

研究員 の眼

AI の時代、どこから先を信じるか ブラックボックスにどう対処する？

保険研究部 主席研究員 篠原 拓也
(03)3512-1823 tshino@nli-research.co.jp

いま世の中では、AI(人工知能)を活用した技術開発が急速に進んでいる。たとえば、AI を用いて、小売り業で消費者のニーズ変化を的確に捉えて商品の品ぞろえを変えたり、鉱工業で物流を最適化して製品の生産効率を高めたりする取り組みが行われている。

情報サービス分野でも、AI の活用が進んでいる。一例として、AI を搭載したチャットツールがある。前もって顧客からの質問とその回答のシナリオを設定しておき、該当するシナリオに最も近い質問をAI が自動で判断する。シナリオを豊富に用意しておけば、まるで人間とやり取りしているような対応が可能となる。

その他にも、安全で快適な自動運転、利便性の高い金融システム、未病を検知する効果的な健康管理など、人々の生活の質の向上を図る観点でも、さまざまなAI の活用が模索されている。

◆ 「ある程度から先は、AI の出した答えを信じるしかない」

AI の活用には、ネットやクラウドを介したビッグデータの解析処理が伴う。その処理は、短時間で大量の計算処理を行うもので、人間の理解が完全には及ばないことが一般的だ。

このため、人間は「ある程度から先は、AI の出した答えを信じるしかない」状態となる。

AI がはじき出した結果を目にする典型的な機会として、テレビやネットでの将棋の実況中継がある。対局がある盤面まで進んだ状態で、次の一手をどのように指すべきか、最善手を見極めたい。

こんなときに、AI はその盤面から先の手を、何億通りも計算して評価する。そして、得られた最善

手を画面上に示す。ところが、ときとして、示された手がどうして最善といえるのか、解説者であるプロの棋士にも理解されないことがある。

◆ 人間の側の AI への接し方が問われる

また、AI の出す答えは道義的に正しいとは限らない。AI は、従来の機械のように、倫理などとは無関係に、ただ安定的に動作して、信頼できるアウトプットを出せばよいというところにとどまらない。

たとえば、ある会社が新入職員のリクルート活動に AI を活用するテストを行った。AI を使って、人材の能力や経験を適切に判定して、採用を行おうという試みだ。ところが、AI は特定の人種や性別に偏った採用可否の判定結果を出してきた。この会社は、この判定結果を用いなかったという。

これは、AI が機械学習をする際に用いたデータに、人間が持つバイアスが反映されていたことに起因する。AI が機械学習によって確立し、さまざまな問題の解決に用いるアルゴリズムには、偏見や差別を含んだり、既存の社会ヒエラルキーを強化したりする可能性が潜んでいるわけだ。

こうしたことを踏まえると、今後は、人間の側の AI への接し方が問われることになる。ただ、やみくもに AI の出した答えを信じるのではなく、AI システムが倫理や道義の面も含めて、フェアに稼働していることをよく確認したうえで、その先を信じるという姿勢が求められることになる。

◆ ブラックボックスはすでに身近にある

AI が登場する前から、人々は機械や技術の核心部分をよく理解せずに、信じて使用してきた。

たとえば、身近にある電卓がそうだ。電卓が出す計算結果をいちいち疑って、暗算や筆算で確認する、という人はいないだろう。これでは、何のために電卓を使っているのかわからなくなってしまう。

だが、電卓が正しく計算する仕組みを完全に理解しているという人は限られるはずだ。たいていの人は、電卓の仕組みをよくわからないまま使っている。多くの人にとって、電卓はブラックボックスといえるだろう。

これは、同じ計算ツールである、算盤(そろばん)とはまったく事情が異なる。算盤は、使う人がその都度正しさを確認しながら、計算を進めていく。一方、電卓は、数字や加減乗除などのボタンを順番に押しているだけで、計算そのものの正しさの確認を意識しなくても答えが出る。

こうしたブラックボックスは、電卓だけでなく、世の中にどんどん広がっている。スマートフォン、パソコンはおろか、自動車もテレビも、故障したときに自力で直せるという人は、普通の人ではまずいない。壊れたときには、購入した店の修理カウンターや、修理工場に持ち込んで直してもらるのが一般的だ。

◆ 「ブラックボックステスト」には限界がある

コンピューターシステムでは、ブラックボックスが正しく動作するかどうかを確かめるために、「ブラックボックステスト」と呼ばれる方法がとられる。これは、ブラックボックスの中身を理解せずに、ブラックボックスへのインプットと、そこからのアウトプットをもとに、テストをするというものだ。

たとえば、適当な年月日をインプットしたときに、実際にそういう日があるかどうかを判定するシステムを考えてみよう。このシステムは、「2022年6月28日」や「2022年10月31日」といった実際にある日のインプットには“Yes”。「2021年9月31日」や「2023年4月31日」といったありえない日のインプットには“No”をアウトプットする。

ブラックボックステストでは、ありうる年月日と、ありえない年月日をつぎつぎにインプットしてみ、アウトプットが正しいかどうかを確認する。

ブラックボックステストは、プログラムの利用者目線からのテストといえる。テストの内容が限られるため、システムにそれほど詳しくない人でも、短時間でテストが実行できるという利点がある。

ただし、ブラックボックステストをどれだけ重ねても、システムが絶対に正しいとは言い切れない。インプットしなかったデータで、なんらかの誤動作を起こす可能性がぬぐい切れないからだ。ブラックボックステストは、計算ロジックを検証しないため、どこまで行っても不十分なのである。

◆ 「ホワイトボックステスト」は負担が大きい

これとは対照的に、「ホワイトボックステスト」というものもある。インプットとアウトプットだけではなく、システムの中身を確認するものだ。プログラムそのものを確認したり、プログラムの条件分岐を通過するインプットデータを網羅的に流してアウトプットを検証したりする。

先ほどの年月日判定システムのテストの例でいうと、「うるう年の対応ができているかどうか」といった確認が該当する。まず、うるう年を判定するプログラムが適切に作られているかどうかを確認する。そして、実際にうるう年の判定が正しくできるかどうか、「〇〇年2月29日」のデータを集中的

にインプットして確かめてみる。たとえば、「2020 年 2 月 29 日」、「2021 年 2 月 29 日」、「2022 年 2 月 29 日」を入れてみる。うるう年である「2020 年 2 月 29 日」には“Yes”。うるう年ではない「2021 年 2 月 29 日」や「2022 年 2 月 29 日」には“No”がアウトプットされることを確認する。

さらに、100 年に一度や、400 年に一度のうるう年の例外を含めた、グレゴリウス暦(*)に対応しているかどうかを確認する。たとえば、2100 年は 100 年に一度の例外でうるう年ではない。そこで、「2100 年 2 月 29 日」をインプットして、アウトプットが“No”となることを確認する。また、2000 年は 400 年に一度の例外でうるう年であった。そこで、「2000 年 2 月 29 日」をインプットして“Yes”と、正しいアウトプットが出力されるか確かめる。こうした確認も「ホワイトボックステスト」に含まれる。

(*) 現在、一般に用いられている太陽暦。西暦の年数が、100 で割り切れるが 400 では割り切れない年は、平年。それ以外で西暦年数が 4 で割り切れる年はうるう年、というルール。

ホワイトボックステストは、プログラムの開発者目線からのテストといえる。システムの内容を熟知して網羅的に行われるテストであるため、基礎となる知識の習得や、テストの手間、一定の時間が必要となる。特に、複雑なシステムをテストする場合には、こうした負担が大きくなる。

◆ 「グレーボックステスト」はテストの中身が重要

そこで、実務上は、この両者を組み合わせた「グレーボックステスト」が行われることが一般的だ。システムの内容をみるために、システムやプログラミングに精通した人がテストを行う。ただし、テストに用いるインプットデータは、必ずしも網羅的であることにこだわらない。テストを行う人が重要と思われるものだけを重点的に流す。こうすることで、テストの負担をある程度抑制しながら、テストの品質を高めることを目指す。

グレーボックステストは、「システムが正当なのか誤っているのか、白黒をはっきりつけたい」という人にとっては満足のいくものではないかもしれない。こういう人は、「自分の頭で理解できたものだけを信じる。少しでもわからない部分が残るものは信じない。」という考え方を持ちやすい。

しかし、AI では、こうした考え方では立ち行かない場合が多い。どうしても人間の理解が及ばない領域が残ってしまう。現実には、さまざまなシステムの開発や稼働場面で、グレーボックステストが行われている。どのようなグレーボックステストを行うか、その中身が問題となる。

AI の時代には、従来以上にブラックボックスとうまく付き合っていかななくてはならない。AI の進化とともに、人間の側も、AI を上手に信じる感性を研ぎ澄ましていく必要があると思われるが、いかがだろうか。

(本稿の執筆にあたり、参考にさせていただいたウェブサイト)

「ホワイトボックステストとは？ブラックボックステストとの違いやテスト技法について解説」

花房輝鑑氏監修(次世代ゲームテスト研究所, AIQVE ONE 株式会社(アイキューブワン), 2021 年 2 月 12 日)

<https://blog.aiqveone.co.jp/whitebox-testing/>

「ホワイトボックステストとブラックボックステスト、どっちが必要？」

(発注ラウンジ, 発注ナビ株式会社, 2017 年 7 月 6 日)

<https://hnavi.co.jp/knowledge/blog/white-box-test/>