

テキストの形態パターン情報を活用した汎用型将来動向予測支援システムの開発

中島 陽子¹ ミハウ プタシンスキ² 本間 宏利¹ 榎井 文人²

Development of a versatility future trend prediction support system using morphosemantic patterns information of text data

Yoko NAKAJIMA, Michal PTASZYNSKI, Hirotoishi HONMA, Fumito MASUI

Abstract:In this research, we developed a versatile future trend prediction support system using future reference sentences extracted automatically from news articles. To achieve that, we firstly developed a method to classify and extract implicit and explicit future reference sentences without the use of simple keywords, and confirmed that it is possible to generate a classification model with a small amount of useful learning data by using a Language Combinatorics-based learning method. Next, we applied automatic future reference extraction to develop a method to support future trend prediction using only future reference sentences, with no need for expert knowledge. The usefulness of the system as a whole was verified in experiments in which sentences referring to the future were extracted from news articles and related to them events were predicted in 1-2 years later.

Key words: 自然言語処理, 意味役割, 将来言及文, 未来動向予測支援, 形態パターン, 機械学習, 分類モデル

1 はじめに

近年, 株価変動や社会情勢動向などを対象とした精度の高い未来予測の実現要求が高まっている. 現在, Web を媒介として, 大量のテキストドキュメント (新聞, ブログ, ツイッター等) が容易に入手可能であることから, これらを対象に自然言語処理技術を利用したデータ解析手法や未来予測手法が注目されている. 未来予測手法においては, 類似尺度から得た因果的推論を使用した将来イベント予測システムの研究 (Kanazawa et al.[1]), 多数のオントロジーと因果関係を適用して, 文書間の類似度を計算した手法 (Radinsky et al. 2012[3]), Twitter データと機械学習によりウイルス拡散予測を試みた研究 (Aramaki et al. 2011[2]) などがある. 最近の研究である (Hurriyetoglu et al. 2018 [4]) は, Twitter 上の Time-to-event (TTE) を予測するために, いくつかのヒューリスティックな時間論理規則とツイートの歴史的な文脈でサポートされた別個の単語と単純な時間表現を適用している.

これらの未来予測システムの多くはテキスト中の時間参照情報を用いた統計処理や機械学習による予測手法を採用している. しかしながら, これらの手法には, 問題固有の統計処理技法や, 大量の学習データ処理が必要であり, 予測対象分野ごとに異なる専門的観点からのヒューリスティクスを要するため, 汎用性に欠けてしまう. そこで, 専門知識を含む将来に言及している文を利用し, 未来動向予測支援システムを提案する.

政府や企業において未来戦略を立てる際には, 各分野に精通している専門家の意見を参考にすることも少なく

ない. 専門家が先を見通す際には, その分野の基礎的な知識やこれまでの歴史や事実, また, 社会的背景や情勢との関連などを包括的に考える. 専門家の先見する文が専門分野に関する包括的な要素が含まれていると仮定すれば, 先見する文は未来動向を見通す支援になり得ると考える.

そこで, テキストデータから表層的また暗示的な将来に言及している文を自動的に抽出する手法およびそれらの文を用いた未来動向予測手法を提案し未来動向予測支援システムを目指す.

本研究の特徴は, 文を構成する形態パターンを考案し将来の事象について言及文を活用する新たな未来動向予測手法を実現することであり, 対象分野の専門知識を用いることなく, 少量の学習データで暗示的な将来言及文を抽出・分類可能であり, 汎用的な未来動向予測支援システムを開発することである.

本論文は, 2章では将来言及形態パターンの考案について, 3章では将来言及文分類モデルの構築, 4章では予測エンジンについて, 5章では未来動向予測システム構築と未来動向予測実験および評価について, 6章では研究成果について, 最終章では考察と展望について報告する.

2 将来言及形態パターンの考案

著者らは, 新聞記事から収集した将来言及文 270 文の表現を調査を行ない, 時間に関する表現 (日時, 明日, 来週など) が 70 表現, 将来の事象に使われる特徴的な語彙 (計画, 予定など) が 141 語であり, 限定した単語を用いた手法では将来言及文は難しいことを確認した (nakajima

¹釧路工業高等専門学校 創造工学科

²北見工業大学 情報システム工学科

et al. [8, 7, 6]). そこで、表層的または暗示的に将来の事象について言及する文（以下、将来言及文）を抽出するために、構文理論のひとつである意味役割と形態素の品詞情報、また、新たに定義する未来語を組み合わせた将来言及文の形態パターンを考案する。将来言及文は、あるイベントに対してその時点以降の未来に着点の動向を言及している文と定義する。以下に例を示す。

1. 三菱化学では2025年前後にLiイオン2次電池を搭載したHEVの本格普及が始まるとみている。
2. 各国から派遣された総勢約500人は23日に担当地域に移動して、対立政党間で脅迫などがいないか監視する。
3. 年末の税制改正大綱の決定に向けて、今後、調整が本格化することになります。
4. 長期的には半導体事業のうちパソコン用などの汎用DRAM事業の割合を減らしていく方針。
5. 国内では12工場を7工場に集約し、800人規模の希望退職を募るなど大規模なリストラに踏み切る。
6. アメリカの部品やソフトウェアの輸出、技術の移転を制限すると発表しました。

2.1 単語の意味役割と品詞情報

意味役割情報は、述語構造解析：文章の述語と項の関係を表した情報であり、文中の各単語の役割を解析した情報である。例を以下に示す。

(入力文) 昨日太郎は花子に本を贈った。

(出力) [昨日] 時間 (点)

[太郎が] 動作主

[花子に] 着点 (人)

[本を] 対象 (生成物)

[贈った] 状態変化あり-位置変化-位置変化 (物理) (人間) -他者への所有物の移動-提供

文の述語 (動詞) に対して、概念を付与し、述語にかかると係り元 (項) に意味的な関係を付与する。例えば「太郎」は「贈った」に対して「動作主」という意味役割であることを示している。意味役割はASA (Argument Structure Analyzer)¹で扱う72種類で意味役割の一覧は岡山大学竹内研究室が作成した述語項構造シソーラス²を用いる。

本研究で対象とする文の解析と意味役割情報の付与にはASAを用いて意味役割情報を得る [9, 10]。ASAは意味役割の他に、かかり先 (Link) や時制 (Tense)、形態素 (Morphemes) などの情報を提供する (Figure 1)。

ASAのシソーラスに存在しない単語や固有名詞などの複合語に対応するために単語の品詞情報を付与する。品詞情報は、オープンソース形態素解析エンジンMeCab³を使用する。

例えば、「第一にSociety 5.0の実装です。」で使われている、「Society 5.0」は一つの単語として扱いたいが、ASA

| | | | |
|-----------|--|-----------|--|
| ID | 0 | ID | 3 |
| Surface | 母が | Surface | 作る。 |
| Link | 3 | Link | -1 |
| CType | elem | CType | verb |
| Main | 母 | Main | 作る |
| Part | が | Similar | 25 |
| Category | ["人"] | Semantic | 状態変化あり-生成・消滅-生成 (物理) -生成- |
| Semrole | 動作主 | Voice | ACTIVE |
| Similar | 7 | Tense | PRESENT |
| Tense | PRESENT | Sentelem | PREDICATE |
| Arg | ["Arg0"] | Polarity | AFFIRMATIVE |
| Frames | ["3-verb"] | Mood | INDICATIVE |
| Morphemes | 0 母 ハハ 母 名詞, 一般 O 1 が が 助詞, 格助詞, 一般 O | Frames | ["0-動作主-Arg0", "1-elem", "2-対象 (生成物) -Arg1"] |
| | | Morphemes | 0 作る ツクル 作る 動詞, 自立基本形 O 1。。。記号, 句点 O |

Figure 1: 「太郎が」と「贈った」に対してASAを実行して得られる情報例

の結果はSociety (名詞), 5.0 (名詞) と別れてしまう。このような場合一つの単語として“Society 5.0”に名詞を付与する複合語処理を行う。

2.2 形態パターン生成

文章を意味役割情報、形態素情報、未来語を組み合わせた形態パターンで表現する。それぞれの形態情報について説明する。意味役割情報と形態素情報については、2.1で既に述べたので、ここでは、未来語について説明する。

将来言及文の中には、明らかに未来を指し示す「見通し」「予定」などの語を含むパターンが存在する。それらの文を確実に取得するために、未来語を定義する。未来語は将来言及文とその他の文の頻出度と比較し、将来言及文に特に頻出する語とする。未来語は予測対象分野別に、科学技術分野は37個、経済分野は37個、国際分野は31個、さらに、分野共通は15個とする (nakajima et al. [11, 12])。

形態パターンは、以下に示す優先順位に従い、形態パターン情報を付与する。1の未来語は、未来語群に該当する単語があれば“未来語”として付与する。2,3,4については、Figure 1で示した複数の情報からSemrole, Semantic, Categoryの3つの情報を以下の優先順位に従い各情報を付与する。

1. 未来語
2. 意味役割情報 (Semrole)
3. 述語意味論情報 (Semantic)
4. カテゴリー情報 (Category)
5. 複合語処理後の品詞

形態パターンの例を以下に示す。

例1: 原文 2023年3月末までの製品化を目標としている。

形態パターン 名詞 時間 対象 未来語 状態変化あり

例2: 原文 2020年度末からの導入を目指し、最終的に100機を配備する。

形態パターン 名詞 モノ 対象 未来語 名詞 助詞 対象 状態変化あり

¹ <http://www.cl.cs.okayama-u.ac.jp/study/project/asa/>

² <http://pth.cl.cs.okayama-u.ac.jp/>

³ <https://taku910.github.io/mecab/>

3 将来言及文取得のための分類モデル

将来言及文取得はテキストデータを分類モデルで分類することで取得する。将来言及文分類モデルについて説明する。

3.1 基本の将来言及文分類モデル

将来言及文分類モデルは、前章2の手法で生成した形態パターンを学習させることで将来言及文分類モデルを生成する。学習には、SPEC(Sentence Pattern Extraction and analysis arChi-ecture) [13]を用いる。SPECは、文の要素の順序を考慮した1~6要素の組み合わせを生成しデータを訓練できる。生成するパターンの要素を n とすると、 $1 \leq k \leq n$ を満たす整数 k 個の組み合わせグループが存在することになる。要素が6以上の場合は、'*'(アスタリスク記号)で置き換えられる。また、学習時には行われる重み計算方法が14種類用意されており、交差検定の結果で最も精度のよいモデルを採用できる。入力には正解データと不正解データが必要である。正解データに将来言及文とし、不正解データにその他の文を入力とする。各データは1文ごとに形態パターンに変換されたのち、SPECで学習を行い将来言及文分類モデルを生成する。

それぞれ同量の将来言及文とその他の文を形態パターンに変換し訓練データとしてSPECに入力し将来言及分類モデルを生成する。分類モデルの評価はニュース記事から将来言及文500文とその他の文500文合計1000文を用いて分類モデルを生成し、260文のテストデータで評価したところF値(適合率と再現率の調和平均)が0.76であることを確認し、実験に必要な基準を満たすことを示している(nakajima et al.[16])。

3.2 将来言及文分類モデルの拡張

未来動向予測に汎用性と分類モデルの精度を向上するために、ニュース記事の分野分類を参考に、科学技術、国際、経済、政策の分野別にテキストデータを学習し科学技術、国際、経済、政策および共通の分類モデルを生成する。予測したい動向に適切な分類モデルを選択する。

4 予測エンジンの構築

予測エンジンは動向予測したいイベントを問題と選択肢形式で入力し、将来言及文から得られる予測支援文を用いて結果を出力する。概要と手法について説明する。

4.1 概要

未来動向予測は、予測したいイベントは質問文、予測結果は解答を選択肢としてあたえ、そのイベントの予測支援文を用いる。予測結果はそれぞれの文に対し、単語極性情報を用い、解答を選択する形式で行う。予測支援文とは、予測したいイベントに関連するニュース記事などから将来言及文として分類した文集合である。

イベント予測問題形式の例を以下に示す。問題：2010年に行われる参議院選挙の結果を予測せよ。選択肢：1. 民主党が単独過半数をとる。2. 民主党だけでは単独過半数を取れないが、民主党を含む与党で過半数をとる。3. 民主党を含む与党で過半数を確保できない。

予測支援文は、問題文と選択肢文から抽出するキーワードでニュースコーパスから検索し、将来言及文分類モデルで将来言及文と判断された文とする。選択肢を選ぶために、極性値、分類モデル出力で得られる将来言及の強さを示す値と予測支援文に対する検索ワードの出現頻度の3つの値を用いてスコアを計算する。

4.2 手法

文章における極性とは、ある文章が良い印象をもつ(ポジティブ)か悪い印象をもつ(ネガティブ)かを表す情報である。本研究では極性を示す数値を極性値と定義する。文の極性の判定には日本語評価極性辞書[14]を用いた。日本語評価極性辞書は、用言を中心に収集した評価表現約5千件のリストに人手で評価極性情報を付与したデータである。しかしポジティブの極性の単語に「賛成」は登録されているが、「賛成」の対義語である「反対」がネガティブの極性の単語には登録されていない。本実験においては登録されている単語のみでは不十分であり、不足な単語は手動でポジティブ極性に11個、ネガティブ極性に17個追加した。

将来言及の強さを表す値は、将来言及文分類モデルで計算される将来言及文らしさの値である。

選択肢の極性値を P_c 、予測支援文の極性値を P_s 、将来言及の強さを示す値を C 、予測支援文と選択肢に共通に含まれるキーワードの個数を n とする。予測支援文1文のスコア S_i は1式で表される。 i は予測支援文の番号であり、予測支援文の個数を I とすると $[1 \leq i \leq I]$ である。また C は文章が将来について言及しているかの指標であり、将来言及分類モデルで分類される際に、将来言及文に使われている形態パターンの使用頻度やその文に使われる形態パターンの構成要素等により算出される(nakajima et al.[6])。キーワードの個数 n は予測したい内容との関連度を表す。

$$S_i = P_c \times P_s \times C \times n \quad (1)$$

予測支援文の数を I とすると選択肢のスコア S_x は2式で表される。 x は解答選択肢番号であり選択肢の個数を X とすると $[1 \leq x \leq X]$ である。

$$S_x = \sum_{i=1}^I S_i \quad (2)$$

予測支援文の1文のスコア S_i はStep1-Step4で計算される。

Step1: 選択肢の極性 P_c を日本語評価極性辞書を参照し、ポジティブの場合は1、ネガティブの場合は-1とする。また、構文が主節と従属節で構成されている場合、接続詞、接続助詞("また", "従って"など)がある場合、それらを境界とし文節に分割して適用する。分割数を A とすると分割されたそれぞれの文のスコアを $S_{x,1}, S_{x,2}, \dots, S_{x,A}$ で表す。Figure:2の選択肢2のように接続助詞がある場合は、「民主党だけでは単独過半数をとれないが」と「民主党を含む与党で過半数をとる」の2文節に分けてそれぞれの文でスコアの計算を行う。接続詞や接続助詞("しかし", "一方"など)の極性を反転させる反転子については、文節に分けることで考慮可能になる。

Step2: 次に予測支援文1文のスコア S_i を1式により算出する。図2は選択肢1を例にスコア S_1 を計算した例であ

る。説明のため予測支援文の将来言及の強さ C は将来言及分類モデルで 2.0 が算出されたと仮定し例を示す。予測支援文 1 文のスコアは S_1 は、 $[S_1 = P_c \times P_s \times C \times n = 1 \times (-1) \times 2 \times 2 = -4]$ となる。

予測支援文の例

民主党が単独過半数を占めるのは難しくそうだ

極性 $P_s = -1$ 将来言及の強さ $C = 2.0$

選択肢 1 のスコア計算

1. 民主党が単独過半数をとる。

極性 $P_c = 1$ 将来言及文との検索ワード共通数 $n = 2$

Figure 2: 予測支援文を用いたスコア S_1 の計算方法

Step3: 各選択肢のスコア S_x は、全支援文 $[1 \leq i \leq I]$ の S_i を求め、総和を求める。ただし、分割して文節ごとに計算した選択肢のスコアは、各文節を統合し、スコア $S_{x,1}, S_{x,2}, \dots, S_{x,A}$ の和を S_x とする。2 式より、選択肢 2 のスコア S_2 は $[S_2 = S_{2,1} + S_{2,2} = -4 + 10 = 6]$ が算出される。

Step4: 最もスコアが高い選択肢を予測結果とする。以上の手順に従い全ての予測支援文と解答選択肢に適用し未来動向予測問題の予測結果を得る。

5 未来動向予測システムの構築

未来動向予測システムの構成と実験および評価について説明する。

5.1 構成

未来動向予測システムは将来言及文分類モデルと予測エンジンから構成される (Figure 3)。入力は予測したいイベントの問題文と解答選択肢とし、出力は選択された解答選択肢である。予測支援文はニュースコーパス、Web 上に公開されているニュース記事や政府が公開している白書などから取得する。

5.2 実験および評価

12 年後の動向予測を第 4 回先見力検定⁴ から選んだ 7 問、著者らが作成した科学技術に関する問題 8 問を用い実験を行いシステムの評価を行う。予測するイベントは既に結果が分かっている過去イベントとして、その実際の結果と比較する。第 4 回先見力検定の予測実験には 2009–2010 年の佐賀新聞⁵ と経済白書⁶、2009 年の毎日新聞コーパス⁷ から獲得した予測支援文用い、2011–2012 年の動向予測を行う。また、著者らが作成した科学技術に関する未来予測問題の予測実験には、2012 年、2013 年の毎日新聞コーパスから獲得した予測支援文を用い、2014 年の動向予測を行う。Web に存在するイベント関連文の収集に

⁴<http://geneki.a.la9.jp/senken/index4.html>

⁵<https://www.saga-s.co.jp/>

⁶<http://www.kantei.go.jp/jp/hakusyo/>

⁷<http://www.nichigai.co.jp/sales/mainichi/mainichi-data.html>

は問題文のキーワードを google 検索機能に入力し Web 上からキーワードを含む記事を検索し用いる。

本実験の結果は第 4 回先見力検定が正答率 71.4%、科学技術に関する問題が正答率 75.0% の精度で予測できた (nakajima et al. [12])。

2009 年に実施された第 4 回先見力検定を受験した受験者の平均正答率は約 30%⁸ である。この結果は、1 年間どんな方法でもデータでも調べて良い条件で受験した結果である。また、Kurokawa et al. [15] が数回行なった第 4 回先見力検定の結果、受験者の平均正解率は 30% 程度に止まると言及している。

6 研究成果

本研究で構築した未来動向予測支援システムについて各部分ごとに考察を述べる。

6.1 形態パターンを用いた将来言及文の抽出と将来言及文の有用性

新聞記事と web ニュース記事から将来言及文とその他の文を各 500 文取得し、形態素情報と文脈解析に用いられる意味役割を併用した形態パターン (Morphosemantic Patterns: MoPs) を生成し、教師あり機械学習を用い将来言及文を分類するための分類モデルを生成した。さらに、精度の向上を目指し、将来言及文の特徴となる語を調査し、それらを未来語として定義し、前述の MoPs に未来語を素性として加え、分類モデルを生成し比較を行なった。また、ニュース記事は国際、科学技術、政治などドメインに分類されていることを利用し、科学技術の記事に注目し実験を行なった。以上の手法により抽出した将来言及文を用いた未来イベント予測を行う際に、将来言及文が有効であるか否かを検証するためのプロトタイプ予測エンジンを設計し実験を行なった。予測エンジンは、将来言及文を入力とし、機械学習により正解データを分類する正解分類モデルを生成し、この分類モデルに問題文を入力とし、2 個以上の選択肢の中から正解を選び解答する。将来イベントを予測する設問 7 問を人間と本手法により実験した結果、人間による予測実験の正解率は 30% 程度に対し、プロトタイプ実験では、50% 以上の結果が得られ、将来イベントを予測するための情報として、将来言及文の有用性を確認できた (Nakajima et al. [16, 11, 18, 17])。

6.2 将来言及文分類モデルの構築

未来動向予測支援システムの中核部である将来言及文獲得のための分類モデルの精度向上と汎用性の強化を実現し最適な分類モデルの構築、および、汎用性を備えた予測エンジンの開発を行った。将来言及文獲得を行う分類モデルの精度向上には、将来言及文に使われる特徴的な語を未来語と定義し、形態パターンの要素 (意味役割: 動作主, 動作など) に「未来語」を新たに追加し分類モデルを生成した。未来語は分野ごとに将来言及文 1000 文を用いて決定した。未来語を含まない潜在的な未来言及文も存在するため、要素に未来語を考慮する学習と未来語を意味役割に置き換え学習する二段階学習構造による分類モデルの生成を試みた。各分類モデルの評価には、新

⁸<http://geneki.a.la9.jp/kentei.html>

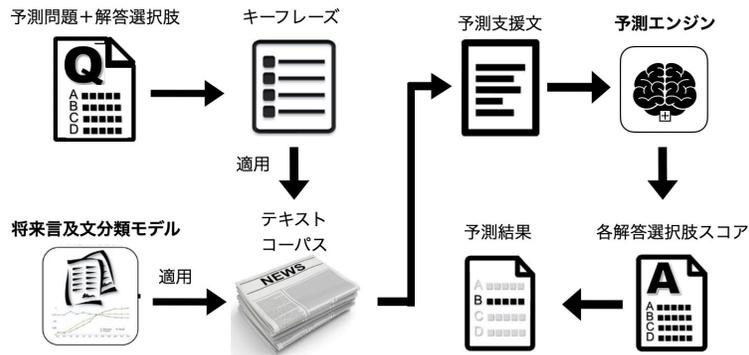


Figure 3: 未来動向予測システム概要図

たに収集したデータ 1200 文を用いて行い、H29 年度の実験結果よりも約 11 ポイントの向上を確認した。分類モデルの汎用性強化と最適化実装は、ニュース記事の分野分類を参考に、科学技術分野の分類モデルに加え、国際、経済、政策の分類モデルを生成し、ニュースドメイン毎に適切な分類モデルを切り替え可能とする機能を追加した。さらに各分野と全分野共通の未来語を定義し、汎用性を考慮した (nakajima et al. [16, 19, 20])。

6.3 予測エンジンの構築

予測エンジン部は、実世界の問題に対応するために予測したい未来動向が [起きるか/起きないか/どちらともいえない] の回答を得る手法として、単語の極性が [ポジティブ/ネガティブ] であるかを表す日本語評価極性を用いた手法を考案し実験を行った。予測エンジンの精度は、プロトタイプ予測エンジンよりも分野混合動向予測問題の予測結果は 23.9 ポイント、分野別の予測結果は 12.5 ポイント高い結果が得られた。しかし、単語の極性が [ポジティブ/ネガティブ] であるかを表す日本語評価極性を用いた手法では辞書依存が原因で日本語評価極性辞書にない単語が頻出すると判定できないという問題を解決するために、分野ごとに頻出単語を辞書補完し、極性判定を定式化し極性分類精度の改善を行った (nakajima et al. [16, 12])。

6.4 未来動向予測支援システムの構築と評価

将来言及文抽出精度、予測エンジンの予測精度を向上させるため、各過程において検討・精査、およびアルゴリズムの改善を行い、未来動向予測システムの構築および、実世界における予測問題とデータを用いシステムの有用性の検証を行った。意味役割付与機構、形態パターン導出機構、将来言及形態パターン導出機構、将来言及文分類モデル、予測エンジンを組合せ、将来動向予測システムを構築した。本システムの評価は、既存の「先見力検定」で出題された問題と既に結果が判明している実世界のイベントから我々が作成した未来予測問題を用い Web ニュース記事を入力として実験を行った。全 15 問の平均正解率は約 73% であり人間の平均予測正解率 30% を大きく上回る結果を得た (nakajima et al. [12, 19, 20])。

7 まとめと展望

本研究が目的とする、対象分野の専門知識を用いることなく、文を構成する形態パターンを用い少量の学習データで潜在的な将来言及文を抽出・分類できる、汎用性を備えた未来動向予測支援システムを構築した。従来手法とは違う手法で未来予測の可能性を示すことができ、他分野への発展が期待できる。また、対象分野の固有表現を持つ文の分類が容易になることで、未来予測以外の他分野への発展が期待できる。本システムは、専門的知識を持たずとも近年 1, 2 年の将来の事象について言及するデータをもとに企業・政策活動などにおいて、中・長期戦略など先を見通すための支援に活用が可能であり、さらには社会技術イノベーション創出の分野に貢献できると考える。

謝辞

本研究は JSPS 科研費 17K00324(2017–2019) の助成を受けたものである⁹。

References

- [1] Kensuke Kanazawa, Adam Jatowt, Katsumi Tanaka. 2011. Improving Retrieval of Future-Related Information in Text Collections. 2011 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), pp. 278–283.
- [2] Eiji Aramaki, Sachiko Maskawa, Mizuki Morita. 2011. Twitter catches the flu: Detecting influenza epidemics using twitter. Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 1568–1576.
- [3] Kira Radinsky, Sagie Davidovich and Shaul Markovitch. 2012. Learning causality for news events prediction. The 21st International Conference on World Wide Web.
- [4] Hurriyetoglu, A., Oostdijk, N., & van den Bosch, A. (2018). Estimating time to event of future events based on linguistic cues on twitter. In Intelligent natural language processing: Trends and applications (pp. 67–97). Cham: Springer.

⁹<https://kaken.nii.ac.jp/ja/grant/KAKENHI-PROJECT-17K00324/>

- [5] M.Higashiyama, K.Inui, Y.Matsumoto. (2008). Learning Sentiment of Nouns from Selectional Preferences of Verbs and Adjectives. Proceedings of the 14th Annual Meeting of the Association for Natural Language Processing, pp.584–587.
- [6] Y.Nakajima, M. Ptaszynski, F. Masui, H. Hirotooshi. (2016). A method for extraction of future reference sentences based on semantic role labeling. IEICE Trans. Information and Systems, Vol.E99D, No.2, pp.514–524.
- [7] Yoko Nakajima, Michal Ptaszynski, Hirotooshi Honma, Fumito Masui. (2016). Automatic Extraction of Future References from News Using Morphosemantic Patterns with Application to Future Trend Prediction. ACM Summer Newsletter AI Matters, Vol.2, No.4, pp.13–15.
- [8] Yoko Nakajima, Michal Ptaszynski, Hirotooshi Honma, Fumito Masui. (2014). Investigation of Future Reference Expressions in Trend Information. Association for the Advancement of Artificial Intelligence (AAAI) Spring Symposium, SS-14-01, pp.31–38.
- [9] 竹内孔一, 森本真依子. (2009). 動詞項構造ソーラスに基づく動詞語義ならびに意味役割付与データの構築. 言語理解とコミュニケーション研究会, NLC2009-9 pp.13–18.
- [10] 原 靖弘, 竹内 孔一. (2016). 係り元の末尾表現に着目した Hierarchical Tag Context Tree を利用した日本語意味役割付与システムの構築, 情報処理学会論文誌 Vol.57 No.7 1611–1626.
- [11] 中島陽子. (2017). 未来イベント予測のための将来言及文における特徴語の調査. 釧路工業高等専門学校紀要, vol.51, pp64–68.
- [12] Yoko Nakajima, Michal Ptaszynski, Hirotooshi Honma, Fumito Masui. (2019). A Proposal of Prediction Method Using Word Polarity Information for Future Event Prediction Support System. Advanced Informatics, Concept, Theory, and Applications (ICAICTA), No.95.
- [13] M. Ptaszynski, R. Rzepka R, K. Araki, Y.Momouchi. (2011). SPEC-Sentence Pattern Extraction and Analysis Architecture. In Proceedings of the Seventeenth Annual Meeting of the Association for Natural Language Processing.
- [14] M.Higashiyama, K.Inui, Y.Matsumoto. (2008). Learning Sentiment of Nouns from Selectional Preferences of Verbs and Adjectives. Proceedings of the 14th Annual Meeting of the Association for Natural Language Processing, pp.584–587.
- [15] T.Kurokawa, H.Kekeya. (2009). Analysis of a tendency to answer in foresight official approval (in Japanese). In:The Fifth media information inspection arts and sciences meeting, pp.50–55.
- [16] Y.Nakajima, M. Ptaszynski, F. Masui, H. Hirotooshi. (2017). A Prototype Method for Future Event Prediction Based on Future Reference Sentence Extraction. In: Proceedings of Workshop on Linguistic and Cognitive Approaches To Dialogue Agents (IJCAI 2017), pp.42–49.
- [17] Yoko Nakajima, Michal Ptaszynski, Hirotooshi Honma, Fumito Masui. (2018). Future Reference Sentence Extraction in Support of Future Event Prediction, International Journal of Computational Linguistics Research, Vol.9, No.1, pp.27–41.
- [18] 中島陽子. (2017). 未来イベント予測のための将来言及文における特徴語の調査. 釧路高専紀要, Vol.51, pp.64–68.
- [19] 中島陽子, 本間宏利, Akmal Hakim, Michal Ptaszynski, 梶井文人. (2020), 将来言及文の分類精度向上を目的とした汎用型分類モデルの構築, 釧路高専紀要, Vol.53, pp.37–42.
- [20] Yoko Nakajima, Michal Ptaszynski, Fumito Masui, Hirotooshi Honma, (2020), Future Prediction with Automatically Extracted Morphosemantic Patterns, Cognitive Systems Research, Vol.59, pp.37–62.