

【研究成果報告時/研究成果の概要】様式-2-ウ

水害対応ヒヤリ・ハット事例の半自動抽出による
継続的な研修教材の更新に関する研究

東京大学 教授 大原美保

地方自治体は、常に最新の知見を踏まえながら職員の災害対応力の向上をはかる必要がある。研究代表者は、地方自治体の職員が災害時に「困る・焦る・戸惑う・迷う・悩む」などの状況に陥る事例を「災害対応ヒヤリ・ハット事例」と定義し、災害対応検証報告書からこれらの事例を抽出し、研修等の人材育成に活用する取り組みを行ってきたが、マンパワーの制約により新たな事例の継続的な抽出は容易ではない。本研究では、収集済のヒヤリ・ハット事例を学習データとして、深層学習モデルBERTを構築し、収集済の事例に類似した文を新たな災害対応検証報告書から新規抽出する技術の開発を行った。正答率が平均97%となる非常に高精度の事例抽出モデルを開発することができたが、最も高い精度で学習が安定するデータ数を探索する必要性も示唆された。また、学習済のモデルを用いて事例候補を表示できる事例抽出支援システムの開発も行い、半自動的にヒヤリ・ハット事例を抽出して研修教材を継続的に更新できる体制を構築することができた。

Key Words : 水害、災害対応、深層学習、人材育成

1. はじめに

災害発生時に円滑な対応を行うには、過去の災害での教訓から学び、起こり得る困難をあらかじめ予測して、必要な事前対策を講じておく必要がある。しかしながら、急速に人口減少が進む我が国では、団塊の世代の大量退職に伴い、災害対応の教訓や経験を地方自治体内で継承していくことはますます困難になっている。災害対応の業務や教訓を学ぶ人材育成の機会がますます重要となっている。近年、地方自治体が、災害後に災害対応検証報告書を公表するようになり、職員自らの視点で、災害対応の課題を知ることができる。研究代表者は、地方自治体の職員が災害時に「困る・焦る・戸惑う・迷う・悩む」などの状況に陥る事例を「災害対応ヒヤリ・ハット事例」と定義し、災害対応検証報告書からこれらの事例を抽出し、研修等の人材育成に活用する取り組みを行ってきた。水害での典型的な事例をまとめた「水害対応ヒヤリ・ハット事例集（地方自治体編）」¹⁾の公表も行っている。災害の頻発に伴って多くの災害対応検証報告書が刊行され、最新の事例を学びたいというニーズはあるが、マンパワーの制約により新たな事例の継続的な抽出は容易ではない。

そこで本研究は、収集済の水害対応ヒヤリ・ハット事例を学習データとして、深層学習モデルBERTを構築し、収集済の事例に類似した文を新たな災害対応検証報告書から新規抽出する技術を開発することを目的とする。

2. 研究方法と実施体制

(1) 研究方法

研究は、以下の4つのステップにより実施した。研究期間は2年間であり、1年目にStep1、Step2を実施し、2年目にStep3、Step4に進んだ。

- Step1 水害対応ヒヤリ・ハット事例の出現傾向の分析
- Step2 収集済の事例を学習データとした深層学習モデルの作成
- Step3 深層学習モデルを用いた事例抽出の精度検証
- Step4 事例の半自動抽出による継続的な研修教材の更新体制の構築

(2) 研究の実施体制

本研究の研究代表者は大原美保（東京大学大学院情報学環総合防災情報研究センター・教授）、共同研究者は山下倫央（北海道大学大学院情報科学研究院・准教授）である。防災分野の研究代表者と、AI分野の共同研究者の協力により、各々の分野では実施不可能である新たな研究展開の創出を目指した。

3. 水害対応ヒヤリ・ハット事例のデータベースを用いた事例出現傾向の分析

(1) 災害対応検証報告書の傾向に関する分析

Step1としてまず、教訓情報を含む災害対応検証

報告書の構成を分析し、深層学習モデルを用いた事例抽出の適性を分析した。図-1に、2000～2020年の災害の発生年に対して都道府県・市町村が刊行した教訓情報を含む報告書数を示す。これらは合計96点となり、広域災害となった平成30年7月豪雨災害、令和元年東日本台風については、多数の市町村が災害対応検証報告書を公表している。

災害対応検証報告書は、被災した地方自治体が自主的に刊行したものであるため、報告書の書き方は地方自治体の考えによって様々であり、特に近年ますます多様になっている。災害対応検証報告書の形式は、概ね下記のA～Eのタイプに分類できる。

- ・A：経過状況、原因や結果、教訓が別のパラグラフに記述され、近くに配置されている
- ・B：原因や結果、教訓は、別のパラグラフに記述されているが、特に、教訓が別の章立て（例えば、改善策、今後の対応等）になっている
- ・C：経過状況、原因・結果、教訓が同一のパラグラフに混ざって書かれている
- ・D：経過状況が文章または表形式で記述された後、原因・結果、教訓が表形式で書かれている
- ・E：原因・結果、教訓が表形式で書かれており、経過状況の記述はほぼ無い

図-2の通り、報告書の約4割強は、DやEのように、事例が表形式で記述されている。表形式の記述は、テキストファイルとしての抽出が難しく、モデルを用いた事例抽出には向かない。残りの約6割の報告書は、モデルを用いた事例抽出は可能と考えられた。

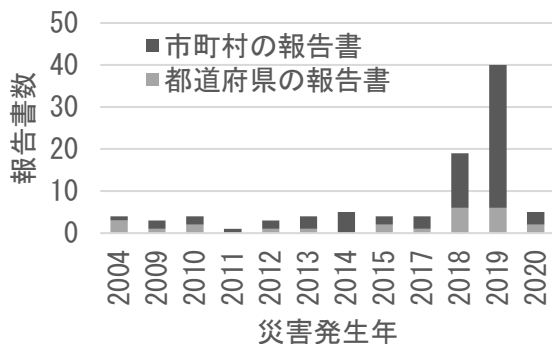


図-1 水害に関する災害対応検証報告書数の推移

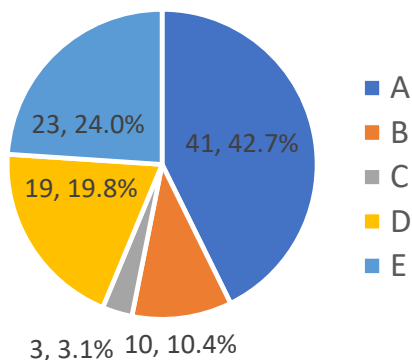


図-2 96の報告書でのタイプ別内訳

2カ年研究用 (R6)

(2) 水害対応ヒヤリ・ハット事例の抽出とデータベース化

研究代表者らによるこれまでの活動において、教訓情報を含む 96 の報告書を精読し、自治体職員が「困る・焦る・戸惑う・迷う・悩む」などの状況を陥っていると示唆される記述を水害対応ヒヤリ・ハット事例として特定し、3,973 の事例を収集している。事例は、畑村 (2005) が提唱する「失敗学」³⁾の考え方を参考に、「事象の概要・経過・原因・結果・教訓」から構成されるデータベースに整理してきた。また、災害対応業務のカテゴリー別、抽出元の災害対応検証報告書の情報も、データベースに入力している。一部の事例についてはデータベース化の作業の途上にあつたため、本研究において、3,973 の全ての事例を整理し、事例データベースとしての整備を完了した。

(3) 水害対応ヒヤリ・ハット事例の出現傾向の分析

図-3は、全事例を業務のカテゴリー別に分類したものである。

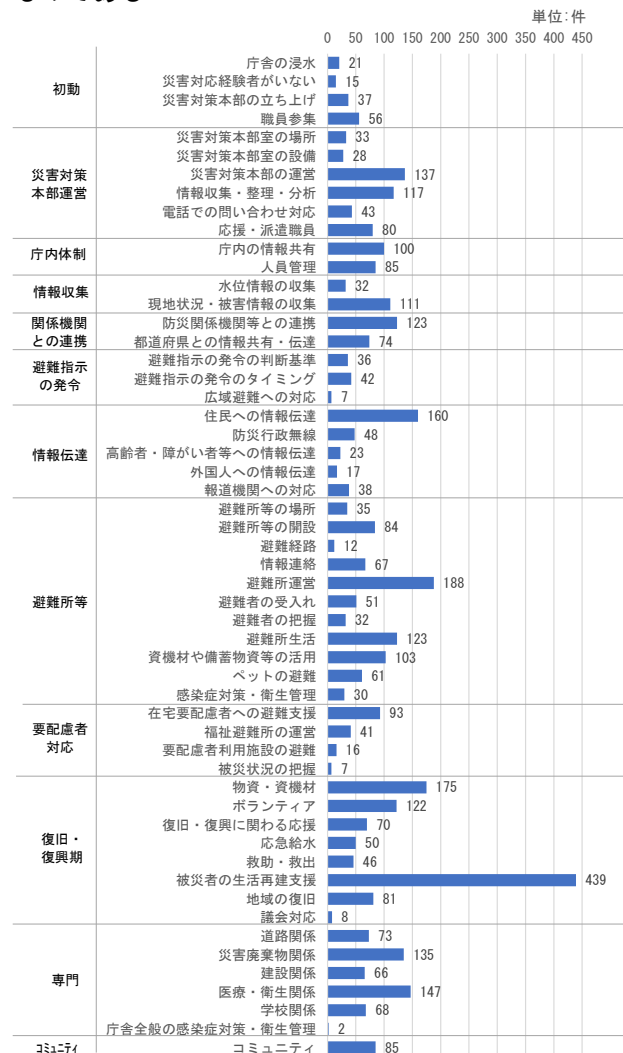


図-2 水害対応ヒヤリ・ハット事例のカテゴリー別分類

初動期においては、住民への情報伝達、災害対策本部運営、防災関係機関等との連携、情報収集・整理・分析などが多い。避難所開設以降は、被災者生活再建支援関連、避難所運営、物資資機材が多い。このような傾向分析から、担当職員が「困る・焦る・戸惑う・迷う・悩む」などの状況に陥りがちな典型的な局面をあらかじめ知ることができる。

4. 事例の深層学習モデルの作成と検証

Step2 では、前章で構築した水害対応ヒヤリ・ハット事例データベースを用いて、深層学習モデルを作成した。BERT モデルを用いて、1 文を入力とし、水害対応ヒヤリ・ハット事例（正例）か、事例ではない（負例）かの確率を出力する文章分類モデルの作成を行った。作成の流れは、データ抽出・前処理、データセット作成、モデル・トークナイザー選定、タスクに合わせたヘッド設定、学習・評価である。

モデルの開発環境は以下である。データの前処理・作成は Windows10 の環境で行い、学習は AWS の SageMaker を使用した。

- ・主な使用言語：Python3.8, 3.9
- ・学習環境 AWS SageMaker
- ・深層学習フレームワーク：Pytorch
- ・仮想環境：Anaconda, docker
- ・OS：Windows 10, Ubuntu 20.04

(1) 前処理および深層学習モデルの作成

提供された事例文章データから、機械学習で問題となりそうなノイズ（記号、ページ表記、部署情報、改行コード等）を削除・整形する処理を行った。1事例に2つ以上文章がある場合は、それぞれ別の事例文章として分割した。

(2) 深層学習モデルの作成

学習には、前章で構築した水害対応ヒヤリ・ハット事例を正例として用いた。一方、学習には水害対応ヒヤリ・ハット事例ではない負例も必要となるた

め、災害対応検証報告書から、水害対応ヒヤリ・ハット事例ではない文章を抽出して、負例データベースを構築した。

HuggingFaceで掲載される事前学習済みのBERTモデル（東北大学）を使用した。トークナイザーには、MeCab with the Unidic 2.1.2を用いた。事例分類モデルのため、ヘッドには2つの出力を持つネットワーク層を追加し、ファインチューニングを行った。

(3) 学習・評価

作成したデータセットを用いて、それぞれ5分割交差検証を行い、学習した結果を評価した。学習率、その他パラメータはPytorchで実装された関数のデフォルト値を使用する。パラメータは、以下の通りである。

- ・バッチサイズ：32
- ・最適化手法：AdamW
- ・学習率：0.00005
- ・エポック：20

(4) 負例数を増やした場合の学習効果の分析

まずは、負例数を変化させた場合の学習効果を把握するため、3つのデータセットで事例分類モデルの学習・評価を行った。正例データとしては、2017年以降の災害対応検証報告書のうち、研究代表者らが以前より事例データベースとして整理済であった1710件を用いた。データセット1は、負例として、事例文章に近い位置からランダムに抽出した事例ではない文章1000件を用いたものである。データセット2は、データセット1より負例数を増やし、事例文章が含まれていないページの文章184,099件としたものである。データセット3は、事例文章が含まれていないページの文章を10～200文字の制限でランダム抽出した10,000件としたものである。データセット3では、正例も、1文章が10～200文字以内となる制限を加え、1670件とした。

表-1 学習結果（上段：データセット1による学習、下段：データセット3による学習）

	パターン1	パターン2	パターン3	パターン4	パターン5	平均
再現率	0.99	1.0	0.995	1.0	0.975	0.992
適合率	0.98	0.99	0.99	0.99	0.98	0.986
F値	0.985	0.995	0.992	0.995	0.977	0.989
Accuracy	0.985	0.995	0.992	0.995	0.978	0.989

	パターン1	パターン2	パターン3	パターン4	パターン5	平均
再現率	0.77	0.92	0.92	0.87	0.94	0.884
適合率	0.87	0.91	0.91	0.92	0.90	0.902
F値	0.82	0.92	0.92	0.90	0.92	0.896
Accuracy	0.95	0.98	0.98	0.97	0.98	0.972

表-2 学習結果 (データセット6)

	パターン1	パターン2	パターン3	パターン4	パターン5	平均
再現率	0.98	0.97	0.97	0.98	0.95	0.97
適合率	0.93	0.94	0.92	0.90	0.93	0.92
F値	0.96	0.96	0.95	0.94	0.94	0.95
Accuracy	0.97	0.97	0.96	0.96	0.96	0.96

Step3では学習効果の評価を行った。データセット1を用いて、それぞれ5分割交差検証を行い、学習した結果を表-1上段に示す。5分割交差検証とは、データを5個に分割して、そのうち1つをテストデータに残りの4個を学習データとして評価を行う手法である。学習効果の評価指標としては、下記の4つを用いる。

- ・正答率 (Accuracy) : 事例かそうでないかを当てた割合
- ・再現率 : 本来事例であるうち、どれくらいを抽出できたか
- ・適合率 : 事例と判定したうち、どれくらい事例だったか
- ・F値 : 再現率と適合率の調和平均

データセット1による5つのパターンの学習の平均は、学習の正答率 (Accuracy) 98.9%、再現率99.2%、適合率98.6%、F値98.9%となった。一般的な学習タスクにおいて99~100%の値は非常に高精度であり、データセット1のファインチューニングは有効であると言える。

一方で、データセット2では、5つの分割パターンで全て負例と分類する結果となった。データセットが大きく不均衡であったことが原因であると考えられる。

データセット2の負例データは文字数が多いものが含まれ、正例データに比べても多様であった。データセット3は、データの偏りを小さくするため、正例データの頻度分布より10~200文字のみを対象としたデータセット10000件とし、学習・評価を行った。

データセット3の学習結果を表-1下段に示す。データセット2で見られたような分類の偏りは見られず、平均再現率は88%という精度で分類できた。BERTモデルをファインチューニングした本タスクの文章分類性能はかなり高いことが分かったが、一方で、負例データを追加しすぎると学習に影響が出ることが分かった。このことから、最も高い精度で学習が安定する適切な負例データのサンプル数を探索する必要がある。

(5) 正例数を増やした場合の学習効果の分析

次に、3つのデータセットを用いて、正例数を変化させた場合の学習効果も分析した。増やした正例は、前章で述べた通り、今回新たにデータベースに加えたものである。データセット4は、前章で用いたデー

タセット3から作成し、正例を、文字数は9~158文字の1415件としたものである。データセット5では、交差検証のため5分割されたデータセット4のそれぞれの正例データに対し、新たに9~158文字の3,348件を追加したものである。データセット6は、正例を全4763件としたものである。負例は、データセット3と同様に10,000件とした。

紙面の都合により、表-2に、データセット6の学習結果のみを示す。正例データの追加により、再現率は向上が見られたが、適合率は低下するということが分かった。一般的にこれら指標はトレードオフの関係にあるため、学習において問題のある現象ではないといえる。実用時はヒヤリ・ハット事例の取りこぼしを防ぐことを優先するため、再現率が重要である。よって、モデル改善には正例データを追加していくという方針に問題がないといえる。なお、クラス間のデータ数に違いがあるため、通常の機械学習モデルの評価値とは異なる点は留意する必要がある。

5. 継続的な事例更新に向けた水害対応ヒヤリ・ハット事例の抽出支援システムの開発

事例抽出プロセスの半自動化を目指して、収集済の事例に類似した文章を抽出して、事例候補として挙げる事ができる事例抽出支援システムの開発を行った。本システムは、前章までで開発した水害対応ヒヤリ・ハット事例の分類モデルを使用し、災害対応検証報告書のPDFファイルと、モデルを用いた事例の推論結果をブラウザ上で表示させるシステムとして設計した。Pythonで実装されたWebアプリケーションフレームワークであるDjangoによって作成しており、GPUを有するパソコンでのスタンドアローン環境での実行を想定している。

図-3に、画面イメージを示す。九州北部豪雨災害で福岡県の災害対応検証報告書を用いた表示例である。左画面は、元の災害対応検証報告書である。右画面には、左側に示したページにおいてヒヤリ・ハット事例（正例）と推論される箇所及び事例ではない（負例）と推論される箇所を表示しており、特に事例（正例）は黄色のハイライト付きで表示している。ページ番号、行数、推論の確からしさ (Score) もあわせて表示することとした。左側のページにおいても、事例（正例）と推論される箇所を黄色のハイライト付きで表示している。これにより、学習済

のモデルが、どの箇所をヒヤリ・ハット事例であると推論しているかが、手軽にわかるようになった。

本システムを用いて、半自動的に事例抽出ができるようになり、手作業による抽出に比べて、マンパワーを格段に減らすことができるようになった。一方で、深層学習モデルは、一般的記述、(一般に、〇〇が困難である、等)、他の文献からの引用、過去の災害での記述、脚注、なども抽出してしまうことがわかった。これらは、テキストとしては正例ではあるが、文脈から見ると、事例抽出は適切ではない。深層学習モデルでは、これらを誤抽出してしまうため、抽出後に、手作業による選別が必要がある点に留意すべきであることがわかった。

6. おわりに

本研究では、収集済のヒヤリ・ハット事例を学習データとして、深層学習モデルBERTを構築し、収集済の事例に類似した文を新たな災害対応検証報告書から新規抽出する技術の開発を行った。BERTモデルを用いて、1文を入力とし、水害対応ヒヤリ・ハット事例(正例)か、事例ではない(負例)かの確率を出力する文章分類モデルの作成を行った。

負例を増やした場合、正例を増やした場合の双方についての検討を行い、負例・正例の数による学習効果の違いを分析した。負例を1000件、10000件と増やした結果、いずれも正答率は97%となり、高精度の結果が得られたが、負例データを追加しすぎると、学習に影響が出ることも分かった。正例を1415件から4763件に増やして検証を行った結果、再現率は向上が見られたが、トレードオフの関係にある適合率は低下した。実用時はヒヤリ・ハット事例の取りこぼしを防ぐことを優先するため、モデル改善には正

2 カ年研究用 (R6)

例データを追加していくという方針に問題がないと考える。今後も引き続き、最も高い精度で学習が安定する負例・正例データ数を探索する必要がある。

最後に、収集済の事例に類似した文章を抽出して、事例候補として挙げるができる事例抽出支援システムの開発を行った。本システムを用いることで、これまでの事例抽出における膨大な手作業を省力化して半自動的にヒヤリ・ハット事例を抽出し、研修教材を継続的に更新することができる体制が構築できた。最新の災害対応検証報告書から新たな知見を迅速に収集することにより、これらを反映した職員研修や図上訓練等の開催につながる。

謝辞: 本研究は、一般財団法人河川情報センターからの研究助成を得て、実施したものである。また、災害対応ヒヤリ・ハット事例は、過去に被災した地方自治体による災害対応検証報告書から抽出したものである。ここに記して、感謝の意を表する。

参考文献

- 1) 大原美保・栗林大輔・藤兼正和, 地方自治体職員が直面する水害対応ヒヤリ・ハット事例の分析, 土木学会論文集F6(安全問題), Vol. 76, No. 2, 1_81-1_88, 2020.
- 2) 国立研究開発法人 土木研究所 水災害・リスクマネジメント国際センター, 水害対応ヒヤリ・ハット事例集(地方自治体編及び別冊新型コロナウイルス感染症への対応編), 2020. https://www.pwri.go.jp/icharm/special_topic/20200625_flood_response_collection_j.htm (参照年月日: 2022. 6. 30)
- 3) 畑村洋太郎, 失敗学のすすめ, 講談社文庫, 2005

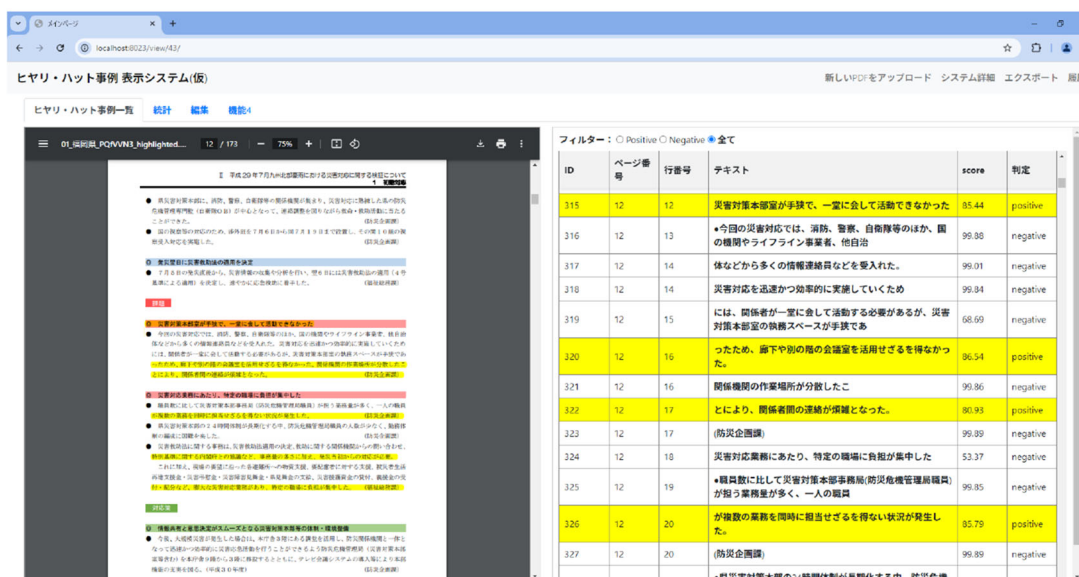


図-3 抽出支援システムの表示画面