

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成25年 6月 7日現在

機関番号：24506

研究種目：若手研究（B）

研究期間：2011 ～ 2012

課題番号：23792550

研究課題名（和文） 看工連携による看護ケアテキスト自動評価及び評価根拠の可視化

研究課題名（英文） Classification and visualization for nursing-care text evaluation by the nursing and engineering cooperation

研究代表者

新居 学 (NII MANABU)

兵庫県立大学・大学院工学研究科・助教

研究者番号：80336833

研究成果の概要（和文）：看護の質向上のために、ネットを介して収集された看護ケアテキストをソフトコンピューティングおよび機械学習を用いて自動分類し、その評価規則を決定木により可視化した。研究成果により概ね50～80%の看護ケアテキストを4段階のクラスに正しく分類できた。また、テキストの構造を考慮した数値化表現により、決定木を用いて分類規則を人間に理解しやすいIf-thenルール形式で可視化できた。

研究成果の概要（英文）：In order to improve the nursing-care quality, we have proposed and developed a nursing-care text classification system. Our proposed system classifies the nursing-care texts using support vector machines and visualizes the classification rules using decision trees. The proposed system can classify 50% ~ 80% of nursing-care texts into four classes correctly. We also proposed a novel feature vector definition that represents structures of nursing-care texts. Using the above mentioned definition, we can easily understand classification rules generated by decision trees.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
交付決定額	2,700,000	810,000	3,510,000

研究分野：医歯薬学

科研費の分科・細目：看護学・基礎看護学

キーワード：看護の質向上，テキスト分類，ソフトコンピューティング，機械学習

1. 研究開始当初の背景

兵庫県立大学の看護学部において内布らは、看護ケアの質を評価するために以下に示すような研究を経て「看護過程データ」の評価尺度を定義し、データの収集に取り組んできた。

[1] 内布：看護 QI 研究会活動報告書 看護 QI プログラムによる第三者評価，看護管理，12(6)，416-421，2002.

[2] 上泉，内布，栗屋：看護ケアの質評価・改善の管理体制づくりに関する研究(厚生科研医療技術評価総合研究事業)2000～2002年度活動報告書。

これらの成果により、現在では看護師に評価尺度に従って看護過程を自由記述文として回答させ、Web アプリケーションにより収集、データベースへ電子的に記録している。このように、データを蓄積する手段は確立されている。

このように蓄積されたデータを利用して、「看護過程情報の分析による看護の質向上のための要因発見」(平成18，19年度科学研究費補助金(若手研究 B) 課題番号18700229)において、蓄積された自由記述文から看護過程の評価に必要な情報を抽出・分析するシステムの開発を行い、看護ケア Web 質評価総合システムにおいて、各看護

師の看護行為の評価を自動的に行えるシステムを開発してきている。

以上のような経過で、看護ケアの質評価を自動的に行うことのできる枠組みが完成し、従来の少数の専門家による評価では実現が困難であった多数の看護師の評価を行うことが可能となりつつある。しかしながら、看護ケア評価の各看護師へのフィードバックであるリコメンデーションの作成は、依然として専門家による手作業で行わざるを得ないという問題が残っている。また、看護のプロセスは日々刻々と変化するため、構築した自動分類システムを安定した分類性能を持つよう改善する必要がある。

2. 研究の目的

国内の看護ケアの質向上のために、看護ケア専門家がそれぞれの看護師が実践した内容をより良くするための示唆を行うことが重要である。専門家が各看護師に対して行う示唆をリコメンデーションと呼ぶ。リコメンデーションの作成は専門家の作業によるが、日本国内の看護ケアの質向上を目指す場合には、この支援システムの開発が必要不可欠である。本研究ではこのような大目的のもとで次の2点に重点をおいて研究を進める。

(1) リコメンデーション作成のためにはまず、収集された看護ケアテキストを正しく評価できなくてはならない。看護ケアテキストは、4段階のクラスに分類される。すなわち自動分類システムがこれら4つのクラスへ正確に分類できることが重要である。

これまでの研究でも分類性能向上のための取り組みを行ってきたが、本研究では、次の(2)で述べる評価根拠の可視化との親和性も考慮して、自由記述文である看護ケアテキストを数値ベクトル化する方法を考え、その有効性を分類性能と(2)で述べる可視化の観点から評価する。その結果として、分類性能と可視化の両面から、性能を両立させる方法を考え、看護ケアテキスト自動分類システムの分類性能向上を目指す。

(2) 看護ケアテキスト自動分類を一定の分類性能で行えると仮定する。このような仮定の下でリコメンデーション作成支援として必要な要件は、自動分類の結果として当該看護ケアテキストが分類された根拠・理由を専門家が知る手段を提供できることである。

これまで、識別に寄与している単語を重要語と位置づけて抽出し、評価者に提示することを試みたが、分類に大きく寄与する単語が必ずしも分類根拠として理解しやすいものではなかったため、分類根拠として理解しやすいIf-then形式の提示方法を考え、評価者の理解しやすい形式で評価根拠の可視化を

行う。

3. 研究の方法

(1) まず、テキストを数値ベクトル化(以降、特徴ベクトルと呼ぶ)するために、過去の回答に含まれる単語(名詞および動詞)を抽出し、単語リストと呼ぶ単語集合を作成する。単語の抽出には、MeCab[1]を使用する。

次に、評価対象とするテキストに対して、名詞と動詞を抽出する。抽出単語が単語リスト中に出現するならば1を、そうでないならば0を割り当てる。こうすることにより、すべての評価対象のテキストは、長さが単語リストの単語数と等しいベクトルとなり、それぞれのベクトルは各テキストに出現する単語に相当する属性値のみが1となる。従来はこのベクトルを用いてサポートベクターマシンにより分類を行っていた。

本研究では、係り受け解析を用いて単語間の係り受け関係も抽出し、この関係を特徴ベクトルとして表現することにより評価対象のテキストの構造を特徴ベクトルへ反映させる。係り受け解析にはCaboCha[2][3]を用いる。またこの時、従来使用しているサポートベクターマシンを変更なしに使用できるように入力の表現を特殊な構造に変更しないこととする。

このような方法で数種類の特徴ベクトルのバリエーションを生成し、2007年と2008年に収集された看護ケアテキストを学習用として使い、2009年収集の看護ケアテキストを評価用として分類性能の比較を行う。

[1] MeCab: Yet Another Part-of-Speech and Morphological Analyzer,

<http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

[2] チャンキングの段階適用による日本語係り受け解析, 工藤 拓, 松本 裕治, 情報処理学会論文誌, Vol. 43, No. 6, 1834-1842, 2002

[3] CaboCha,

<http://code.google.com/p/cabocho/>

(2) (1) で考案する複数の特徴ベクトル表現について、C5.0 と呼ばれる決定木を用いて木構造でIf-thenルールを抽出する。ルール抽出の目安として、サポートベクターマシンと同程度の分類性能が達成されることを前提とする。考案した複数種類の特徴ベクトル表現に対して、抽出されたIf-thenの木構造表現を比較し、理解のしやすさという観点からその評価を行う。

(3) 上記(1)および(2)において、同一の看護ケアテキストを複数の特徴ベクトル表現し、同一の分類システムまたは決定木

により分類性能または If-then ルール抽出を行うのであって、本研究で言う分類性能の向上または悪化は提案する特徴ベクトル表現に由来するものであり、サポートベクターマシンまたは決定木の性能を議論するのではないことに注意する。

4. 研究成果

(1) 表 1 に従来の特徴ベクトル表現による結果と、本研究の成果である係り受け構造を表現した特徴ベクトル表現による結果を示す。

表 1 分類性能の比較

テキスト集合名	従来法	提案法	総テキスト数
Q123	57.2%	59.5%	827
Q131	63.8%	65.1%	769
Q132	64.9%	62.4%	638
Q212	49.9%	50.8%	718
Q213	53.7%	55.9%	735
Q221	47.8%	51.7%	768
Q222	77.8%	79.9%	703
Q322	69.4%	71.0%	604
Q411	80.9%	80.9%	566
Q423	85.5%	81.6%	636
Q425	66.8%	62.9%	690
Q431	81.9%	81.9%	659

なお、2012 年に上記の各テキスト集合に対して、専門家の評価の揺れや年度ごとの評価基準の多少の違いを修正したため、その前後で分類性能に違いが出ていることを付記しておく。

一部のテキスト集合においては性能の悪化が見られる。これらの場合は、テキストの評価基準として単語そのものに重要性を見出しているのか、複数の単語関係からなる表現が重要なのかの違いも考慮しなければならないと考察している。

結果から明らかなように、係り受け関係の特徴ベクトルに取り入れた提案手法の分類性能は、従来法に比べて向上しているため、本研究の目的を達成しているといえる。

表 1 において高い分類性能で、かつ、従来法と提案法に違いがないものは、元のデータにおいてクラスの偏りが非常に大きいものである。このようなデータは、本研究における取り組みでは性能向上が難しい。このようなデータに対する分類性能向上は今後の課題である。

(2) 図 1 に従来法による特徴ベクトルの場合の決定木の一部を、図 2 に提案法による特徴ベクトルの場合の決定木の一部をそれぞれ示す。抽出される決定木は、上部左側から下に向かって If-then 構造をとる。たとえば、図 1 の場合、「コンタクト」という単語があ

るかどうかを見て、なければ「協議」という単語の存在に着目することを意味している。

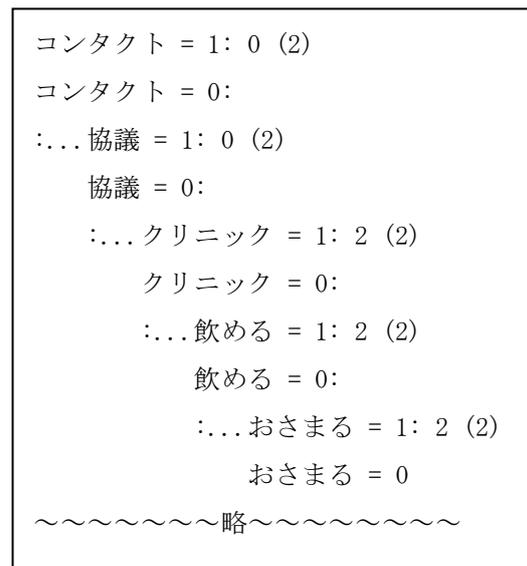


図 1 従来法での決定木の一部

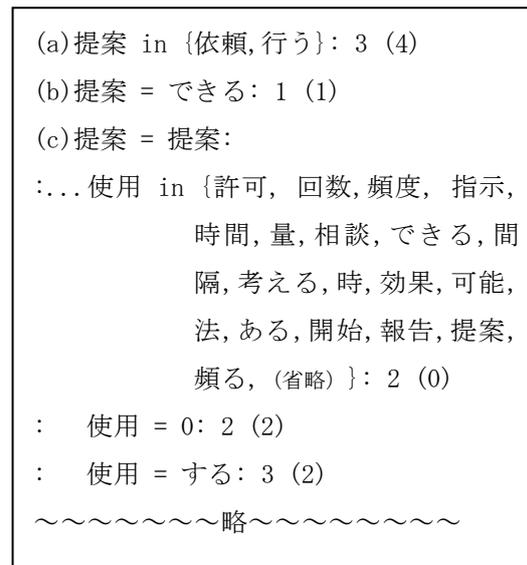


図 2 提案法での決定木の一部

図 2 では(a)「提案」が「依頼」または「行う」へ係る場合、(b)「できる」へ係る場合、(c)「提案」は係り先を持たないが「使用」が {} の中に示される単語に係る場合等が表現されている。この結果より、図 1 のような結果よりも図 2 のような結果が評価対象テキストの内容をよく表現できているため、本研究の目的である分類根拠の理解しやすい形式での表現が達成できているといえる。

決定木での分類性能はサポートベクターマシンのそれよりわずかに劣るため、図に示したようなルール表現をサポートベクターマシンから抽出するか、もしくは決定木の分類性能の向上を図ることが今後の課題であ

る。さらに、図1および2の表現でも、直感的に理解しやすいとはいえないため、例えば看護師個人個人がこのような評価根拠を目にする場合を考慮し、人間に理解しやすい表現方法を追求する必要もあると考えている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計2件)

(1) Manabu Nii, Yoshinori Hirohata, Atsuko Uchinuno, Reiko Sakashita, Dependency relation based feature vectors for classifying nursing-care texts, International Journal of Intelligent Computing in Medical Sciences & Image Processing, 査読有, 2013, doi:10.1080/1931308X.2013.799319

(2) Manabu Nii, Takafumi Yamaguchi, Yutaka Takahashi, Reiko Sakashita, Atsuko Uchinuno, Genetic Algorithm Based Term Selection for Nursing-Care Freestyle Japanese Text Classification, International Journal of Intelligent Computing in Medical Sciences and Image Processing, 査読有, Vol. 4, No.2, pp.119-126, 2011

[学会発表] (計5件)

(1) Manabu Nii, Yutaka Takahashi, Atsuko Uchinuno, Reiko Sakashita, An Approach using Conceptual Fuzzy Sets for Nursing-care Text Classification Proc. of World Automation Congress 2012, WAC-2012-1569535307, 2012

(2) Manabu Nii, Yoshinori Hirohata, Atsuko Uchinuno, Reiko Sakashita, New Feature Definition for Improvement of Nursing-care Text Classification, 2012 IEEE International Conference on Systems, Man, and Cybernetics, pp.2610-2615, 2012

(3) Manabu Nii, Yoshinori Hirohata, Atsuko Uchinuno, Reiko Sakashita, Feature Definition using Dependency Relations between Terms for Improving Nursing-care Text Classification, Proc. of 2012 Fifth International Conference on Emerging Trends in Engineering & Technology, pp.110-115, 2012

(4) Manabu Nii, Takafumi Yamaguchi, Yusuke Mori, Yutaka Takahashi, Atsuko Uchinuno, Reiko Sakashita, Nursing-care

Text Classification using Additional Term Information from Web, Proc. of 2011 IEEE International Conference on Fuzzy Systems, June 27-30, 2011, Taipei, Taiwan, pp.1442-1446, 2011

6. 研究組織

(1) 研究代表者

新居 学 (NII MANABU)
兵庫県立大学・大学院工学研究科・助教
研究者番号：80336833

(2) 研究協力者

片田 範子 (KATADA NORIKO)
兵庫県立大学・看護学部・教授
研究者番号：80152677

上泉 和子 (KAMIIZUMI KAZUKO)
青森県立保健大学・健康科学部・教授
研究者番号：10254468

内布 敦子 (UCHINUNO ATSUKO)
兵庫県立大学・看護学部・教授
研究者番号：20232861

坂下 玲子 (SAKASHITA REIKO)
兵庫県立大学・看護学部・教授
研究者番号：40221999

鄭 佳紅 (TEI KEIKO)
青森県立保健大学・健康科学部・准教授
研究者番号：20363723