する（前章で紹介したようなシステムの「出力」の修正でなく「訓練済みモデル自体」の修正を行う）といったことはいまだ実現できておらず、今後の発展が期待される分野であるといえる。

☆14 <https://github.com/pfnet/PaintsChainer>

☆15 <https://dl.acm.org/citation.cfm?id=3201370>



■ 図-11 FaceShop. システムによる生成結果（左上）に対してユーザが希望する結果を指定すると（左下）、新しい結果が示される（右）

参考文献

- 1) Sacha, D., et al. : Human-Centered Machine Learning through Interactive Visualization : Review and Open Challenges, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, pp.641-646 (2016).
- 2) Amershi, S., Cakmak, M., Knox, W. B. and Kulesza, T. : Power to the People : The Role of Humans in Interactive Machine Learning, AI Magazine, 35 (4) , pp.105-120 (2014). (2018年9月3日受付)

☆16 https://link.springer.com/chapter/10.1007/978-3-319-10590-1_53

☆17 <https://ieeexplore.ieee.org/abstract/document/8237336/>

☆18 <http://proceedings.mlr.press/v70/koh17a.html>

■ 五十嵐健夫（正会員） takeo@acm.org

2000年東京大学工学系研究科 博士（工学）。2002年同大学情報理工学系研究科講師，2005年助教授，2011年教授。2007～13年JST ERATO 研究総括。2017年よりJST CREST「データ駆動型知的情報システムの理解・制御のためのインタラクション」研究代表者。学術振興会賞，SIGGRAPH 若手科学者賞等受賞。