

東京海上研究所 様

社会調査のポイントを
テキストマイニングがあぶりだした

■ Theme (目的)

内閣府のムーンショット型研究開発制度において、台風の影響に関する社会調査を行う。

■ Point (内容・効果)

- 台風の影響について非物理的な被害まで含めて網羅的に記述した文献は少ないため、調査に必要な資料収集から着手した。
- 大量に収集したテキストからテキストマイニングにより情報を整理、それをもとに質問項目を設計し、アンケート調査を実施。
- 調査を行った結果の自由記載の回答の解析にもテキストマイニングを活用。

Interview

株式会社東京海上研究所様では、台風の影響に関する社会調査に**Text Mining Studio** (以下、**TMS**) を活用。大量のテキスト情報から過去の台風がもたらした影響を網羅的に捉え、それをもとに質問項目を設計し、アンケート調査を実施。台風の被害者が何に対して困難や不安を感じたのかを調査した。

2万の新聞記事、4千の景況調査から台風の影響を浮き彫りにする

台風の影響の実態解明に、テキストマイニングを活用されたそうですね。

加藤 当社は内閣府の研究プログラムに参画しており、その中で過去の台風の影響に関する調査をアンケート形式で行うことになりました。しかし、過去の被害やそれによる影響について、アンケートで確認したい項目を網羅的にまとめた文献が見つかりませんでした。そこで、過去に大きな被害を出した台風に関する文書を集め、被害や影響を類型化・整理し、それをもとにアンケートの質問項目をつくりこみました。

作業にあたって対象となる台風を3つ選定しました。2018年に関西国際空港の連絡橋にタンカーが衝突するなどした台風21号、2019年に千葉県で暴風により住宅の屋根を飛ばすなどの被害を出した台風15号、さらに同年、北陸新幹線の車両を水没させた台風19号です。それらに関する新聞記事を全国紙から地方紙、業界紙なども含め検索し抽出しました。また、内閣府や政府系金融機関等が公開している景況調査も調べました。そこには事業主コメントとして「台風により経営している旅館のキャンセルが増えた」といった記述などがあります。最終的に21,212件の新聞記事、4,142件の景況調査コメントが集まりました。他にも、大企業への影響を確認するため、有価証券報告書から458件のコメントを収集しています。

ただ、この膨大な文書を全て私が読み込み、思い込みや偏りを排除して客観性を確保しながら整理・分析するのはまず不可能です。ここにテキストマイニングを活用することにしたのです。

TMSで収集したテキストをどのように解析されたのでしょうか。

加藤 まず単語頻度解析を**TMS**のデフォルト条件で行ったところ、台風ごとの特徴が見えてきてとても興味深かったです。単語は品詞ごとに確認できますが、今回は固有名詞が役立ちました。抽出された地名を見れば被害の大きかったエリアがつかめますし、会社名を見れば被害の内容が想像できました。

次に係り受け頻度解析で「台風－影響」「被害－発生」といった、収集したコメントの内容からして当然出現するような係り受けを削除しました。この作業を何度か繰り返すことで、意味を持った係り受けが残ります。それを態度表現解析で「否定」や「困難」の設定にすると



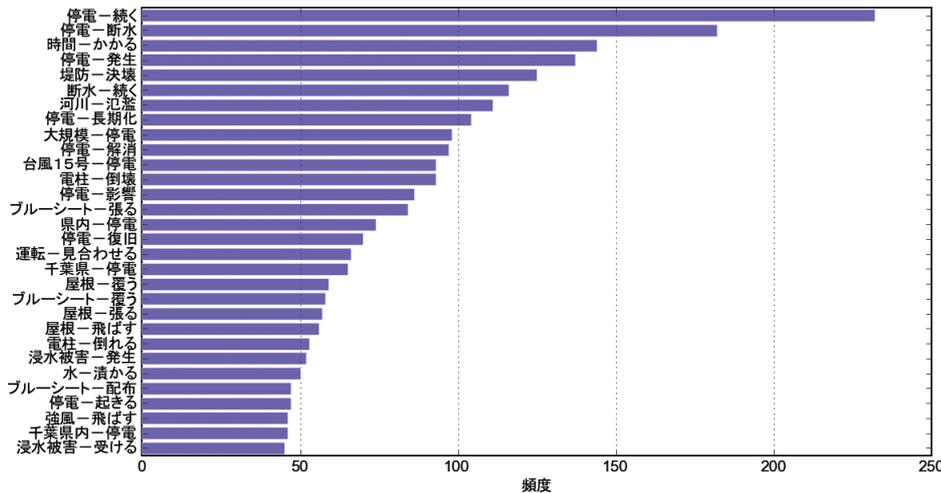
主任研究員 気象予報士
加藤 大輔 様

Profile

株式会社
東京海上研究所 様

東京海上グループのシンクタンクとして、中長期的視点から保険会社の経営に大きな影響を及ぼしうる環境変化に関する研究を行っている。地球温暖化と自然災害リスク、変化・進化する未来社会、この2つを主な対象領域として東京海上グループの経営判断・戦略策定に資する研究を行っている。

「〇〇できない」「〇〇で困った」といった、台風によるネガティブな影響を具体的に把握できました。ここまで行えば重要度の高いキーワードが多数出てくるので、それを係り受け頻度解析の係り元、係り先にそれぞれ指定して深掘りしていきました。例えば千葉の台風被害で多かった「停電」を入れると、断水などの停電の影響などが見えてきます。この作業を数十回繰り返すことで、各台風がもたらした影響の特徴が明確になっていきました。こうして解析した結果をもとに100項目程度の質問を作り、2万人以上を対象にアンケート調査を行いました。



係り受け頻度解析の結果の例。これらを確認することで、各台風がもたらした影響の特徴が浮かび上がってくる。

TMSは解析がブラックボックス化しないので、学術研究に向いている

TMSを導入された背景について教えてください。

加藤 この調査は公的研究費を原資としているため、資材購入には入札が条件でした。打診した会社の中でNTTデータ数理システムは解析デモのほか、私のやりたい解析ができるかどうか確認するための、無料トライアルの機会も提供してくれました。

私が**TMS**に感じたのは、学術研究にとっても適したツールだということです。ビジネスでは解析結果に目が行きがちですが、研究では全プロセスをエビデンスとしてオープンにする必要があります。**TMS**であれば解析で設定したパラメータが詳細に保存・出力できます。また、処理プロセス(計算式など)は技術文書として公開されています。

テキストマイニングは初めてのご経験だったそうですね。

加藤 私はもともと、物理モデルを用いた水害のシミュレーションを担当していました。テキストマイニングの経験はありませんでしたが、少し調べてみて使えそうだと感じました。

TMSを使うにあたって最初に見たのは動画による手順解説です。実際の画面上でどのようにマウスを操作して解析を進めていくかが分かりやすく、これで全体像をつかみました。作業を進める上ではチュートリアルが役立ちました。比較的短時間で操作方法が把握できますし、さらに詳細を知りたいときはマニュアル本編で確認できます。また、私がやりたい解析の手順をNTTデータ数理システ

アンケート回収後も、TMSが役立ったと聞きました。

加藤 アンケートでは自由コメント欄を1つ設定していました。台風が原因で生活や仕事に不安や困難を感じた経験があれば、程度を含めて自由に記載してもらおうものです。その回答が7,000件ほど得られたので、内容のカテゴリ分けや解析にも**TMS**を使いました。コメントを態度表現解析にかけると、何ができなくて不安に思ったのかといった、被害者の状況や心情がつかめます。それらの結果は次の研究や調査の素材として活用しようと思っています。

ムからアドバイスしてもらい、どうしても解決できない点は問い合わせで解決できました。こうしたサポートもあり、導入がとてもスムーズに進みました。

今回の成果や今後の展望をお聞かせください。

加藤 当社が参画しているのは、我が国発の破壊的イノベーション創出を目指す「ムーンショット型研究開発制度」であり、全部で9つある研究プログラムのひとつに所属しております。目標は、「2050年までに、激甚化しつつある台風や豪雨を制御し極端風水害の脅威から解放された安全安心な社会を実現」することです。気象を制御するという意欲的な技術開発ですが、私はこの技術の社会的影響の評価も必要だと考えました。極端な風水害は人々にどんな困難や不安をもたらしているのか、それが気象制御によって解消あるいは軽減できるのか。それを評価することは、研究の価値を確認することにもつながります。その第一歩として、これまでの実態を把握し整理したいと考え、今回の社会調査を企画・実施しました。その結果、過去の台風がもたらした影響や、影響ごとの社会的関心の差を確認することができました。研究チームのメンバーや外部のアドバイザーからもこの調査は非常に有用だというご意見を頂き、活動を継続することになっています。今後は、視点を変えた調査をあと2回は行いたいと考えています。