

# 介護事故防止に向けた事故事例テキストの分析

○ 峯崎智裕<sup>†</sup> 井上創造  
九州工業大学

## Text Analysis of Incident Reports for Preventing Accidents in Caregiving

○ Minezaki Tomohiro and Sozo Inoue  
<sup>†</sup> Kyusyu Institute of Technology University

**Abstract** : To detect the patterns of incidents in caregiving facilities, and to increase the sensing efficiency for accident preventions, we performed text mining with incident reports. As a result, important places and times in a day were extracted.

### 1. まえがき

近年, 介護分野における事故が問題になっている. 介護事故を防止するためには, 事故の原因を調査する必要がある, そのための研究が年々活発に行われている. それらの研究の中には, センシング技術を使用するものが多く存在する[1]. しかし, センシングにはコストがかかってしまう. そこで, 本研究では, 事故事例テキストをテキストマイニングすることにより, 事故が起きやすい状況を取得し, さらに, どの場所にセンサを取り付ければ有用なデータが得られるか考察する.

### 2. 背景

近年の介護分野における問題として, 「ベッドからの転落」, 「段差につまずき転倒」などといった事故の割合が大きく, このような事故を検出・予防するには, 事故の原因を知ることが大事である. しかし, 事故は「介護者」, 「被介護者」, 「状況」などといった環境要因の影響が大きく, 同じような事故でも原因はまったく違うこともある.

現在, 文書の中に現れる単語を行列にした単語・文書行列を用いた潜在意味分析が多くの研究で使われている. 本研究では, この技術を用いて, 事故事例テキストにより事故が起きやすい状況を検出することを目的とする. さらに, 事故防止に向けてセンシングをする際, どこにセンサを設置すれば有用なデータが得られるかも分析する.

### 3. 分析

#### 3.1 事故事例テキスト

本研究では, ある介護サービス事業者の, 介護施設で起きた事故・ヒヤリハットを記載した事故事例テキストを使用し, テキストマイニングを行う. このテキストには事故詳細の他に日時, 年齢, 介護度, 認知度, 事故歴, 事故の種類, 場所, 目的, 重症度といった属性情報も記載されている.

施設全体では, 以下の通りである.

- 事故件数は 384 件(内ヒヤリハット 328 件)
- 重症度: 重度 85 件 (22.1%), 軽度 187 件 (48.7%), なし 112 件 (29.2%)
- 転落・ずり落ちが 211 件で 1 番多い

#### 3.2 分析手順

上記の事故事例テキストから, 「就寝介助時, 物音がする為, 確認行くと転倒されていた。」といった事故詳細部分を取り出し, MeCab を用いて分かち書きを行う. それを元に形態素解析を行い, 以下のような単語・文書行列を作成する.

row.names	文書1	文書2	文書3	文書4	文書5	文書6
トイレ	1	0	1	2	0	0
ベット	1	1	1	0	0	0
ベッド	1	1	0	0	0	0
下	1	0	0	0	0	0
看護	1	0	0	0	0	1
左	1	0	1	0	0	0
際	1	0	0	1	0	0
湿布	1	0	0	0	0	0
車椅子	1	0	0	2	0	1
首	1	0	0	0	0	0
床	1	0	1	1	1	0
対応	1	0	0	0	0	0
膝	1	0	0	0	0	0
報告	1	0	0	0	0	0
訪室	1	1	1	1	1	0

図 1. 各事故事例における単語・文書行列

図 1 の  $i$  行  $j$  列の数値は,  $j$  列目の文書に  $i$  行目の単語が含まれている回数を表す. なお, 1 つの文書にしか現れない単語は除外してある.

次に, この図の「ベット」と「ベッド」のように, 同じような意味の単語が別の単語として認識されている事がわかるので, k-means 法によるクラスタリングを行い, 610 単語を 50 個のクラスタに分類する.

クラスタリングした上記の単語・文書行列と事故事例テキストの「場所」, 「介護度」, 「時刻」,

「曜日」, 「利用区分」, 「年齢」, 「目的」, 「事故歴」, 「認知度」といった属性を結合した多変量データを作成する. そして, この多変量データを

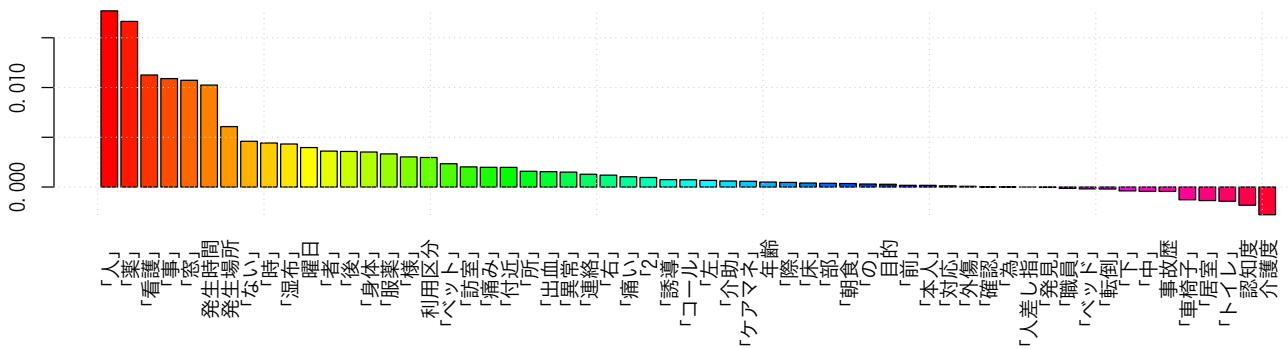


図 2. 属性・単語の重要度

説明変数としたランダムフォレスト法を用い、目的変数である事故の重症度を推定する。

ランダムフォレスト法には以下のような特徴があり、今回の分析に適切であると考えられる。

- ・ 自動的に過学習を抑え、汎化能力の高いモデルを生成する
- ・ 変数間の交互作用を中和した上での説明変数の目的変数への重要度が分かる
- ・ 弱学習器の決定木を抽出すれば、枝の分岐条件を直感的に理解しやすい

最後に、ランダムフォレスト法によって得られた重要度の高いクラスタに含まれている単語を調べ、同様に属性の変数を用いて決定木を行い、属性の詳細を分析する。

#### 4. 評価

ランダムフォレスト法を用いて、予測を行った結果は図2の通りである。

属性のみを説明変数とした場合 48.63%、図2のように属性と単語を説明変数に用いた場合は 64.78%の精度に上昇した。

図2より、属性は「発生場所」、「発生時間」、「曜日」、「利用区分」、「年齢」の順に重要な変数であり、また、「人」、「薬」、「看護」、「事」、「窓」というクラスタの重要度が高いことがわかった。これらのクラスタには「臥床」、「転落」、「服薬」、「マットコール」といった単語が含まれており、これらの単語も重要度が高いといえるだろう。ただし、「事」、「窓」クラスタには雑多な単語が多く含まれており、重要度の信頼性は低いと考えられる。

さらに、重要度の高い属性である事故発生場所・時間を説明変数、重症度を目的変数として決定木を行った結果、「ダイルーム」、「居室」、「共有トイレ」、「浴室」、「廊下」といった場所で、6時12分から6時44分の間起きた事故の85.6%が重度の事故であることがわかった。また、「食

堂」で12時23分以降の時間帯で起きた事故の52.9%が重度の事故であった。

これらの結果より、今後の研究で介護施設にセンサ機器を設置する際、「食堂」や「ダイルーム」、「居室」、「共有トイレ」、「浴室」、「廊下」といった場所に取り付けると有用なデータが取得できることが見込まれる。

#### 5. 考察

ランダムフォレスト法で説明変数に属性のみを使用した場合に比べ、属性と検索語・文書行列を使用した場合の方が精度は上がったが、それでも精度が高いとはいえない結果であった。原因として、事前に行った単語のクラスタリングの精度が低かった事が考えられる。これは、同じクラスタに分類された単語が似ていないことや、似ている単語（「ベット」と「ベッド」等）が別のクラスタに分類されていることにも当てはまる。

これらの改善点として、データを増やしてクラスタリングを行うことにより、ランダムフォレスト法の精度も上がるのではないかと考えられる。

#### 6. あとがき

介護事故事例テキストに対して形態素解析を行い、事故の属性・単語の重要度を調べた結果、「臥床」、「転落」、「服薬」、「マットコール」といった単語や「発生場所」、「発生時間」といった属性の重要性が高いことがわかった。また、受容度の高い属性を用い決定木を作成した結果、重度の事故割合が高い場所と時刻の関係性もわかった。

#### 7. 参考文献

- [1] 平澤桂一, 松村成宗, 金丸直義, 阿部 匡伸 (NTT サイバーソリューション研究所). 病院・介護施設における転倒転落事故予防システム. NTT 技術ジャーナル (2008. 1), pp. 32-35. <http://163.137.191.238/journal/0801/files/jn200801032.pdf>