

UBIC、人工知能の成長を可視化する コンサルティングサービスを提供開始

テキスト解析の状態を視覚的に把握し、効率的に精度を向上

米ナスダックと東証マザーズに上場し、人工知能を駆使したビッグデータ解析事業を手がけるUBIC（本社：東京都港区、代表取締役社長：守本正宏）は、人工知能の成長を可視化することに成功、人工知能の成長過程を視覚的に把握し、テキスト解析の精度を効率的に向上するコンサルティングサービスの提供を開始すると発表しました。

UBICは、人工知能によるEメール自動監査システム「Lit i View EMAIL AUDITOR（以下、EMAIL AUDITOR）」により、企業内の大量のメールを自動的に監査し、企業のお客様が防ぎたいリスクを発見するソリューションを提供しています。

「EMAIL AUDITOR」に搭載された人工知能「バーチャルデータサイエンティスト（以下、VDS）」は、豊富な経験をもつ監査人（社内外の監査スタッフや弁護士など）から、不正を判断する特徴を学び、監査人がすべてチェックすることが不可能な大量の電子メールのテキストから、情報漏洩やカルテルなど不正行為への関与の疑いがあるコミュニケーションや行動を、監査人に代わって見つけ出します。

成長を可視化するコンサルティングサービスは、「EMAIL AUDITOR」の導入を始めたお客様より“人工知能によるメール監査がどの程度機能しているのか？”、“運用開始からどれぐらいの期間で実用的な監査ツールとして稼働できるのか？”などを知りたいという要望に応えるために開発されました。

従来は表の一覧で示していた解析結果を、グラフ上でのメール文書の分布、時系列での推移を見ることで、より分かりやすく確認できるようになりました。今後、「EMAIL AUDITOR」を導入されるお客様に対するコンサルティングサービスとして提供します。

成長を測定するポイントは、監査人が「不正に“関連する”と判断した文書」と「不正に“関連していない”と判断した文書」が、人工知能によって正しくスコア付けされているかどうか、で判断します。監査人が関連性の有無を判断したメール文書が全体のどの位置に分布し、時系列でスコア付けがどう変わるかという動きを追うことで、人工知能の成長度を測ることができます。

人工知能の成長プロセスは、1. 成長初期 2. 成長期 3. 成熟期 の3つの段階に分けることができます。成熟期に到達すると、安定したメール監査機能として稼働が確認できたこととなります。

成長段階における分布の変化（典型的なサンプルと実際の案件での文書分布比較）

下記の図は、人工知能の各成長段階における典型的なサンプルと、ある実際の案件におけるメール文書の分布を並べたものです。典型的なサンプルと実際の案件を比較し、同様の傾向があらわれているか見ることで、実際の案件において人工知能が正しく稼働しているかどうか確認できます。グラフの横軸には最新のスコア付け、縦軸には1つ前の解析時のスコア付けを取ることで時系列での変化をみることができます。

図1：成長初期

運用初期の段階で典型的なサンプルは、監査人が「不正に関連する」と判断したメールであっても、人工知能は低いスコアを付けるなど、監査人の判断と人工知能のスコアに乖離があり、全体的にスコアのばらつきが発生します。実際の案件（運用開始時→10日目）でも同様にばらつきが発生しており、判断の学習が十分でない状態が見られています。

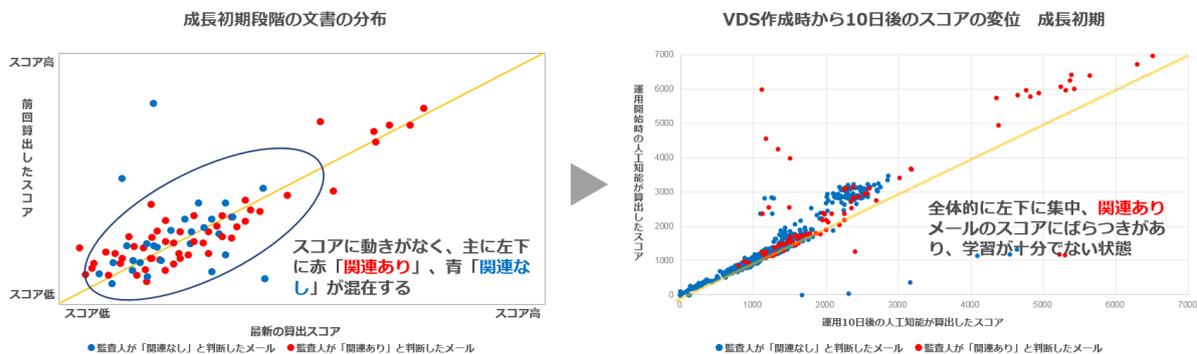


図2：成長期

学習期間を経ると、監査人が不正に関連すると判断したメールについて、高いスコアが付く件数が増加するため、典型的なサンプルでは、「関連あり」の分布は「過去のスコアが低く、現在のスコアが高い」ことを意味する右下のゾーンに移動します。また逆に、「関連なし」の分布が「過去のスコアが高く、現在のスコアが低い」ゾーンに移動する傾向も発生します。実際の案件（17日後→24日後）では、関連ありのスコアの値が高まっており、人工知能が適切に成長していることが分かります。

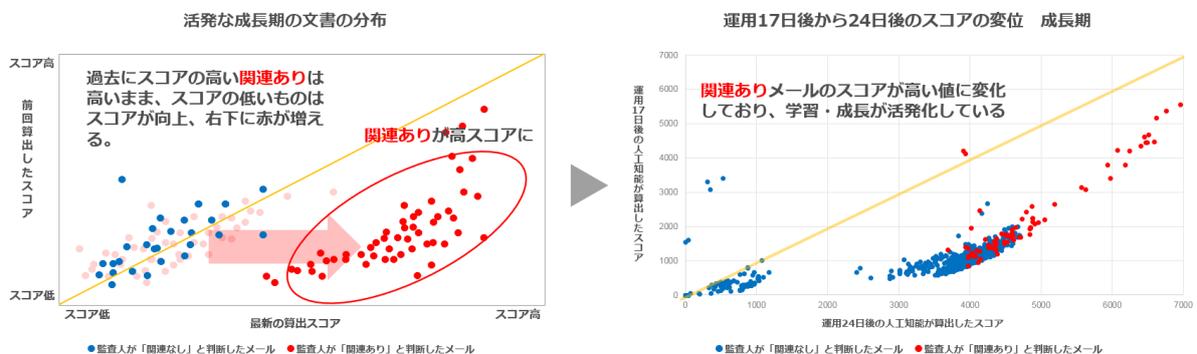
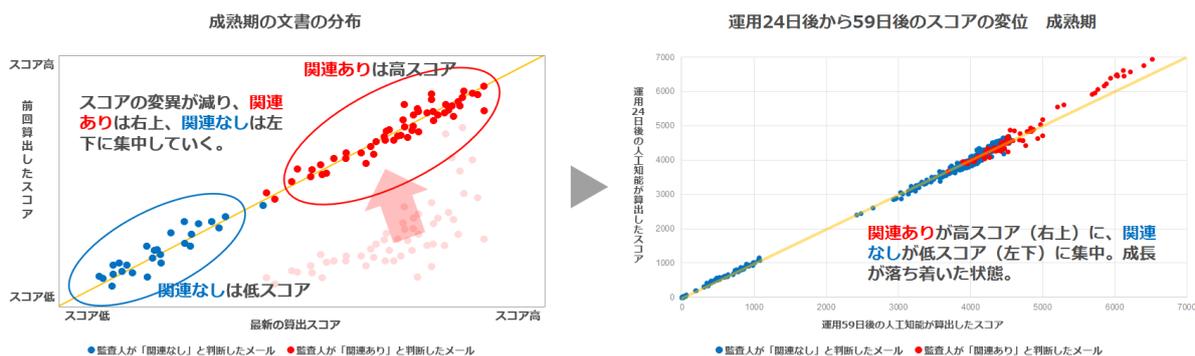


図3：成熟期

さらに学習期間を経ると、典型的なモデルと実際の案件（同 24日後→59日後）のどちらも、監査人が関連すると判断したメールは高いスコアに集中し、関連なしと判断したメールについては低いスコアに集中します。グラフ中の成長基準線付近に集中した状態は人工知能の学習が理想的に進んだことを意味し、人工知能が成熟した状態を表しています。



注) 成長の日数は1つの例であり、案件ごとによって異なります。

UBICでは、高度情報解析課においてVDSの学習機能が適切に働いているか、分析コンサルティングを行っています。今回の可視化の手法は、UBICの高度情報解析課の彦根佑紀主任による分析作業の中で発見された傾向や手法を元に開発されています。可視化の手法を用いることによって、従来よりも、より直感的に人工知能の成長過程を把握することが可能になり、成長の精度やスピードが適正であるかどうかを容易に判断することができます。

「VDSは、監査人の判断が追加される度に、その判断の特徴を追加的に学習し、学習した特徴がより強く表れるようなスコアへと最適化していきます。図のように、学習後の最適化スコア再付与結果から、関連文書のスコアと関連しない文書のスコアとの間の相対差はより拡大しており、VDSが監査人の判断ポイントに沿ってこれらの文書間の差異をより明確に学習していることが分かります。」(UBIC 行動情報科学研究所 シニア・フェロー 蓮子 和巳)

UBICでは、今後も人工知能による分析の可視化手法の開発を通じて、より実用性の高い人工知能分析ソリューションの提供に取り組んでいきます。

【UBICについて】 URL: <http://www.ubic.co.jp/>

株式会社UBICは、独自開発の人工知能「バーチャルデータサイエンティスト (VDS)」により、ビッグデータなどの情報解析を支援するデータ解析企業です。国際訴訟などに必要な電子データの証拠保全と調査・分析を行うeディスカバリ(電子証拠開示)や、コンピュータフォレンジック調査を支援する企業として2003年に創業。自社開発のデータ解析プラットフォーム「Lit i View® (リット・アイ・ビュー)」、アジア言語に対応した「Predictive Coding® (プレディクティブ・コーディング)」技術などを駆使し、企業に訴訟対策支援を提供しています。訴訟対策支援で培った人工知能「VDS」は、専門家の経験や勘などの「暗黙知」を学び、人の思考の解析から、未来の行動の予測を実現、最近では医療やビジネスインテリジェンス、マーケティングなどの領域にVDSを活用し、事業の拡大を進めています。2003年8月8日設立。2007年6月26日東証マザーズ上場。2013年5月16日NASDAQ上場。資本金1,688,433千円(2015年3月31日現在)。

〈本件に関するお問合せ先〉
 株式会社UBIC 広報担当 池内・小越
 TEL: 03-5463-6380 FAX: 03-5463-6345