株式会社アナリティクスデザインラボ 様

テキストマイニング × AIで、効果的な業務コンサルティングを展開

産業技術総合研究所や大手コンサルティング会社でデータサイエンティストとして活躍し、現在はアナリティクスデザインラボで代表を務める野守 耕爾様は、テキストマイニングツール「Text Mining Studio」と複数のAI（人工知能）を組み合わせた新たな解析手法を開発。テキストマイニングを進化させるとともに、その解析データにより効果の高い業務コンサルティング活動を展開している。

株式会社アナリティクスデザインラボ
代表取締役 野守 耕爾 様

Interview

Text Mining Studioと2つのAI技術をマッチング

新開発の解析手法の概要を教えてください。

野 守 数理システムのテキストマイニングツール「Text Mining Studio」に2つのAI技術を組み合わせています。ひとつは、PLSA（Probabilistic Latent Semantic Analysis）と呼ばれる技術で、複雑なビッグデータをシンプルに解釈できるようにするクラスタリング技術です。これは、数理システムの汎用データマイニングツール「Visual Mining Studio」の二項ソフトクラスタリングとほぼ同じ機能です。もうひとつは、変数間の確率的な因果関係を探索するモデリング技術であるベイジアンネットワークで、数理システムのベイジアンネットワーク構築支援システム「BayoLink」を使用します。これら数理システムのツールは同一プラットフォームで連係して動作できるため操作性も良好です。この新しいテキスト解析手法は特許取得済みで（特許第6085888号）、「Nomolytics（Narrative Orchestration Modeling Analytics）」という名称で商標出願もしています。

Nomolyticsの概要

新しい解析手法のメリットを教えてください。

野 守 ①膨大なテキストデータを理解しやすくできること、②テキスト情報に潜む要因関係を可視化できること、③状況変化に応じたシミュレーションができることです。
テキストマイニングはテキスト（文章）に含まれる単語を抽出してテキスト全体の記述傾向の現状把握ができる技術ですが、膨大なテキストデータでは抽出される単語も膨大で大変複雑な解析結果となり、解釈が難しくなります。そこでその抽出された単語をPLSAでグルーピングし、使われ方が似ている単語を文脈を考慮しながらトピックというかたちに集約します。テキストマイニングで抽出される数千・数万単位の膨大な単語も、これによりわずか数十個程度のトピックにシンプル化され、容易に解釈できるようになります。
さらにベイジアンネットワークにより、抽出したトピックの周辺の確率的な因果関係を可視化し、どのような条件でどのような話題がされやすいのかなどテキスト情報に潜む要因関係を把握できます。加えてベイジアンネットワークでは確率シミュレーションができますので、現状把握だけでなく、与える条件を変えることによってどんな影響が起きるのか予測もでき、施策の効果を事前に確認することなどができます。

［裏面に続く］

PROFILE

野守 耕爾 様

2012年3月、早稲田大学大学院 創造理工学研究科経営システム工学専攻 博士課程修了。博士（工学）。2012年4月より（技術研修生としては2008年より）独立行政法人産業技術総合研究所 デジタルヒューマン工学研究センター入所。2012年12月よりデロイトトーマツグループ 有限責任監査法人トーマツ デロイトアナリティクス入所、データサイエンティストとしてビッグデータ解析に基づく業務コンサルティングおよび解析技術の研究開発に従事。2017年6月、株式会社アナリティクスデザインラボ設立。

株式会社
アナリティクスデザインラボ
http://www.analyticsdlab.co.jp
office@analyticsdlab.co.jp

現場の特徴が定量化される、気づきが生まれる

実際に、この解析手法はどのように活用できますか。

野 守　様々な業務のテキストデータに活用できます。
例えば、コールセンターの問い合わせデータに適用すれば、その問い合わせをいくつかのトピックに変換し、製品別や顧客属性別に問い合わせトピックの特徴を示すことで、製品の問題や顧客ニーズを把握することができます。データに「解約」や「満足度」などの情報があれば、ベイジアンネットワークにより、例えば製品や顧客の属性情報と問い合わせのトピックから解約確率をシミュレーションしたり、解約確率を抑制するための要因や複雑なルールも見つけることができます。
マーケティング分野でいえば、Web上の口コミデータへの適用もあります（数理システムユーザーコンファレンス2016にて発表）。温泉旅行の口コミデータへの適用事例では、観光客の属性別に関心のあるトピックと観光地域別に特徴的なトピックを定量的に示しました。それらを照合することで、各地域は誰にどのような価値を提供すべきかというマーケティング戦略を検討できます。また口コミの満足度に対する要因関係を可視化したところ、男性はカップルで行き温泉で温まることで、女性は女友達や家族などで行き砂湯で楽しむことで満足度が高まることが示されました。こうした結果からターゲットに応じた満足度の高いプロモーション施策を検討でき、経験豊富な業務担当者に新たな気づきを与えることができます。
また経営に直結する例として、特許文書への適用があります（数理システムユーザーコンファレンス2017にて発表）。特許技術の内容をトピック化しそのトレンドを把握したり、競合他社の技術動向を把握できますので、今後の技術開発の戦略や、企業M&Aなどによる競争戦略を検討できます。また特許文書内の「課題」と「解決手段」の記述項目からそれぞれ用途のトピックと技術のトピックを抽出し、それらの関係をベイジアンネットワークでモデル化することで、ある用途展開のために必要な技術や代替できる技術を把握したり、保有技術の新たな用途への展開などを検討することができます。

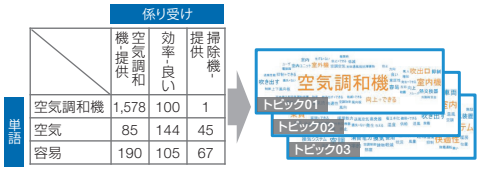
特許文書から用途と技術のトピックを抽出し、新規用途の探索などを行う例

用途と技術のトピック抽出

特許文書の「課題」と「解決手段」をテキストデータに使用
・「課題」から用途トピックを抽出
・「解決手段」から技術トピックを抽出

テキストマイニング　テキストマイニングを実行して単語と係り受け表現を抽出

PLSA　「単語×係り受け」の共起行列を作成し、これにPLSAを適用してトピックを抽出

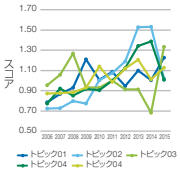


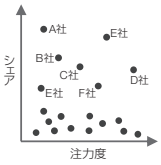
トピックの特徴集計

全特許データに対し各トピックのスコア（該当度）を計算

出願年集計
トピックスコアを出願年で集計してトピックのトレンドを把握

出願人集計
トピックスコアを出願人で集計して、各トピックにおける出願人の特徴を把握

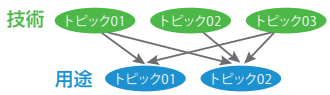


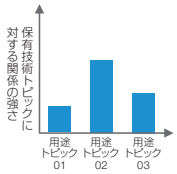


用途と技術の関係分析

用途トピックと技術トピックの統計的な関係性をベイジアンネットワークでモデル化

ベイジアンネットワーク





保有技術と関係のある用途トピックのうち、まだ想定していない用途を探索し、それに関連する元の特許文書を確認することで具体的な新規用途を検討

Nomolytics解析に興味をお持ちの企業様にメッセージをお願いします。

野 守　ビッグデータの時代となり、またAIなどそれを処理する技術も多く利用可能になってきましたが、これらを使いこなすのは容易ではありません。一番の目的は経営や業務の課題を解決することにあり、その手段としてデータ解析があるわけですが、その目的を達成するためにどのようなデータからどのようなアウトプットを出せばよいのか、そのためにはどのような技術を選択すればよいのかという設計力や構成力が求められます。NomolyticsはPLSAとベイジアンネットワークで構成していますが、この選択にはきちんとした意味があり、大量で複雑なテキストデータを人間がいかにシンプルに理解できるか、そして課題解決に重要な特徴をいかに顕在化させられるか、それを可能にする技術として開発しました。それぞれの技術の特徴を理解し、解決すべき課題に応じてそれを使い分けることが重要で、当然場合によってはPLSAやベイジアンネットワーク以外の手法が適していることもあり、それを判断するのはやはり人です。技術に固執せず、解決したい課題から適切なデータ解析・活用の設計を行い実行すること、これが私共の提供させていただいているコンサルティングサービスです。

野守さんの誠実な業務コンサルティングサービスをお薦めします。

近年、データ分析だけでなく、その結果に基づいた業務コンサルティングやソリューションを望まれるお客様が増えています。その際の有力なパートナーがアナリティクスデザインラボです。代表の野守様は学生の頃からのText Mining Studioのユーザーでもあり、近年はVisual Mining StudioやBayoLinkを用いた業務コンサルティングの実績も数多くありますので、実力と実績ともに当社が自信をもってお薦めできる心強いパートナーです。

株式会社NTTデータ数理システム
データマイニング部
グループリーダー 主任研究員
岩本 圭介

NTT DATA　株式会社 NTTデータ 数理システム　お問い合わせ：平日10:00-17:00（e-mail、FAXは随時受付）

Text Mining Studio担当

〒160-0016 東京都新宿区信濃町35番地 信濃町煉瓦館1階　TEL 03-3358-6681　FAX 03-3358-1727
〈e-mail〉tmstudio-info@msi.co.jp　〈URL〉http://www.msi.co.jp/tmstudio/