

キャラクターを用いたテキストマイニングスキルの獲得支援 Assistance for Acquiring Text Mining Skills Using Characters

後藤 賢悟¹ 砂山 渡² 川本 佳代¹
Kengo Goto¹ Wataru Sunayama² Kayo Kawamoto¹

¹ 広島市立大学大学院情報科学研究科

¹ Graduate School of Information Sciences, Hiroshima City University

² 滋賀県立大学 工学部

² School of Engineering, The University of Shiga Prefecture

Abstract: 近年、膨大なテキストデータから有用な情報を取り出すテキストマイニングの技術が注目され、多数のマイニングツールが開発されている。しかし、データの分析を行うことができる人は少なく、マイニングの初心者がスムーズに分析を行える環境が求められている。そこで本研究ではテキストマイニングツール TETDM の機能として、初心者ユーザを対象に、テキストマイニングスキルの獲得を目的としたチュートリアルシステムを提案する。

1 はじめに

近年、ネットワークや記憶装置の発達により、世の中には膨大な量のテキストデータが蓄積されている。そこで、それらの大量のテキストデータから有用な情報を取り出すテキストマイニングの技術が注目を集めており、現在までに多数のテキストマイニングツールやソフトウェアが開発されている。しかし、現状ではデータの分析を行うことができる人間は少ない。そこで、これまでテキストマイニングを行ってこなかった人が、スムーズにツールを使って分析を行える環境を構築することが重要となってくる。

テキストマイニングのソフトウェアの一つ、TETDM (Total Environment for Text Data Mining: テキストマイニングの統合環境) [1] では、利用者のテキストマイニングの実力に応じて「モード」を変更できることから、テキストマイニングの知識がない初心者にとっての入門ツールとして気軽に利用できることが期待される。しかしながら、初心者ユーザは自力でテキストマイニングのスキルを獲得することは非常に難しい。

そこで本研究では TETDM の機能として、初心者ユーザを対象とした、アシストキャラクターを用いたチュートリアルシステムを構築することで、ユーザのテキストマイニングスキルの獲得と向上を目指す。

2 関連研究

キャラクターを使用して、ユーザーの支援を行う研究として、益子らの「キャラクターを用いた学習継続

支援ツールの開発」[2] や、渡辺らの「セルフレARNING型授業におけるエージェントキャラクターによる学習支援」[3] がある。これらのアシストキャラクターの存在は、ユーザのシステムに対する利用意欲を高めると共に、ユーザの学習スキル向上に大いに貢献していることがわかる。実際、ここで取り上げた以外にも、多くの学習システムにはこのようなアシストキャラクターによる支援システムが搭載されている。

しかしながら、テキストマイニングの分野においては、このようなアシストキャラクターを使ったユーザのテキストマイニングスキル向上支援を行う研究は行われていない。

数あるテキストマイニングツールの中で、ユーザのテキストマイニングスキルの獲得を目的としたチュートリアルを搭載したツールはほとんど存在しない。データマイニングツールではあるが、チュートリアルに近い機能を搭載している KEEL [4] というツールがある。KEEL はデータマイニングのために、回帰、分類、クラスタリング、パターンマイニングのようなアルゴリズムを実装したツールである。KEEL では、実際にデータマイニングを行うときに利用する一連の機能をまとめた実験を設計する機能が提供されている。初めてデータマイニングを行うとき、ユーザはアルゴリズムがどのように動作するかを学ぶ必要がある。そのための教育のための機能が搭載されており、ユーザはアルゴリズムのリアルタイムビューに着目しながら短期間で実験を実施することによって、アルゴリズムの違いを学ぶことができる。本研究でも同様に、テキストマイニングのスキルを獲得を目的とした練習を短時間で繰り返し練習を行うことができるチュートリアルを実装する。

3 TETDM と利用モード

TETDM(Total Environment for Text Data Mining) は、テキストデータマイニングのための統合環境として作成された。TETDM では豊富なツールが提供されており、それらのツールを使用することで利用者は文章のより詳細な分析を行うことができ、新たな発想が得られる可能性がある。現在、最新版として TETDM サイト内 [5] で公開されている。

TETDM には、利用者のテキストマイニングの実力に応じて、以下の4つの起動モードを切り替えることができる。初めて TETDM を利用する人は、最初はスーパーライトモードから利用し、慣れたらライトモード、通常モードと順に利用モードを切り替えていくことを推奨している。

- スーパーライトモード
 初心者が最初に利用出来るモード。あらかじめ用意された8つのツールセットを利用出来る。
- ライトモード
 少しだけツールを操作してみたい人向けのモード。
- 通常モード
 いろいろなツールを操作してみたい人向けのモード。より汎用性がある実用的な分析ができる処理ツールと可視化ツールが選択できるようになる。
- 拡張モード
 より詳細な分析をしたい人向けのモード。専門的な内容を含む利用を想定した処理ツールと可視化ツールを選択できるようになる。

4 テキストマイニングのプロセスとスキル

本研究では、テキストマイニングを行う際に、ユーザが最終的に知識を獲得するまでの理想的なテキストマイニングプロセスを図1のように定義する。以下で、テキストマイニングの各プロセスの内容と各プロセスで必要とされるスキルについて述べる。

4.1 テキスト収集&テキスト整形

テキスト収集は、ユーザが実際に分析したいテキストを準備するプロセスである。テキスト整形は、分析したい内容に関するテキストを準備し、そのテキストをシステムに入力可能な状態にするプロセスである。例としては、テキストの段落分けや、句点の挿入処理、キー

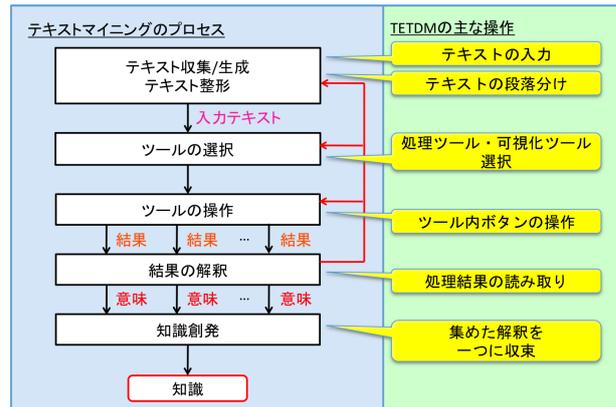


図 1: テキストマイニングのプロセス

ワード設定などが挙げられる。TETDM ではテキスト収集はテキストファイルを用意するか、TETDM 上でコピー&ペーストするかどちらかで達成できる。テキスト整形は段落の区切りとなる場所を空行としておけば、ボタン一つで整形可能となっている。

テキスト収集プロセスでは、テキストを集める時に、分析の目的に応じてどのようにテキストを集めてくればいいのかをユーザが把握するスキルが必要となる。テキスト整形のプロセスについては、集めたテキストの量や、分析の目的に応じた段落分けを行うスキルが必要となる。また、テキスト整形では扱う品詞の設定ができるスキルも必要となる。文章中に出てくるものや人物などを知りたい場合は名詞、文章中の行動や動作を示した単語を特定したいなら動詞、大小や高低などの状態を知りたい場合は形容詞をそれぞれキーワードとして設定することが必要だ。

4.2 ツール選択

ツール選択は分析の目的に応じたテキストマイニングツールを選択するプロセスである。TETDM では処理ツール(入力されたテキストに対し、テキストマイニング処理を行うツール)と可視化ツール(処理ツールによるテキストマイニング処理の結果を画面へ出力するツール)がある。それらを組み合わせることでツールのセットを行うことで達成できる。現在の TETDM(Ver2.21) では処理ツールが46種類、可視化ツールが35種類あるため、様々な組み合わせを試してツールの役割を覚えることが重要となる。

ツールの選択プロセスでは、分析の目的に適したツールを選択できるスキルが必要となる。その為には、各ツールを選択することでどのような処理結果を得ることができるかを把握する必要がある。

4.3 ツール操作

ツール操作は有意義な結果を得るため、ツールを操作するプロセスである。TETDMでは、ツールの選択後に処理ツール内や可視化ツール内に配置されているボタンを押すなどの操作が当てはまる。

ツールの選択プロセスでは、操作の必要があるツールを選択した場合、分析の目的に適したツールの操作ができるスキルが必要となる。その為には、各ツールを操作することでどのような処理結果を得ることができるかを把握する必要がある。

4.4 結果の解釈

結果の解釈プロセスでは、ツールの出力結果（評価値）と、それが表す意味の解釈（どのような点が良いのか・どのように修正するのか）を行う。解釈は一度だけ行えば終わりではなく、場合によっては他のツールを選んだり、先ほどとは異なる操作を行うなど、別の視点から新たな解釈を行い、複数の解釈を集める必要がある。TETDMでは各パネルに結果と解釈を登録できるフォームがあり、このフォームに結果と解釈を重要度（結果と解釈が自分にとってどのくらい重要なのか）と共に記述することで達成できる。

結果の解釈プロセスでは、処理結果のどこに着目するか、どのように結果を読み解いていくかについてのスキルが必要となる。また、次の知識創発プロセスで知識創発を行うためには、多くの解釈を集める必要があるため、場合によっては再度キーワード設定やツールの選択・操作を行う必要がある。そのため、次にどのような解釈を集めるべきか、そのためにはどのように行動すればいいかを考えるスキルも必要となるだろう。

4.5 知識創発

知識創発は得られた解釈を一つにまとめることによって大局的な知識を得るためのスキルとする。信頼性が高い知識を得るためにはより多くの解釈をまとめることが重要となる。TETDMでは登録した解釈を一つにまとめることができる知識創発ウインドウを利用することで達成できる。

知識創発プロセスでは、集めた解釈をまとめて、より多角的な視点から見た説得性の高い解釈へとまとめ上げることができるスキルが必要となる。そのためには、集めた解釈をどのように結合していくのかをどの順番で、どのように筋道を立てて解釈をまとめていくかが重要となる。



図 2: キャラクターの吹き出しによる説明文の表示



図 3: 指アイコンでのボタン提示のアシスト

5 キャラアシストチュートリアル

本章では、初心者のテキストマイニングスキルの獲得をアシストする、チュートリアルシステムについて述べる。

5.1 チュートリアルの枠組み

本チュートリアルは、「TETDM の操作方法の説明」と「テキストマイニングの流れ」の2つの課題群から構成されている。このうち、テキストマイニングのスキルに関わるのは後者となり、スキル獲得のための前提となる操作説明を別途設ける。

テキストマイニングスキル獲得を目指したチュートリアルを行う前に、まずはTETDMの基本的な操作方法をユーザに習得してもらう必要がある。「TETDMの操作方法の説明」のチュートリアル課題群では、TETDMの環境の操作説明チュートリアルである「操作の説明」と、TETDMで選択して利用出来るテキストマイニングツールの利用方法のチュートリアルである「ツールの説明」の2つの課題群に課題を分類した。環境の説明とツールの説明の課題を分けた理由は、ツールの説明に関してはツールの利用方法がわからない時に確認したいツールの説明のチュートリアルだけを確認できるようにするためである。また、TETDMは多くの開発者がツールの開発を行っており、仮にツールがこれから増えた際、チュートリアルもそれに依って拡張しなければならぬが、ツールの説明のチュートリアルだけが独立していることで、課題の拡張が容易であるというメリットもある。

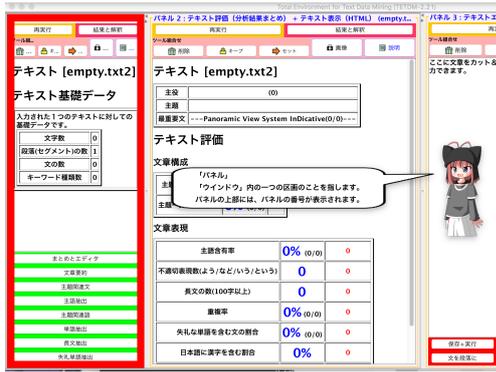


図 4: 四角でウィンドウを囲むアシスト

TETDM の利用方法のチュートリアル課題を終え、環境やツールの使い方を習得したユーザを対象に、テキストマイニングスキルの獲得を目指したチュートリアル「テキストマイニングの流れ」を行なってもらう。本チュートリアルの課題は、テキストマイニングのプロセスの流れに則り、実践的な練習を行うことで、ユーザにテキストマイニングの流れと、各プロセスにおけるマイニングスキルの獲得を促す。

5.2 キャラクターによるアシスト内容

5.2.1 吹き出しを用いた説明文の表示

本チュートリアルにおける説明文は、キャラクターの台詞という形でキャラクターの左側に表示(図 2)される。説明文は、最大で 1 行 30 文字 × 3 行まで表示され、ユーザは吹き出しをクリックすると次の説明文が表示される。吹き出しを右クリックした場合は、一つ前の説明文に戻るようになっている。また、ユーザによっては吹き出しを連打して説明文を読まずにチュートリアルを進める恐れもあったため、新たに表示された説明文は説明文 1 行につき 0.4 秒間はクリックしても次の説明文が表示されない仕様になっている。また、キャラクターと吹き出しはドラッグ&ドロップで好きな位置に移動が可能のため、仮にボタン上にキャラクターや吹き出しが現れた場合は簡単に移動できる。キャラクターを移動したら、移動後に表示される吹き出しは移動したキャラクターに追従して表示されるようになる。仮にキャラクターを画面左に移動した場合、吹き出しは画面外に出ず、キャラクターの右側に表示されるようになっている。

表 1: 「TETDM の操作方法 (環境)」スーパーライトモード用課題

課題番号	課題名
1	はじめに
2	TETDM の概要
3	スーパーライトモードでできること
4	TETDM の各部の説明
5	TETDM のボタンの説明
6	TETDM の用語解説
7	ゲームモードの説明
8	テキストマイニングの流れ
9	テキストの入力方法
10	テキストの段落分け
11	ツールの選択方法

表 2: 「TETDM の操作方法 (ツール)」スーパーライトモード用課題

課題番号	課題名
1	まとめとエディタ
2	文章要約
3	主題関連文
4	主語抽出
5	主題関連語
6	単語抽出
7	長文抽出
8	失礼単語抽出

5.2.2 指アイコンおよび枠囲みによるボタン等のポインタ表示

キャラクターの台詞による説明では、文章ではわかりにくく、説明が難しい内容も多い。例えば、ボタンの位置や各種マイニングツールの操作方法などの説明を行う場合がそれに当たる。そこで、ボタン位置等を指アイコンを用いてユーザに提示するアシストの導入(図 3)を行う。このようなアシストは、TV ゲームのチュートリアル画面などではよく見られるが、ツールやアプリケーションの使い方の説明にはあまり用いられていない手法である。また、ウィンドウ全体を示すなど、指アイコンでは説明の限界がある場合は、ウィンドウを四角で囲む(図 4)アシストを行う。

5.3 TETDM の操作方法についてのチュートリアル課題

本節では TETDM の操作方法についてのチュートリアル課題について、各起動モードで利用出来る機能とともに述べる。

5.3.1 スーパーライトモード

スーパーライトモードでは、テキストエディタにテキストをコピー&ペーストすることでのテキストの入

表 3: 「TETDM の操作方法 (環境)」ライトモード用課題

課題番号	課題名
1	ライトモードでできること
2	TETDM のボタンの説明 2
3	TETDM のボタンの説明 3
4	キーワード設定について
5	句点の挿入
6	キーワード設定 (1)
7	キーワード設定 (2)
8	キーワード設定 (3)
9	アプリボタンによるツールの選択

表 4: 「TETDM の操作方法 (ツール)」ライトモード用課題

課題番号	課題名
1	2 ちゃんまとめ
2	タイピング

力や、テキストの段落分けの機能など、簡単な機能のみが扱える。また、選択できるツールは、ツールは処理ツールと可視化ツールがあらかじめ組み合わせられたテキスト評価アプリケーションの 8 つのツールセットを利用出来る。チュートリアルも、この内容に沿って作成を行う。

スーパーライトモードでのチュートリアル課題は、TETDM の環境の操作説明の課題については表 1 に示す 11 題、ツールの使い方の課題については表 2 に示す 8 題を用意している。チュートリアルでは、サンプルテキストを用いて、環境の基本的な操作方法やツールの利用方法を学んでいく。

5.3.2 ライトモード

ライトモードでは、新たにキーワード設定の機能が追加される。また、選択できるツールセットも、新たに 2 種類増える。ライトモードでのチュートリアル課題は、TETDM の環境の操作説明の課題については表 3 に示す 9 題、ツールの使い方の課題については表 4 に示す 2 題を用意している。

5.3.3 通常モード

通常モードでは、テキストファイルを選択することでのテキストの入力や、処理ツールと可視化ツールを自由に組み合わせるツールの選択方法が新たに追加されており、ツールの種類数も処理ツール 27 種類 + 可視化ツール 18 種類と、一気に増加している。また、新たに処理結果と解釈の登録機能と、知識創発の機能が追加された。通常モードの操作方法のチュートリアルも、

表 5: 「TETDM の操作方法 (環境)」通常モード用課題

課題番号	課題名
1	通常モードでできること
2	TETDM のボタンの説明 4
3	テキストの入力方法 2
4	ツールの選択方法 2
5	ツールの選択方法 3
6	結果と解釈の登録
7	知識創発 1
8	知識創発 2

表 6: 「TETDM の操作方法 (ツール)」通常モード用課題

課題番号	課題名
1	単語情報まとめ
2	文情報まとめ
3	セグメント情報まとめ
4	国語辞書
5	単語間関連度
6	単語冗長文抽出
7	類似文抽出
8	再帰的クラスタリング
9	漢字の割合
10	タグクラウド

この内容に沿って展開され、サンプルテキストを利用した練習が可能である。

通常モードでのチュートリアル課題は、TETDM の環境の操作説明の課題については表 5 に示す 8 題、ツールの使い方の課題については表 6 に示す 10 題を用意している。なお、通常モードで行われるツールの説明の課題は、複数のツールを組み合わせる利用することが前提のツールについては一つの課題で組み合わせられる全てのツールについての説明を行っている。

5.3.4 拡張モード

拡張モードでは、複数のテキストファイルを同時に入力したり、テキストの一部分を部分テキストに設定して部分的にテキストマイニングできる機能などが追加された。ツールについては、処理ツール 46 種類 + 可視化ツール 35 種類が選択可能で、分析の用途ごとに用意されたツールセットを搭載したアプリケーション 4 種類が新たに選択可能になっている。

拡張モードでのチュートリアル課題は、TETDM の環境の操作説明の課題については 6 題、ツールの使い方の課題については 17 題用意している。なお、拡張モードで行われるツールの説明の課題は、アプリケーションで選択できるものには課題名に通し番号が入力されている。

表 7: 「TETDM の操作方法 (環境)」拡張モード用課題

課題番号	課題名
1	拡張モードでできること
2	TETDM のボタンの説明 5
3	テキストの入力方法 3
4	ツールの選択方法 4
5	ツールの選択方法 5
6	部分テキストの設定方法

表 8: 「TETDM の操作方法 (ツール)」拡張モード用課題

課題番号	課題名
1	1-1: テキスト集合評価アプリケーション
2	1-2: テキスト分類
3	1-3: テキスト間類似度 A
4	1-4: テキスト間類似度 B
5	1-5: レポート評価
6	2-1: 部分テキスト作成アプリケーション
7	3-1: 文章推敲アプリケーション
8	3-2: 主題語含有率
9	3-3: 段落間類似度表示
10	3-4: 段落間木構造
11	4-1: プロフィールチェックアプリケーション
12	形態素解析
13	意見文抽出
14	単語置換
15	辞書再構築
16	単語抽出
17	段落の並び替え

表 9: 「テキストマイニングの流れ」スーパーライトモード用課題

課題番号	課題名
1	はじめに
2	テキストマイニングの流れ
3	テキストを集めよう
4	テキストを整形してみよう
5	ツールを選択して結果を読み取ろう (1)
6	ツールを選択して結果を読み取ろう (2)
7	ツールを選択して結果を読み取ろう (3)
8	ツールを選択して結果を読み取ろう (4)
9	ツールを選択して結果を読み取ろう (5)
10	ツールを選択して結果を読み取ろう (6)
11	集めた解釈をまとめてみよう
12	処理結果を画像処理しよう

表 10: 「テキストマイニングの流れ」ライトモード用課題

課題番号	課題名
1	はじめに
2	テキストマイニングの流れ
3	キーワード設定を試みよう
4	簡単なツールの操作を試みよう (1)
5	簡単なツールの操作を試みよう (2)
6	解釈のまとめ

発の機能は通常モードから利用出来る機能なので、解釈を集めるところまでは行わない。

5.4 テキストマイニングのスキル獲得に向けたチュートリアル課題

本節ではテキストマイニングの流れのチュートリアル課題について、各起動モード毎に述べる。

5.4.1 スーパーライトモード

スーパーライトモードでのスキル獲得を目指したテキストマイニングの流れのチュートリアル課題は、表 9 に示す 12 題を用意している。スーパーライトモードでは、最初にテキストマイニングの全体的な流れの説明を行った後、テキスト収集&テキスト整形プロセスから結果の解釈プロセスまでの流れを練習できる。

「テキスト収集&テキスト整形」プロセスでは、テキストの効果的な集め方や目的に応じた段落分けの方法が学べる。ツール選択プロセスは、スーパーライトモードで扱える 8 つのツールセットを、どのような場面で選択すれば良いかを学べる。ツールの操作プロセスは、スーパーライトモードはツールの操作を基本的には行わないようにしたので、ここでの説明は省いている。結果の解釈プロセスでは、ツールの処理結果の読み取り方を学べる。結果と解釈の登録機能と知識創

5.4.2 ライトモード

ライトモードでのスキル獲得を目指したテキストマイニングの流れのチュートリアル課題は、表 10 に示す 6 題を用意している。ライトモードでは、新たに追加されたキーワード設定の機能を軸に、スーパーライトモードで学んだテキスト収集&テキスト整形プロセスから結果の解釈プロセスまでの流れを補足する形で学んでいく。また、ライトモードでは、スーパーライトモードでは行わなかったツールの操作を行い、処理結果がどのように変化するかを学ぶことができる。

5.4.3 通常モード

通常モードでのスキル獲得を目指したテキストマイニングの流れのチュートリアル課題は、表 11 と表 12 に示す 18 題 (9 題 2 パターン) を用意している。通常モードでは、結果と解釈の登録機能と、知識創発インタフェースが新たに利用可能になったことから、本格的にテキストマイニングの全プロセスを通した分析の練習が可能である。

表 11: 「テキストマイニングの流れ」通常モード用課題

課題番号	課題名
1	はじめに
2	テキストマイニングの流れ
3	テキストファイルからのテキスト入力
4	テキストの大筋を見ていこう
5	背景語を見てみよう
6	単語の出現頻度を見てみよう
7	単語の出現頻度を見てみよう(2)
8	再帰的クラスタリング
9	集めた解釈から知識創発をしてみよう

表 12: 「テキストマイニングの流れ」通常モード用課題(別パターン)

課題番号	課題名
1	はじめに
2	テキストファイルからのテキスト入力
3	テキストの大筋を見ていこう
4	主題の一貫性を見てみよう
5	主題の一貫性を見てみよう(2)
6	不適切表現の有無を見てみよう
7	単語冗長文の有無を見てみよう
8	集めた解釈から知識創発をしてみよう
9	知識創発で原因と結果を把握しよう

6 キャラアシストチュートリアル の有効性検証実験

6.1 実験内容

本節では実験内容について述べる。本実験の目的は、TETDM キャラアシストチュートリアル利用者のテキストマイニングスキル獲得・向上の検証を行うことである。被験者は、TETDM の利用経験がなく、なおかつテキストマイニングの経験がない大学生、大学院生の男女計 16 名とした。16 名の被験者のうち、8 名は今回作成したキャラアシストチュートリアルのうち、「TETDM の操作方法」「テキストマイニングの流れ」の両方の課題に取り組む提案群とした。残りの 8 名を比較群とし、「TETDM の操作方法」のチュートリアル課題のみに取り組み、テキストマイニングの流れに関しては、テキストマイニングの各プロセスについての操作方法やアドバイスを記入した紙媒体を元に、実際に TETDM を操作しながらテキストマイニングの流れについて練習するという形をとった。

実験期間は 3 日間で、最初の 2 日間で課題を行ってもらい、最終日には事後テストを行った。事後テストは、提案群・比較群同様に同じテキストを TETDM で分析してもらった。事後テストに用いたのテキストは、ポータルサイト Amazon で販売される nikon 製カメラのレビューを収集したものである。事後テストの内容は、被験者は nikon とは別の会社の社員であると仮定して、ライバル会社 nikon に勝つためのカメラのアイデ

表 13: 根拠に基づいたカメラのアイデア数(被験者平均)

提案群	比較群
3.1	1.5

表 14: キーワード設定回数(被験者平均)

提案群	比較群
2.3	0.6

アを nikon のカメラのレビューを TETDM を用いて分析することで考えるというものである。

6.2 実験結果と考察

本節では実験結果と考察について述べる。

最初に、事後テストで被験者に考えてもらったカメラのアイデアの中で、根拠に基づいたアイデアの個数についての比較を行った結果を表 13 に示す。t 検定の結果、両軍のアイデア数には有意な差があると確認できた。このことから、提案群のほうが分析を行った根拠に基づいたアイデアが多い傾向にあり、テキストマイニングのスキル獲得は提案群のほうが優れていたことがわかる。

また、各被験者の分析手順の差を比較するために、事後テスト後に回収した TETDM の操作履歴のログデータから、テキストマイニングの各プロセスにおける、提案群と比較群のテキストマイニングの手順の差を比較した。

テキスト収集&テキスト収集プロセスでの両群のキーワード設定を行った回数の差を表 14 に示す。キーワード設定回数は、提案群のほうが比較群よりはるかに多い結果となった。

ツールの選択プロセスでの、ツールの選択回数を表 15 に、ツールの操作プロセスの、ツールの操作回数を表 16 にそれぞれ示す。この結果から、提案群のほうがより多くツールの選択や操作を行ったことが確認出来る。

提案群のキーワード設定やツールの選択、操作の回数が増えた要因には、テキストマイニングプロセスの実践に関して紙で説明されるよりも、チュートリアルによって実際に試す機会を与えることで各プロセスでの分析や操作をより多く行っていたからだと考えられる。これらの操作回数増加は、より良い分析結果を得るための試行錯誤の回数が増加したと考えることができる。

結果の解釈プロセスでの、TETDM のツールの処理結果と解釈の登録機能を用いて解釈をまとめた個数の

表 15: ツールの選択回数 (被験者平均)

提案群	比較群
50.4	30.1

表 17: 結果と解釈の登録数 (被験者平均)

提案群	比較群
8.1	3.5

表 16: ツールの操作回数 (被験者平均)

提案群	比較群
32.5	11.5

表 18: 知識創発を行った被験者数

	提案群	比較群
1つの結論にまとめた	3	1
複数の結論にまとめた	3	2
知識創発を行っていない	2	5

データを表 17 に示す。この結果から、提案群のほうがより多く結果と解釈を登録したことがわかる。

このような結果となった理由は 2 つ考えられる。まず一つ目は、提案群はキーワード設定やツールの選択・操作でより良い処理結果を集めようと試行錯誤を何度も繰り返した結果、理想的な処理結果を集めることができたために、解釈の登録数も増加したと考えられる。また、二つ目の理由としては、提案群はテキストマイニングスキル獲得のためのチュートリアル課題で何度も実践練習を行ったことで、処理結果を読み取って自分なりにまとめる力が比較群よりも向上していることも考えられる。従って、結果の解釈プロセスでは提案群のほうがテキストマイニングスキルを獲得できたと言える。

知識創発プロセスでの、提案群と比較群で各ユーザが知識創発を行ったかについてを、表 18 に示す。この結果から、提案群のほうが知識創発の機能を用いて集めた解釈をまとめようとしていたことがわかる。このような結果になった理由として、提案群は多くの解釈が集まり、自分の頭だけでは解釈をまとめきれなくなった結果、積極的に TETDM の知識創発機能を用いて解釈をまとめるように意識することができたからだと考えられる。従って、知識創発プロセスでは、提案群のほうがテキストマイニングスキルを獲得できたと言える。

7 結論

本研究では、TETDM の機能の一つとして、TETDM 利用者のテキストマイニングスキルの向上させるためのキャラクターを用いたチュートリアルシステムの開発を行った。その後、評価実験により本システムが TETDM を始めて利用するユーザのテキストマイニングスキルの獲得・向上に有効であることを検証した。

今回はどんな分析の目的のテキストマイニングにも通用する基本的なスキルを獲得することを目的としたチュートリアルの構成を行ったので、今後は分析の目的に応じた、より専門的なスキルを獲得できるチュートリアル課題の拡張なども行っていきたい。

参考文献

- [1] 砂山渡・高間康史・西原陽子・梶並知記・串間宗夫・徳永秀和: 統合環境 TETDM を用いたマイニングツールの開発と利用の実践, 人工知能学会論文誌, Vol.29, No.1, pp.100-112, (2014)
- [2] 益子博貴, 松村敦, 宇陀則彦: キャラクターを用いた学習継続支援ツールの開発, ARG 第 1 回 Web インテリジェンスとインタラクション研究会 (ARG SIG-WI2), (2012)
- [3] 渡辺博芳, 水柿恵: セルフラーニング型授業におけるエージェントキャラクタによる学習支援, 情報科学技術フォーラム一般講演論文集 (N. 教育・人文科学), Vol.3, No.4, pp.335-337, (2004)
- [4] J.Derrac, J.Luengo, Alcal'a-Fdez, A.Fern'andez, S.Garc'la: Using KEEL-Software as a Educational Tool: A Case of study Teaching Data Mining, Next Generation Web Services Practices (NWeSP), 2011 7th International Conference on, Salamanca, pp.464-469, (2011)
- [5] TETDM:(URL) <http://tetdm.jp/>
- [6] 中垣内李菜, 川本佳代, 砂山渡: 統合環境 TETDM を用いたテキストマイニングにおける初心者のためのツール選択支援, 第 27 回人工知能学会全国大会, 3B3-NFC-01a-1, (2013)
- [7] キャラクターなんか機:(URL) <http://khemix.sakura.ne.jp/download.shtml>