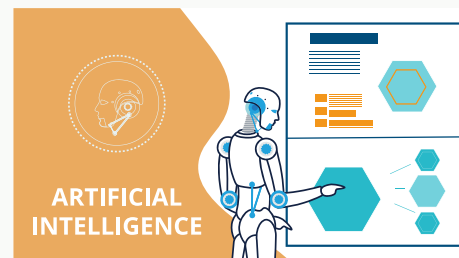


求められる XAI（説明可能な AI）

XAI（説明可能な AI）が企業に必要

- ・なぜそうなったの？
- ・なぜ他じゃダメなの？
- ・いつになったら成功するの？
- ・いつ失敗するの？
- ・いつになったら信頼できるの？
- ・どうやったらエラーを修正できるの？

XAI のゴールは、ハイレベルのパフォーマンスを保ったまま、説明できるモデルを作り、人が理解して、信頼でき、効率的に管理ができるようにすること。



ディープラーニングを中心に大量なデータを利用して、AI によって今まで人の判断では不可能であった高度な未来予測が可能になると、到底人間では、その予測結果の根拠を説明することができません。

AI が、活躍すればするほど AI に対する依存度が高まります。DeepMind 社の AlphaGo と囲碁のプロ棋士との戦いは有名ですが、誰一人予想もしない手を AI が選択したことについて、もはや人間では説明できないとまで言われています。

今、企業では、人の採用や、医療の診断などあらゆる分野に AI が進出していますが、すでにそのことが、AI による判断に人の運命まで影響を与えるほどの重要なポジションを、AI に依存してしまっているという現状があります。

たとえば、大企業の採用の 1 次審査で、書類選考を AI に任せると、年齢や性別など何を優先して選考しているかわからず、優秀な人を足切りしている可能性もあります。医療では、医師と AI の診断結果が異なる場合に、なぜ違うのかが解明できなくて困ってしまい、有効な治療に進めないというリスクもあります。

ディープラーニングに代表される機械学習では、必ずしも教師データを与える手法ばかりではなく、AI が自立して学習することさえ可能であるため、こういった特徴のデータがどれくらいの影響を与えているのかわからないという判断基準がブラックボックス化されてしまう問題があります。

そのために、近年世界中で XAI（Explainable AI）という説明可能な AI の研究がなされ、多くの論文も発表され、いろいろな手法が開発されています。すでに、各企業によって、実用的なソリューションとして展開されてサービス化されている例も多くなってきました。

米国防高等研究計画局 (DARPA) の XAI の研究は有名ですが、日本においても、内閣府の「人間中心の AI 社会原則検討会議」において、「公平性、説明責任及び透明性の原則」にて、「AI を利用しているという事実、AI に利用されるデータの取得方法や使用方法、AI の動作結果の適切性を担保する仕組みなど、用途や状況に応じた適切な説明が得られなければならない」としています。これにより、AI を利用した企業は、その処理の決定過程の説明責任が求められるようになり、今後、業務自動化に AI を組み込む場合などに考慮が必要になってきます。

そもそも、人間に説明が出来ない AI の内部処理なので、AI が何を考え、どうしてそのような結果を出したのかを解析するための別の AI を設置することが、今後流行っていくのかもしれない。