

介護サービス向上に向けた介護事故事例テキストの分析

峯崎 智裕¹ 井上 創造¹

概要：本論文は、事故が起きやすい状況を取得することを目的として事故事例テキストをテキストマイニングした結果について述べる。さらに結果より事故減少のための対策の考察を行う。介護施設での事故のパターンを検出するため、我々は、過去に起きた事故事例テキストをテキストマイニングした。事故事例テキストから、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、単語・文書行列を作成する。次に、k-means法によるクラスタリングを行い、2701単語を50個のクラスタに分類する。クラスタリングした単語・文書行列と事故事例テキストの属性変数を結合した多変量データを作成する。そして、この多変量データから目的変数を除いたデータを説明変数とし、ランダムフォレスト法を用い、目的変数である属性をそれぞれ推定する。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単語クラスタと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。分析の結果、行動の目的と起きやすい事故の種類の関係性といった過去の介護事故傾向がわかった。

Text Analysis of Incident Reports for Preventing Accidents in Caregiving

TOMOHIRO MINEZAKI¹ SOZO INOUE¹

1. はじめに

近年、介護分野における事故が問題になっている。対策の1つとして、要支援者・要介護者に対する介護士数を増やせば介護事故は減少するだろう。しかし、それは現状不可能であり、今後もそのような見込みもない。日本では今後、少子高齢化が進み、労働人口が減少する一方、必要となる介護職員数は増加する。これでは、他の対策を考えざるを得ない。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り、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][2]。今後、センシング実験を行う予定であり、その事前準備として過去の事故傾向を分析する必要がある。今までにもインシデントレポート解析のためのテキストマイニング手法は提案されているが [5]、実データでテキストマイニングしたものは少ない。そこで、本論文は、事故が起きやすい状況を取得することを目的として事故事例テキストをテキストマイニングした結果について述べる。さらに結果より事故減少のための対策の考察

を行う。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。事故事例テキストから、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、単語・文書行列を作成する。次に、k-means法によるクラスタリングを行い、2701単語を50個のクラスタに分類する。クラスタリングした単語・文書行列と事故事例テキストの属性変数を結合した多変量データを作成する。そして、この多変量データから目的変数を除いたデータを説明変数とし、ランダムフォレスト法を用い、目的変数である属性をそれぞれ推定する。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単語クラスタと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。分析の結果、行動の目的と起きやすい事故の種類の関係性といった過去の介護事故傾向がわかった。

2. 背景

近年、介護分野における事故が問題になっており、対策の1つとして、要介護（要支援）者に対する介護士数を増

¹ 九州工業大学

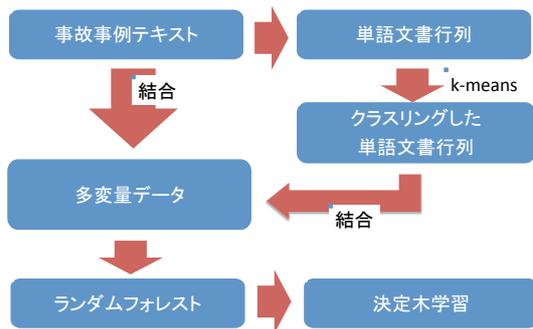


図 1 分析の流れ

やせば介護事故は減少すると考えられる。しかし、それは現状不可能であり、今後もそのような見込みもない。日本では今後、少子高齢化が進み、労働人口が減少する一方、必要となる介護職員数は増加する。要介護（要支援）認定者数は現在、約 546.1 万人である。それに対し介護福祉士の従業者数は約 66.1 万人である（平成 24 年 9 月）[3]。これでは、他の対策を考えざるを得ない。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。今までにもインシデントレポート解析のためのテキストマイニング手法は提案されているが、実データでテキストマイニングしたものは少ない。また、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。そこで、本論文は、事故が起きやすい状況を取得することを目的として事故事例テキストをテキストマイニングした結果について述べる。さらに結果より事故減少のための対策の考察を行う..

3. 分析

この章では、分析に使用したテキストの詳細と分析手順について述べる。

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。

使用したデータは施設全体で、以下の通りである。

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 2 単語文書行 (一部)

- 事故件数は 5189 件
- 重症度：重度 1134 件 (21.9%)、軽度 2428 件 (46.8%)、なし 1627 件 (31.4%)
- 転落・ずり落ちが 2986 件で 1 番多い

3.2 分析の種類

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。

図 2 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、2701 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、この多変量データから目的変数を除いたデータを説明変数とし、ランダムフォレスト法を用い、目的変数である以下の 4 つの属性をそれぞれ推定する。

- 1 利用区分
- 2 事故の種類
- 3 診断
- 4 重症度

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、汎用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

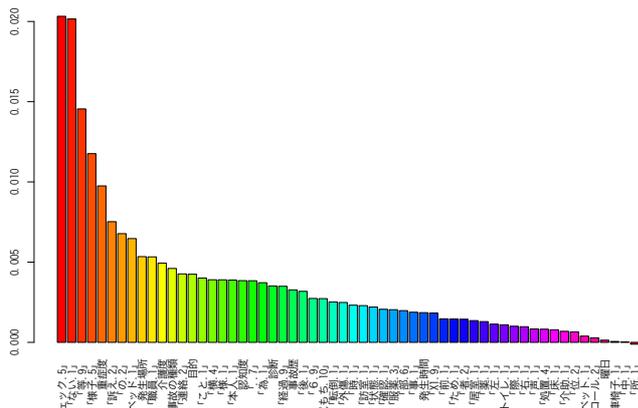


図 3 利用区分を目的変数としたランダムフォレストを用いて予測した時の重要度

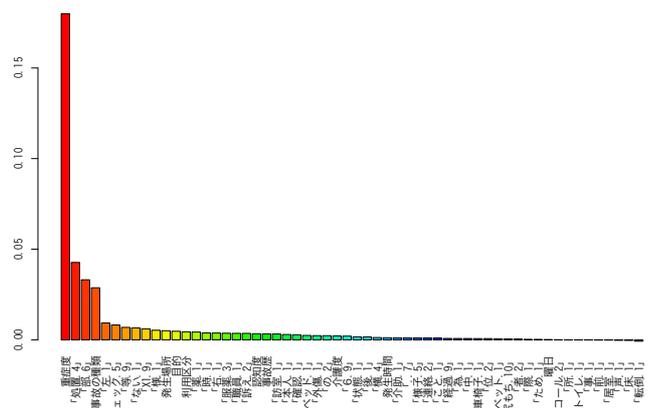


図 5 診断を目的変数としたランダムフォレストを用いて予測した時の重要度

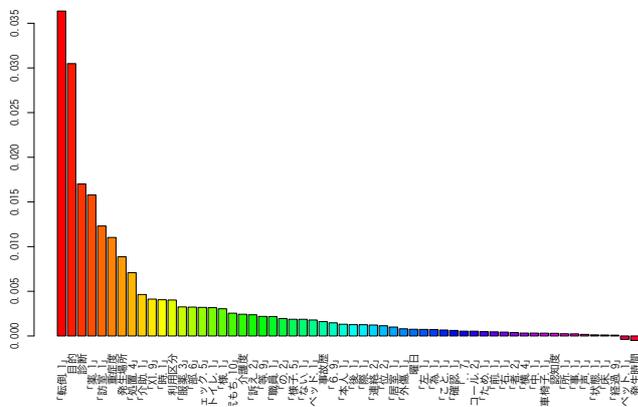


図 4 事故の種類を目的変数としたランダムフォレストを用いて予測した時の重要度

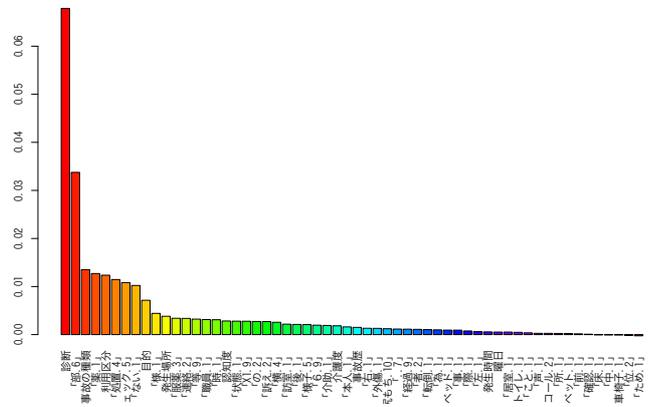


図 6 重症度を目的変数としたランダムフォレストを用いて予測した時の重要度

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、目的変数である以下の4つの属性を予測した。

- 1 利用区分
- 2 事故の種類
- 3 診断
- 4 重症度

結果はそれぞれ図3, 図4, 図5, 図6の通りである。

4.1 利用区分を目的変数として分析を行った場合

図3より、属性は「重症度」、「発生場所」の順に重要な変数であり、また、「チェック」、「等」、「様子」というクラスタの重要度が高いことがわかった。また、図の単語の後ろに付属している数字は、そのクラスタに存在する単語数

を表す。これらのクラスタには「バイタル」、「測定」、「観察」、「移乗」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。

さらに、重要度の高い属性「重症度」、「発生場所」と単語クラスタ「チェック」、「等」、「様子」を説明変数、利用区分を目的変数として決定木を行った結果を、図7に示す。

4.2 事故の種類を目的変数として分析を行った場合

図4より、属性は「目的」、「診断」、「重症度」、「発生場所」の順に重要な変数であり、また、「転倒」、「薬」、「訪室」、「処置」、「部」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「剥離」、「出血」、「骨折」、「病院」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、「転倒」、「薬」の単語クラスタには1単語のみ含まれており、これらの単語はジオの種類に直結するので、重要度の信頼性は高いが、決定木学習の説明変数には含めない。

さらに、重要度の高い属性「目的」、「診断」、「発生場所」と単語クラスタ「訪室」を説明変数、事故の種類を目的変

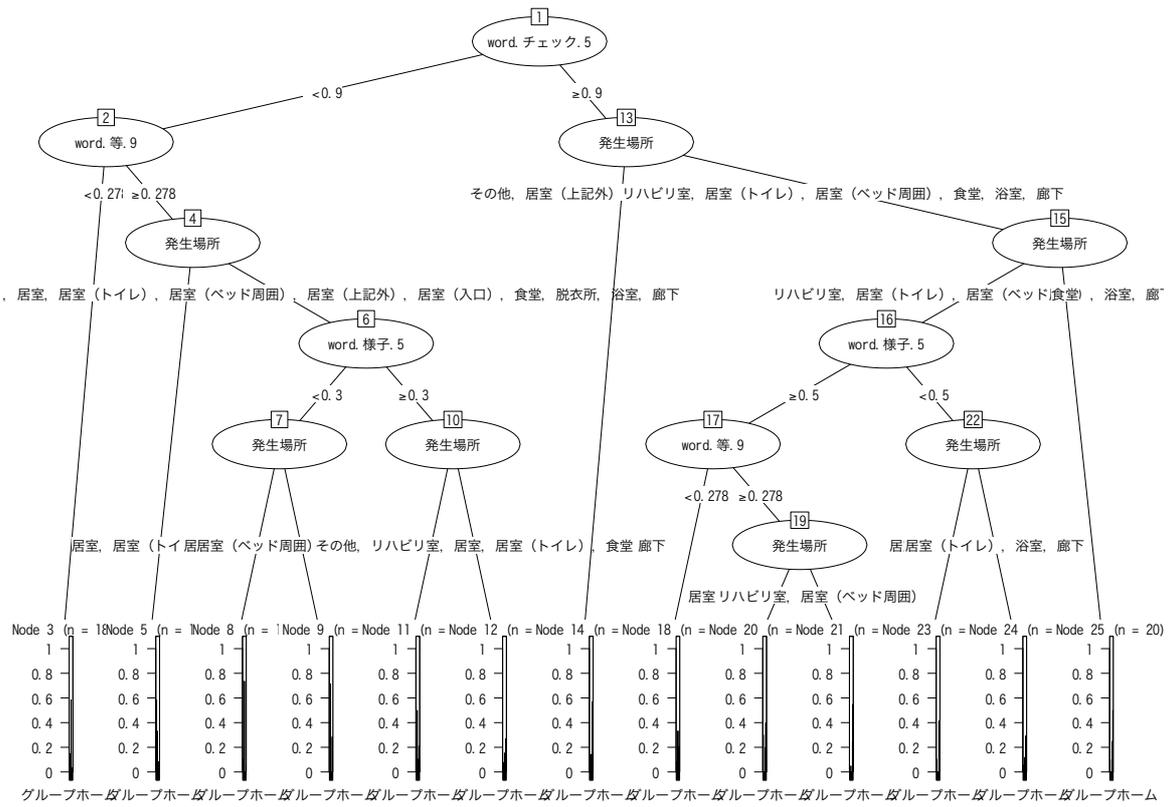


図 7 利用区分を目的変数とした決定木学習

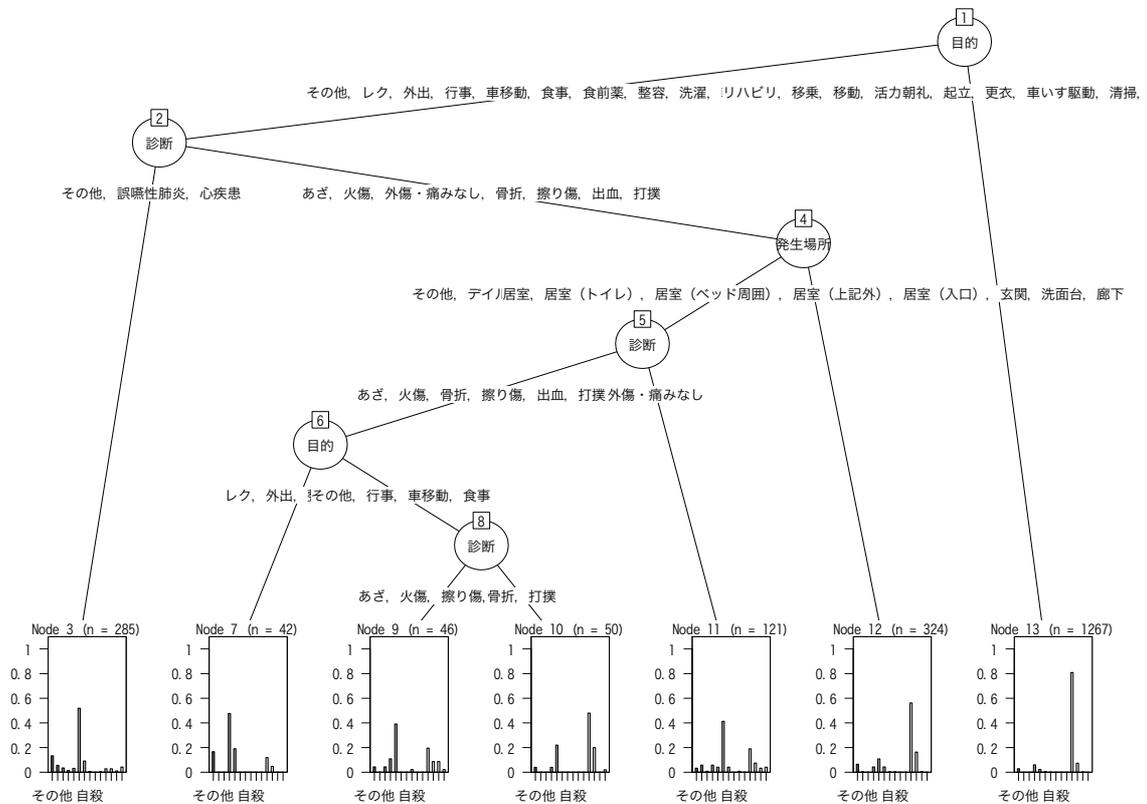


図 8 事故の種類を目的変数とした決定木学習

数として決定木を行った結果を図8に示す。図8の1番右のグラフより、「リハビリ」、「異常」、「移動」、「活力朝礼」、「起立」、「更衣」、「車イス駆動」、「清掃」、「排泄」、「歩行」を目的とした行動中に起きた事故では圧倒的に「転倒・ずり落ち」事故が多いことがわかる。さらに、「レク」、「外出」、「行事」、「車移動」、「食事」、「整容」、「入浴」をを目的とした行動中に起きた事故では「誤薬・誤配」や「あざ・外傷」といった事故が多いことがわかる。

4.3 診断を目的変数として分析を行った場合

図5より、属性は「重症度」、「事故の種類」の順に重要な変数であり、また、「処置」、「部」、「チェック」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「剥離」、「出血」、「受診」、「病院」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。

さらに、重要度の高い属性「重症度」と単語クラスタ「処置」、「部」を説明変数、診断を目的変数として決定木を行った結果を図9に示す。これより、擦り傷や出血の場合は介護士が処置を行っている傾向にあることがわかる。また、介護士が処置を行っていない事故には、外傷・痛みなしや打撲といった軽度事故と判断されたものと、骨折や重度の打撲といった重度事故と判断され、病院で受診したものが含まれていることがわかる。「部」クラスタには、受診や病院といった単語が含まれている。

4.4 重症度を目的変数として分析を行った場合

図6より、属性は「診断」、「事故の種類」、「利用区分」、「目的」、「発生場所」の順に重要な変数であり、また、「部」、「薬」、「処置」、「チェック」という単語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「受診」、「病院」、「剥離」、「バイタル」、「測定」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。

さらに、重要度の高い属性「利用区分」、「目的」と単語クラスタ「部」、「薬」、「処置」、「チェック」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を図10に示す。「診断」と「事故の種類」はランダムフォレストを用いた予測で重要度が高いと判断されたが、これらの属性は事故の重症度に直結するので、決定木では説明変数から省いた。

この図より、単語クラスタ「部」、「薬」が事故詳細に含まれている場合、重症度の事故が起きている傾向にあることがわかる。単語クラスタ「部」、「薬」には「受診」、「病院」、「薬」といった単語が含まれている。また、介護士によって処置が行われた事故では、軽度の事故である場合が多いことがわかる。

5. 考察

評価で利用区分を目的変数にランダムフォレスト法や決定木学習を用いて分析を行ったが、利用区分は事故の結果

を表すものではなく、事故の要因になりうるものである。それに対し、事故の種類、診断、重症度は事故の結果から得られる情報であり、これらの属性を目的変数にして、分析を行った結果の方が、より過去に起きた事故の傾向がわかった。

事故の種類を目的変数にし、分析した結果、「リハビリ」、「異常」、「移動」、「活力朝礼」、「起立」、「更衣」、「車イス駆動」、「清掃」、「排泄」、「歩行」を目的とした行動中に起きた事故では圧倒的に「転倒・ずり落ち」事故が多いことがわかった。さらに、「レク」、「外出」、「行事」、「車移動」、「食事」、「整容」、「入浴」をを目的とした行動中に起きた事故では「誤薬・誤配」や「あざ・外傷」といった事故が多いことがわかった。主に行動の目的と起きやすい事故の種類との関係がわかり、行動の目的毎に介護士が注意すべき項目が明確に分類できると考えられる。

診断を目的変数にし、分析した結果、事故が擦り傷や出血と判断された場合は介護士が処置を行っている傾向にあることがわかった。また、介護士が処置を行っていない事故には、外傷・痛みなしや打撲といった軽度事故と判断されたものと、骨折や重度の打撲といった重度事故と判断され、病院で受診したものが含まれていることがわかった。「部」クラスタには、受診や病院といった単語が含まれている。この結果は外傷に関する結果であり、今までに多く施設毎に対策されているが減少していないので[4]、説明変数を変えて分析をする必要があると考えられる。

重症度を目的変数にし、分析した結果、事故詳細に単語クラスタ「部」、「薬」が事故詳細に含まれている場合、重症度の事故が起きている傾向にあることがわかった。これの原因は、主に誤薬であると考えられる。被介護者に薬を与える際は十分に注意が必要である。また、介護士によって処置が行われた事故では、軽度の事故である場合が多いことがわかった。これは少なからず、介護士が処置を行ったことにより軽度の事故になったもの含まれている。介護士が事故を発見した際の処置により、事故の総数は変化しないかもしれないが、重度事故を軽度事故に抑えることができるので、介護士の事故処理能力が大きく関係することが考えられる。今後センシング実験を行う際は、居室や介護士にセンサを取り付け、介護士が被介護者に対応をできているかの検証を行う必要がある。

また、これらの結果を見てみると、被介護者が1人で行動している時にはもちろん事故が起きているが、そうでない場合にも事故が起きていることがわかる。これは主に誤薬が多いと考えられる。また、「レク」や「行事」、「活力朝礼」中における外傷事故がおきている。これは、介護士がふと目を離した際に起きていることが考えられるので、怪我を起しやすの人に注意して行動することや介護士毎に担当の被介護者を割り振るといった対策が必要である。

また、今回多変量データを作成する際、k-means法によ

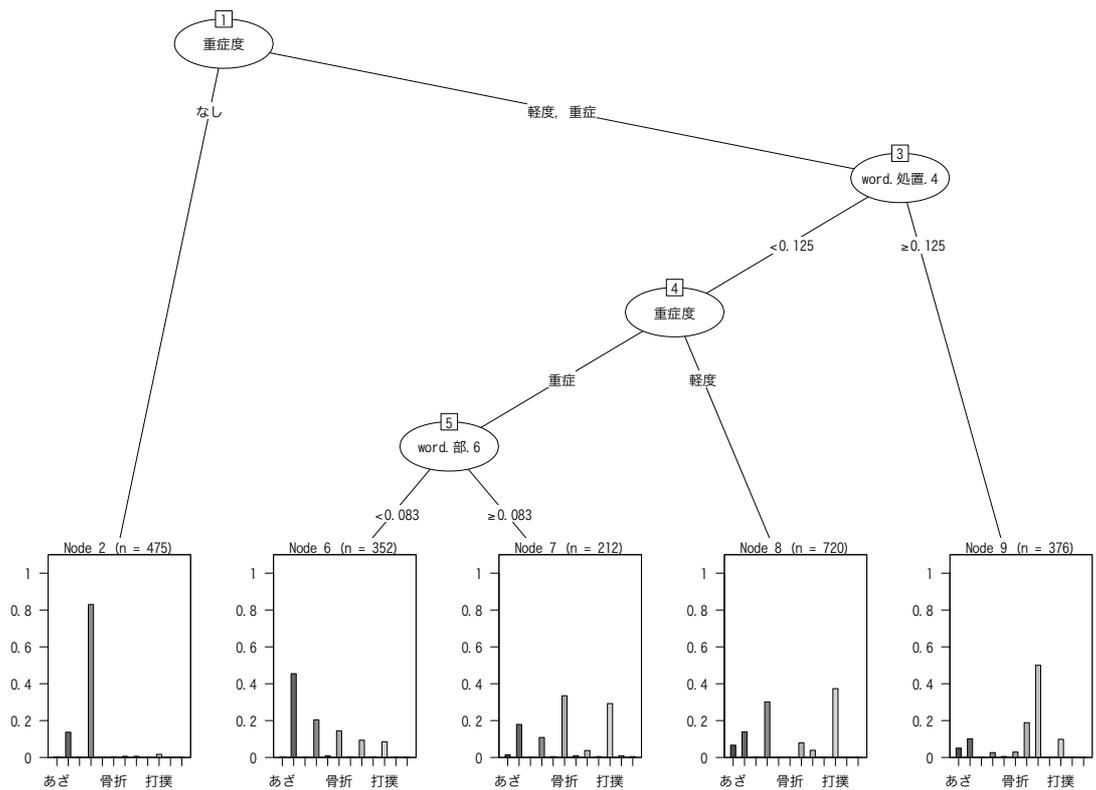


図 9 診断を目的変数とした決定木学習

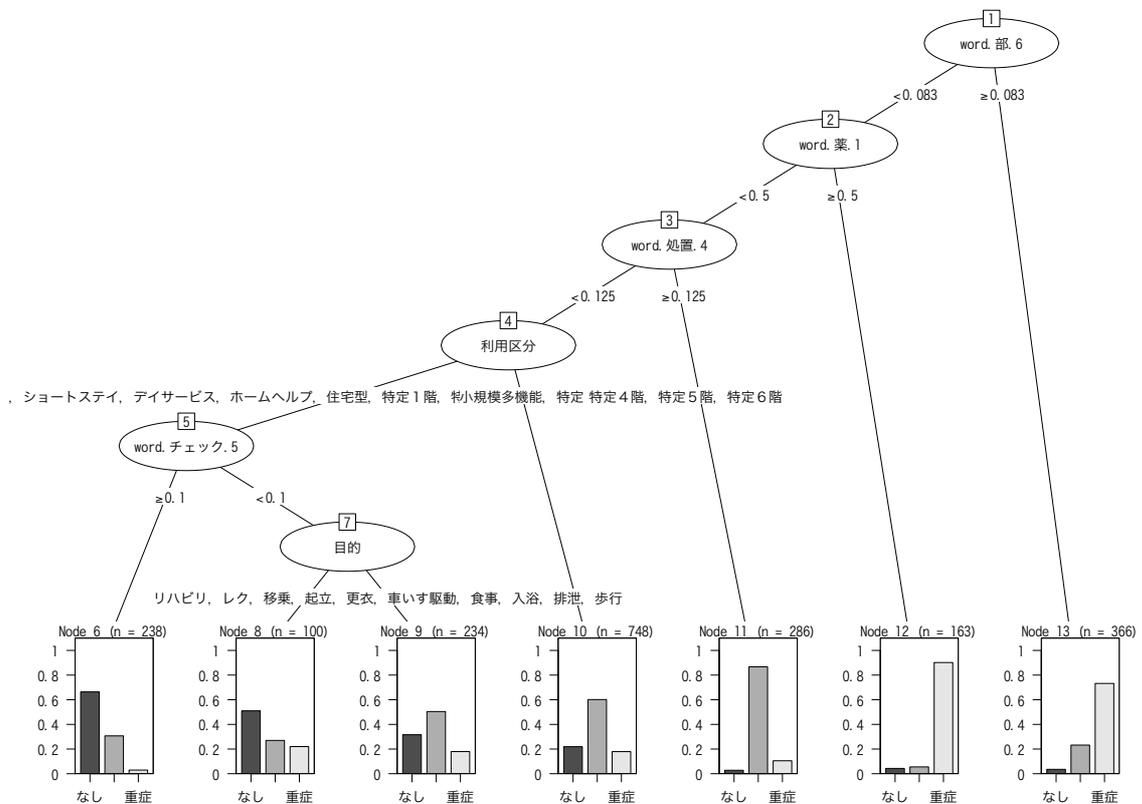


図 10 重症度を目的変数とした決定木学習

るクラスタリングを行い、2701 単語を 50 個のクラスタに分類を行ったが、1つのクラスに 2400 単語近くも分類されているクラスタがあった。これはその他の雑多な単語としてまとめられたものと考えられるが、2701 単語の内、約 2400 単語が1つのクラスタにまとめられると残りの約 300 単語が 49 のクラスタに分類されたことになってしまう。この問題の対策としては、事故詳細テキストから単語を抽出する際、複数の文章に含まれている単語のみを抽出することによって、分散が少なくなり解決されると考えられる。今回は2つ以上の詳細テキストに含まれている単語を抽出したので、最低含数を 2 より大きく設定する必要があると考える。

6. あとがき

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った。事故事例テキストから、事故詳細部分を取り出し、MeCab を用いて分かち書きを行った。それを元に形態素解析を行い、単語・文書行列を作成した。次に、k-means 法によるクラスタリングを行い、2701 単語を 50 個のクラスタに分類した。クラスタリングした単語・文書行列と事故事例テキストの属性変数を結合した多変量データを作成した。そして、この多変量データから目的変数を除いたデータを説明変数とし、ランダムフォレスト法を用い、目的変数である属性をそれぞれ推定した。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単語クラスタと属性の変数を用いて決定木を行い、属性の詳細を分析した。介護事故事例テキストに対して形態素解析を行い、事故の属性である「利用区分」、「事故の種類」、「診断」、「重症度」に対する事故の属性・単語の重要度を調べ、重要度の高い属性・単語を用い決定木を作成した結果、それぞれの目的変数と説明変数の関係がわかった。この結果、行動の目的と起きやすい事故の種類は関係が深い、事故の重症度と介護士の事故処理能力は関係があるといった過去の介護事故傾向がわかった。

謝辞

実験データの提供、そして本研究に際して、多くのご指摘を下さいました介護施設の皆様に感謝いたします。

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. 入手先 (<http://163.137.191.238/journal/0801/les/jn200801032.pdf>)
- [2] Jakkula, Vikramaditya, and Diane J. Cook. "Anomaly detection using temporal data mining in a smart home environment." *Methods of information in medicine* 47.1 (2008): 70-75.

- [3] 介護保険事業状況報告 月報(暫定版) | 厚生労働省入手先 (<http://www.mhlw.go.jp/topics/0103/tp0329-1.html>)
- [4] 林泰史. "高齢者の転倒防止." *日本老年医学会雑誌* 44.5 (2007): 591-594.
- [5] 岡部貴博, 吉川大弘, and 古橋武. "インシデントレポート解析のための多重接続型階層的テキストマイニング手法の提案." *日本知能情報ファジィ学会 ファジィシステムシンポジウム 講演論文集* 22.0 (2006): 54-54.