

# つぶやきから To-Do を発掘する： 入力負荷を軽減した To-Do 管理システムの提案

谷岡 遼太<sup>1,a)</sup> 宮部 真衣<sup>2,b)</sup> 吉野 孝<sup>3,c)</sup>

概要：近年の携帯情報端末の発達に伴うデジタルツールの普及により，一般向けの To-Do 管理ツールはより日常的に感じられるものとなった．しかし，To-Do 管理の堅苦しい印象は，学生が積極的な管理習慣を身に付け始めるための障壁になっていると考えられる．To-Do 管理の普及には，個人のもつ To-Do 管理への抵抗感の減少や，動機づけが必要である．また，To-Do 管理を促進するためには，To-Do 管理ツール独自のインタフェース上の操作や，文字情報の入力に関する負荷の軽減が不可欠である．そこで本稿では，To-Do 管理における入力負荷軽減のため，マイクロブログ上の To-Do 自動抽出手法を用いた To-Do 管理システムを提案する．To-Do 管理のためのカテゴリを定義し，Twitter のツイートデータの分類によりコーパスの構築を行った．また，精度検証実験を行い，「希望」カテゴリについては，比較的高精度に判定可能であることを示した．

## 1. はじめに

ビジネス社会では個人の自己管理能力や生産向上性が求められる．日立ソリューションズの調査によると，20代のビジネスパーソンのうち，92%は時間・タスク管理術を必要と感じており，社会人にとって，To-Do 管理の重要性は高い [1]．仕事は，実行時間が決められているアポイントメントと，決められていないタスクに分かれる．水口は，タスク（仕事）比の大きい社会人に備え，アポイントメント（授業）の多い学生のころから，効率的な時間管理術を養うべきだと指摘しており [2]，アポイントメントを予定表に書き込み，タスクを To-Do としてリスト化する方式を推奨している [3]．本稿では，タスクのうち，現在実行している作業や近い未来で実行しうる作業を「To-Do」と呼び，To-Do 管理の支援方法を検討する．

To-Do 管理の重要性は高いものの，その実践率は高いとはいえない．To-Do 管理を実践している 20 代のビジネスパーソンは，42%に留まっている [1]．また，工学部の大学生を対象に，To-Do 管理と SNS に関する意識調

査を実施した結果 [4]，To-Do 管理をしていない学生は約 19%（63 名中 12 名）であった．To-Do 管理をしない理由としては，「面倒」「書くタイミングがない」「習慣がない」といった点が挙げられた．本研究では，大学生を対象とした To-Do 管理支援を実現することにより，卒業後ビジネス社会で働く際に日常的に To-Do 管理を実践可能にすることを目指す．

To-Do 管理の普及には，To-Do 管理への抵抗感の軽減や，動機づけが必要であると考えられる．佐々木らは，スマートフォンと，文字情報の少ない To-Do 管理との好相性について指摘している [5]．一方で，日本人の To-Do 管理ツールの利用率は他国に比べて低い [6]．2013 年の調査によれば，デジタルツールでのスケジュール管理をやめた理由として，「操作が面倒（41.1%）」が最も多く挙げられている [7]．このことは，デジタルツール独自のインタフェースやスマートフォンによる文字入力の操作が，手書きの場合と比べ，利用者に負担を与えていると考えられる．つまり，To-Do 管理ユーザにとっての操作や入力に関する負担の大きさは，ツールの利用を継続させる際の重要な問題であり，To-Do 管理の利用を促進するためには，これらの負荷の軽減が不可欠であると考えられる．

そこで我々は，大学生の利用率の高い，マイクロブログに着目する<sup>\*1</sup>．これまでに我々は，日常の行動を対

<sup>1</sup> 和歌山大学大学院システム工学研究科  
Graduate School of Systems Engineering, Wakayama  
University, Wakayama 640-8510, Japan

<sup>2</sup> 京都大学学際融合教育研究推進センター  
Center for the Promotion of Interdisciplinary Education  
and Research, Kyoto University, Kyoto 606-8507, Japan

<sup>3</sup> 和歌山大学システム工学部  
Faculty of Systems Engineering, Wakayama University,  
Wakayama 640-8510, Japan

a) enlosph@gmail.com

b) mai.miyabe@gmail.com

c) yoshino@sys.wakayama-u.ac.jp

<sup>\*1</sup> マイナビの調査によると，2015 年卒大学生のうち，マイクロブログである Twitter の利用者は約 6 割である [8]．我々が実施した調査 [4] においても，Twitter を利用している学生は約 78%（63 名中 49 名）であり，大学生のマイクロブログ利用率は高い．

象とした個人向けの To-Do 管理システムの開発を行ってきた [9]. 大学生が日常的に利用しているマイクロブログと, To-Do 管理システムを連携させることにより, To-Do 管理における入力負荷の軽減を目指す. 具体的には, マイクロブログ上に発信されたテキストから To-Do を自動抽出し, To-Do 管理システムと連携する. 普段発信しているマイクロブログ上のテキストから To-Do を抽出できれば, 利用者に To-Do の入力を意識させることなく To-Do が登録されることになる. 本研究では, まず, To-Do のカテゴリを定義し, Twitter<sup>\*2</sup>上のツイートを用いて, 分類用のコーパスを構築した. また, 構築したコーパスをもとに, 精度検証実験を実施した.

本稿では, 2 章において関連研究について述べる. 3 章において, 本システムについて説明する. 4 章において, コーパスの構築について述べる. 5 章で精度検証実験について述べる. 最後に 6 章で本稿の結論についてまとまる.

## 2. 関連研究

### 2.1 個人のタスク管理に関する研究

情報共有のためのグループウェアシステムは従来から研究が進められており, その支援内容は多岐にわたる. その多くは組織内のタスク処理を目的としており, スケジュール管理・プロジェクト管理・ワークフロー管理などが挙げられる.

個人の予定の共有を目的とした研究では, 濱崎らが学会会議における共有型スケジュール支援システムの開発と運営を実現した [10]. この研究では, 開発システムが会議開催前の情報収集支援について有用性があることを確認した. また, 大向らは人間関係ネットワークに基づく情報フィルタリングを用いたタスクスケジューラの提案を行っている [11]. ここで述べる人間関係ネットワークは, 複数人の間で公開された個人のスケジュール情報とタスクの依頼関係から生成される. 本研究では, 個人の To-Do 管理の利用意識の強化のため, 共有内容を限定していないマイクロブログ上のテキストを利用する.

他者との共有や一般公開が可能な To-Do 管理 Web サービスに, checkpad<sup>\*3</sup>が挙げられる. また, RTM<sup>\*4</sup>などの To-Do 管理 Web サービスの一部には Twitter との連携が可能であり, ハッシュタグやダイレクトメッセージを利用した To-Do の登録や通知の機能が実装されている. 例えば, TwitDo<sup>\*5</sup>と一度連携した Twitter ユーザは, 「#todo タスク」と呟くことで To-Do を登録し, 「#done タスク」と呟くことでその To-Do を完了する.

また, Facebook<sup>\*6</sup>などの一部のマイクロブログでは, ユーザ同士で個人の予定を公開し共有する機能が実装されている. 以上のサービスは, ユーザの意識的な To-Do 管理のために利用される. 本研究では, To-Do 管理を意図せずとも, 普段どおりに従来のツールを使うだけで, To-Do を登録可能な環境を構築し, To-Do の入力負荷の軽減を目指す.

### 2.2 マイクロブログ上の行動に関する研究

村益らは, マイクロブログを利用する大学生の投稿内容について調査した [12]. その結果によれば, 「自分について」を示す Twitter 上の投稿は全体の 55.3%であり, 同氏の決める分類項目の中では最多となった. また, 「趣味」「困難な経験」「自慢」を示す内容は, それぞれ全体の 47.7%, 11.2%, 1.2% を占めることがわかった. これらには, To-Do として取り扱うことが可能な内容が含まれると考えられる. そこで本研究では, Twitter 上のツイートを To-Do の収集源として取り扱う.

Naaman らは, Twitter ユーザの発言内容について, 自分の現状や意見などの 9 つのカテゴリに基づき分類を行った [13]. 調査結果によると, 自分の現状についての発言は全体の約 41% を占める. 本研究では, この自分の現状に関するツイートの内容について, さらに 7 つのカテゴリを定義し, 自動分類を試みる.

ミンテイらは, Twitter からの人間行動属性の自動抽出を提案した [14]. この研究では, 条件付き確率場と自己教師あり学習を用いて, 人手によるラベル編集を必要としない行動の基本属性 (行動主, 動作, 対象, 時間, 場所) の自動抽出を実現した. 人間行動属性は, 本研究における To-Do のような個人の日常的行動を包括する. ただし, この研究では行動モデルの構築や行動推測を目的としている. 本研究では, 以上の行動を管理可能な To-Do としてマイクロブログユーザに提示し, To-Do 管理の促進を試みる.

## 3. To-Do 自動抽出手法を用いた To-Do 管理システム

### 3.1 システムの目的

To-Do 管理の普及においては, 解決すべきさまざまな課題が存在する. なかでも, 「未経験者への動機づけ」や「モチベーションの維持」は, とくに重要な課題であると考えられる. これらの課題の解決においては, ユーザの作業負担の軽減が, 重要な役割と果たすと考えられる. そこで本研究では, ユーザがマイクロブログ上に発信した情報から, 管理対象となる To-Do を自動抽出し, 入力負荷を軽減した To-Do 管理システムを開発する. システムは, 自動抽出した To-Do 候補をユーザに提示し, ユーザの To-Do 登録負荷を軽減する. 前述した 2 つの

<sup>\*2</sup> <https://www.twitter.com/>

<sup>\*3</sup> <http://www.checkpad.jp/>

<sup>\*4</sup> <https://www.rememberthemilk.com/>

<sup>\*5</sup> <http://twitdo.com/>

<sup>\*6</sup> <https://www.facebook.com/>

## (1) Twitter



## (2) To-Do 管理システム



図 1 Twitter と To-Do 管理システムとの連携

課題に対する、本提案手法のアプローチを以下に示す。  
未経験者への動機づけ

学生は、To-Do 管理経験の有無にかかわらず、マイクロブログを利用している場合が多い。本手法により、マイクロブログユーザは、普段のマイクロブログの利用から半無意識的に To-Do 管理を実践できる。また、マイクロブログから抽出される To-Do 候補は、ユーザのどのような行動が To-Do となるのかを教示する。これらのことは、ユーザが抱く To-Do 管理への抵抗感の軽減や動機づけをもたらすと考えられる。

### モチベーションの維持

To-Do 候補の生成において必要な文字情報は、マイクロブログ上に存在するため、To-Do 登録や管理における操作負担はない。このことは、To-Do 管理を主目的としないマイクロブログユーザの利用促進や、To-Do 管理への習熟度が高いマイクロブログユーザの利用継続をもたらすと考えられる。

### 3.2 システムの構成

本研究では、To-Do 管理システムを、以下の機能をもつ Web システムと定義する。

- (i) ユーザ登録、ログイン認証
- (ii) 手動による To-Do の登録
- (iii) 登録された To-Do の閲覧、管理

本研究では、機能 (ii) における入力負荷を軽減するために、Twitter から To-Do を自動抽出する To-Do 管理システムを提案する。そこで本システムでは、上記の機能 (i) について、Twitter のアカウントによるユーザ・ログイン認証を利用する。また、機能 (ii) について、Twitter 上のテキストを用いた To-Do の自動登録の機能を追加する。図 1 に Twitter と To-Do 管理システムとの連携のイメージを示す。ユーザは、まず、Twitter のアカウント (図 1(1)) を用いて、To-Do 管理システム (図 1(2))

にログインする。ユーザは、(A) ユーザ自身が一から入力を行う従来手法と、(B) Twitter 上のテキストをもとに To-Do を自動抽出する提案手法のいずれかを用いて、To-Do 管理システムに To-Do を登録する。(B) の手法では、利用者がツイートした時点で To-Do の自動登録を行う。ただし、「( ??? )を買えた」(図 1 の To-Do 管理システムにおける (B) の (c) )のように、To-Do 登録に必要な情報が不足している場合、(B) では To-Do 候補の提示のみを行う。ユーザは本システムにログインした後、テキストの一部や抽出先のツイートを参考に、登録する To-Do の情報を補完できるようにする。

Twitter 上のテキストから To-Do を自動抽出する手順を以下に示す。

- (1) Twitter 上のテキストについて、SVM (Support Vector Machine) [15] により構築した分類器を用いて、後述する To-Do カテゴリへと分類する。
- (2) 条件付確率場 (CRF: Conditional Random Fields) [16] を用いて、To-Do となる部分を判定する。

本論文では、手順 (1) における To-Do カテゴリ分類のためのコーパスの構築と精度検証実験について、4 章と 5 章で述べる。

### 3.3 To-Do カテゴリ

To-Do は、達成状態や実行期日、依頼関係などのさまざまな情報を含むことがある。そこで、これらの情報を考慮し、マイクロブログ上の自分の現況に関するツイートについて、7つのカテゴリを定義した。定義にあたっては、一部 GTD<sup>\*7</sup>の観点を参考にした。

表 1 に To-Do カテゴリの定義と該当するツイート例を示す。以下に定義について述べる。

- ツイート内容が To-Do を含む場合 (例: 「レポートを終わらせよう」「炊飯器が欲しい」)、実行優先度

<sup>\*7</sup> Getting Things Done の略。2002 年に Allen が提唱した生産性を上げる仕事術 [17]。

表 1 To-Do カテゴリの定義とツイート例

カテゴリ名	グループ*1 と定義	ツイート例
(1) 予定	P 自分がすること・しようとしていること・ 予定・計画・To-Do.	17時までにレポートを終わらせよう。 はやく帰って旅行の準備しないと
	N 自分がしないこと・してはいけないこと・しない予定.	試験終わるまでゲームしない
(2) 希望	P 自分がしたいこと・なりたいこと・ 欲しいこと・希望.	早めに寝たい。 こういう時に炊飯器が欲しいよねー!
	N 自分がしたくないこと・なりたくないこと・ 欲しくないこと.	正解がないのに、性格は間違えたくないって思っちゃう 甘いものは要らないのに...
(3) 状態	P 自分が実行していること・進行していること・ 続けていること・継続した予定.	毎日ジョギングで歩いています ギリシャを旅行中です!テンション上がる(・_・)
	N 自分が実行していないこと・続けていないこと.	最近タバコを吸っていない
(4) 保留	P-1 自分がすべきことを、していないこと・しなかったこと・未着手.	まだ宿題やってないなーどうしよう.
	P-2 自分ができないこと・しにくいこと・できなかったこと・不可能.	報告書 5 つ...1 つしか終わらせられなかった(; ;)
	N 自分がやってはいけないことをしてしまうこと・してしまったこと.	ついカ食いをしてしまう... う〜ん
(5) 達成	P 自分が過去にしたこと・できたこと・達成・完了.	よーし仕事終わったああああああ
	N 自分がやってはいけないことをしなかったこと・しないでいられたこと.	今日は珍しく遅刻をしなかった!
(6) 依頼	P 他人にさせること・してもらふこと・他人に依頼する予定.	ガイドさんに道案内してもらおうつもりです
	N 他人にさせないこと.	勝手に荷物を置かせないようにしないと。
(7) スキル	P 自分が今、いつでもできること・可能・自慢.	実は自分もギターが弾ける...

\*1 P をポジティブグループ, N をネガティブグループとする.

の大小から, ツイートは表 1(1) 予定または (2) 希望に属する.

- ツイート内容が長期間実行される To-Do を含む場合 (例:「毎日ジョギングで歩いている」), ツイートは (3) 状態に属する.
- ツイート内容が達成状態のわかる To-Do を含む場合 (例:「宿題をやっていない」「仕事を終わらせた」), ツイートは (4) 保留または (5) 達成に属する.
- ツイート内容が依頼関係を示す To-Do を含む場合 (例:「ガイドさんに道案内をしてもらおう」), ツイートは (6) 依頼に属する.
- ツイート内容が自己のいつでも達成できる状態 (可能) を含む場合 (例:「ギターを弾ける」), ツイートは (7) スキルに属する.

上記のカテゴリに属するツイートには, 実行に対する肯定・否定の意思が含まれている. そこで, 以上のカテゴリについて, (P) ポジティブなグループと (N) ネガティブなグループとした.

また, 上記のカテゴリは, 1 つの To-Do における状態遷移を示している. 例えば, 「レポートを作成したい」(希望) というツイートをしたユーザは, そのあとに「明日までにレポートを作成する」(予定), 「レポートを作成できた」(達成) というツイートを行う可能性がある. このとき, 開発する To-Do 管理システムは, これらを同一の To-Do に関する状態遷移としてとらえ, 自動的に状態管理ができるようにすることを想定している.

## 4. Twitter 上のデータを用いたコーパスの構築

### 4.1 コーパスの構築方法

本研究では, 3 章で述べた 7 カテゴリに関して, 該当するツイートの自動抽出を試みる. そこで, Twitter のツイートをを用いて分類用コーパスを構築した.

まず, 工学部の大学生 10 名 (男性 5 名, 女性 5 名) が 2014 年 6 月 7 日までに公開した最新のツイート各 100 件\*8, 合計 1000 件のツイートを収集した. 次に, 評価者に, To-Do カテゴリへの分類作業を依頼した. 評価者は, 評価対象のツイートを発信した 10 名以外の工学部の大学生 3 名 (男性 3 名) である. 評価者 3 名は Twitter の利用経験がある. 評価者には, 提示された各ツイートに対し, 各カテゴリの 2 値分類 (属する/属さない) 作業を行ってもらった. なお, 各ツイートは複数のカテゴリに属する可能性がある.

### 4.2 各カテゴリへの分類結果

表 2 に, 3 名の評価者による分類結果および分類の一致率 (統計量) を示す. 各カテゴリへと分類した結果, 見られた特徴を以下に示す.

#### (1) 予定

「予定」カテゴリについては, 判定件数が多く, 「ご飯食べたら行く」「課題しよーっと」など, ツイートユーザの行動意思が含まれるツイートに関して一致率が高かった. 「多数決」での分類結果 104 件について, 投稿時点から実行を開始する To-Do を含むツイート (例:「仕入れ行ってこよー」「また持つ

\*8 リプライや公式リツイートを除き, 非公式ツイートを含む 100 件を評価対象とした.

ていくかな、傘」など)は, 61件(58.7%)だった。また, 「今日は」「6時に」「そろそろ」など, To-Doの実行日時あるいは時期を示す語句のあるツイートは, 36件(34.6%)だった。以上の両方の条件を満たすツイート(例:「今から研究室行きます!」など)は4件(3.8%)だった。このことから, 「予定」カテゴリのツイートには, 時間指定のないリアルタイムの To-Do が含まれやすいと考えられる。

(2) 希望

「希望」カテゴリについては, 判定件数が多かった。また, 多数決の結果である78件のうち, 「~たい」という表現が含まれるツイートは, 37件(47.4%)であり, 「希望」カテゴリに分類されたツイートには, 「たい」の文字を含みやすいと考えられる。

(3) 状態

「状態」カテゴリについては, 評価者間で判定件数の差が大きかった(最大120件, 最小18件)。1名の評価者が多く判定したのは, 「動作の継続状態」以外の広義の「状態」を含んでいたためである。広義の「状態」には, 現在の状況を示す内容(例:「暑い」など)が多く含まれていた。

(4) 保留

「保留」カテゴリについては, 評価者間のばらつきが大きかった。定義が複雑であり, 評価者が混同した可能性がある。

(5) 達成

「達成」カテゴリについては, 評価者間で判定件数の差が大きかった(最大141件, 最小21件)。ツイートには「~した」「~できた」を含む表現が多く, これらの表現が意図する達成度の差の解釈に関して, 評価者間で意見が異なると考えられる。また, ツイートしたユーザ以外が達成したと考えられる内容についても「達成」と判定される場合があった。

(6) 依頼

「依頼」カテゴリについては, 判定件数が少なかった。その理由の一つとして, マイクロブログ上のユーザが第三者に関係するツイートを公開しにくいことが考えられる。また「多数決」の5件のツイートは, 第三者を特定しない願望を含む内容(例:「誰か!俺に昼飯食わせてくれ!」など)だった。

(7) スキル

「スキル」カテゴリは, 判定件数が少なかった。明確に能力に関して記述してはいないが, スキルを暗示する内容のツイートがカテゴリに該当すると判定するかどうかで, 評価者間の意見が分かれた。例えば, 1名の評価者は, 「グループミーティングは出ます」という内容のツイートに関して, 「出ます」という表現が「出ることができる」という能力を暗示すると考え, 「スキル」に属すると判定した。

表2 To-Do カテゴリへの分類結果と一致率

カテゴリ名	判定件数*1				統計量*2
	A	B	C	多数決	
(1) 予定	143	134	81	104	0.45
(2) 希望	111	113	77	78	0.51
(3) 状態	39	120	18	16	0.11
(4) 保留	37	52	20	12	0.08
(5) 達成	141	85	21	56	0.22
(6) 依頼	14	6	18	5	0.16
(7) スキル	5	41	24	4	0.03

\*1 A, B, C は各評価者を示す。

\*2 0.4以上の統計量を太字で示す。

#### 4.3 Twitter から抽出可能なカテゴリ

表2に示した判定件数は, 各評価者(A, B, C)が1000件のツイートを分類した結果, 各カテゴリに該当すると判断したツイートの数である。また, 表2中の統計量は, 3名の評価者の各組み合わせ(A-B, B-C, C-A)における平均値である。

表2より, 「(1) 予定」と「(2) 希望」に関する統計量は比較的高い値を示している。つまり, この2つのカテゴリは, 人間による分類のぶれが比較的発生しにくいカテゴリであると考えられる。一方, それ以外のカテゴリに関する統計量は低い値を示しており, 評価者によって判断が分かれやすいカテゴリであると考えられる。また, 「(3) 状態」と「(5) 達成」は, 評価者間で判定件数の差が大きかった。「(4) 保留」, 「(6) 依頼」, 「(7) スキル」は, 判定件数が少なく, そもそもTwitter上にはほとんど出現しない内容である可能性が高い。

#### 4.4 Twitter からの To-Do の抽出

本節では, TwitterからのTo-Do抽出可能性について考察する。

表2で示したように, 3名の評価者の多数決により7つのカテゴリへの分類を行った結果, ツイート1000件中50件以上の該当ツイートがあったカテゴリは「(1) 予定」「(2) 希望」「(5) 達成」の3カテゴリのみであった。この3カテゴリについては, Twitter上のツイートにある程度出現しうるものであると考えられ, 分類精度の改善により, 自動抽出できる可能性がある。

「(3) 状態」「(4) 保留」については, 4.2節で考察したように, 評価者間で評価結果が揺れており, 十分な事例が収集できなかった可能性がある。これらのカテゴリについては, 評価者間での判定の揺れがなくなるようにカテゴリの定義を見直した上で, Twitter上に該当するツイートが投稿されうるのかどうかを検証し直す必要がある。

「(6) 依頼」「(7) スキル」に関しては, 多数決による分類の結果, 該当するツイートが5件以下であり, そもそもTwitter上に投稿されず, 抽出が困難である可能性

表 3 「希望」カテゴリにおける重みの大きい素性の一覧

	素性*1	重み
1	たい	0.373
2	ぐらい	0.106
3	CD	0.093
4	午前	0.091
5	たへん	0.091

\*1素性には、形態素の原形を用いた。

がある。文献 [12] の調査結果によると、「自慢」に関する Twitter の投稿は少なく、特に「(7) スキル」については、Twitter にはほとんど投稿されないと考えられる。ただし、「(6) 依頼」については、4.2 節で示したような、不特定の第三者に対する依頼などは存在する可能性もあるため、Twitter 上からの抽出可能性について今後再検証を行う。

また、より多人数のツイートデータの分析から、Twitter に To-Do が含まれる可能性について検証する。

## 5. 精度検証実験

4.1 節で述べたコーパスを用いて、To-Do カテゴリごとの分類精度を検証する。

今回は、表 2 における「多数決」の結果を正解データとして扱うこととする。なお、人手で分類した結果、一部のカテゴリについては十分なツイート数が存在しないため、「多数決」での分類結果が 50 件以上存在し、評価者間の一致率も比較的高かった「(1) 予定」「(2) 希望」のみ分類精度を検証することとした。

分類器の素性として、形態素\*9の原形 (1-gram, 2-gram) を用いて、SVM により学習を行った。本研究では、TinySVM\*10 を利用し、学習には多項カーネル ( $d=2$ ) を、パラメータはデフォルト値を用いた。

分類精度については、10 分割交差検定により検証した。図 2 に「(1) 予定」カテゴリに関する検証結果を、図 3 に「(2) 希望」カテゴリに関する検証結果をそれぞれ示す。いずれのカテゴリも、1-gram のほうが精度が高い。1-gram における F 値は、「(1) 予定」カテゴリにおいて 0.465、「(2) 希望」カテゴリにおいては 0.607 となり、「(2) 希望」カテゴリについては比較的良好な結果が得られた。とくに、適合率 (Precision) が 0.745 であり、誤った抽出は少ないと考えられる。

表 3 に、「希望」カテゴリにおける重みの大きい素性の一覧を示す。表 3 より、「希望」カテゴリにおいては、希望を意図する表現である「たい」の重みが大きい。4.3 節で述べたように、「希望」カテゴリにはこのような表現を含むものが多いため、高精度に分類できたと考えられる。したがって、今回構築した分類器によって、「(2)

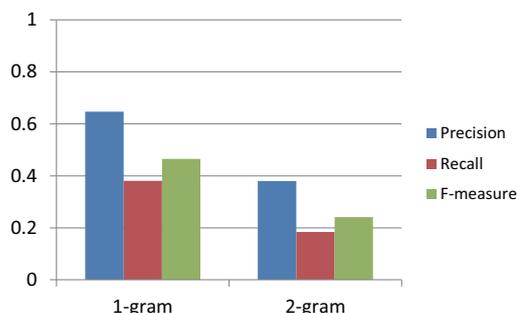


図 2 「(1) 予定」カテゴリに関する 10 分割交差検定結果

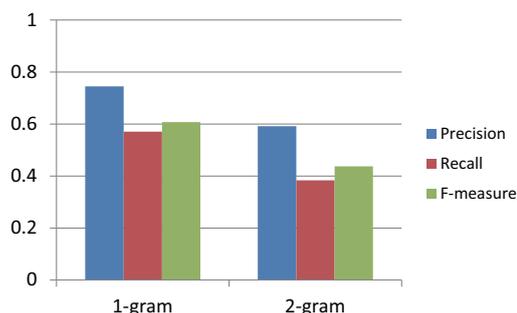


図 3 「(2) 希望」カテゴリに関する 10 分割交差検定結果

希望」カテゴリに該当する To-Do 候補を、ある程度高精度に抽出可能であることが示唆される。

## 6. おわりに

本稿では、To-Do 管理促進のためのマイクロブログ上の To-Do 自動抽出手法を提案した。本研究の貢献は以下の 3 点にまとめられる。

- (1) To-Do 管理の利用促進のため、利用者の入力負担を軽減する方法として、マイクロブログからの To-Do 自動抽出手法を提案した。
- (2) To-Do 管理のためのカテゴリを定義し、Twitter のツイートデータの分類によりコーパスの構築を行った。
- (3) 複数評価者間の一致率が高かったカテゴリについて分類器の構築を行った結果、「希望」カテゴリについては、比較的高精度に判定可能であることを示した。今後は、本手法についての見直しを行い、各 To-Do カテゴリを高精度に判定するための手法を検討する。また、To-Do 自動抽出手法と To-Do 管理システムとの連携を進める。さらに、以下の点を考慮した長期間の評価実験から、To-Do 管理ユーザに対する影響を検証する。

- ユーザの「To-Do 管理の習熟度」と「Twitter への公開欲と投稿数」
- To-Do 管理システムや Twitter のリプライやハッシュタグを利用した意識的な To-Do 管理との比較

\*9 形態素解析には JUMAN (<http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>) [18] を用いた。

\*10 <http://chasen.org/~taku/software/TinySVM/>

## 参考文献

- [1] 日立ソリューションズ: 20代の時間管理に対する意識の調査結果, <http://www.hitachi-solutions.co.jp/column/tashinami/jikan/> (参照 2014-08-25).
- [2] 水口和彦: 残業ゼロ! 時間管理のコツ39, 学研パブリッシング (2012).
- [3] 水口和彦: 世界で一番ゆるい王様の時間術, ダイヤモンド社 (2010).
- [4] 谷岡遼太, 宮部真衣, 吉野孝: To-Do 管理のためのマイクロブログ上の To-Do 自動抽出手法の提案, 平成 26 年度情報処理大会関西支部大会, E-01, pp.1-4 (2014).
- [5] 佐々木正悟, 大橋悦夫: スマホ時代のタスク管理「超」入門, 東洋経済新報社 (2013).
- [6] Microsoft: Survey Shows Increasing Worldwide Reliance on To-Do Lists, <http://www.microsoft.com/en-us/news/press/2008/jan08/01-14NGOMPR.aspx> (参照 2014-08-25).
- [7] 日本能率協会マネジメントセンター: 「あなたの手帳の流儀」調査, <http://www.jmam.co.jp/new/newsrelease/> (参照 2014-08-25).
- [8] マイナビ: 2015 年卒 マイナビ大学生のライフスタイル調査, [http://saponet.mynavi.jp/enq\\_gakusei/lifestyle/](http://saponet.mynavi.jp/enq_gakusei/lifestyle/) (参照 2014-08-25).
- [9] 谷岡遼太, 吉野孝: タスクの公開揭示による ToDo リスト利用促進システム, 情報処理学会, エンターテインメントコンピューティング 2013 (EC2013), pp.196-199 (2013).
- [10] 濱崎雅弘, 武田英明, 大向一輝, 市瀬龍太郎: 学術会議における共有型スケジュールリング支援の開発と運用, 日本データベース学会 letters, Vol.2, No.4, pp.7-10 (2004).
- [11] 大向一輝, 武田英明: 人間関係ネットワークに基づく情報フィルタリングを用いた協調的タスクスケジューラ, 電子情報通信学会論文誌, Vol.J87-D1, No.11, pp.1020-1029 (2004).
- [12] 村益有那, 實雪: ソーシャルネットワーキングサイト上における若者の自己開示と感情表現に関する研究 - Twitter と Facebook の内容分析及び大学生へのインタビュー調査から -, 第 31 回情報通信学会大会, 第 1 会場, pp.1-31 (2014).
- [13] M. Naaman, J. Boase, and C. Lai: Is it Really About Me?: Message Content in Social Awareness Streams, CSCW'10, Computer supported cooperative work, pp.189-192 (2010).
- [14] ミンテイゲン, 川村隆浩, 田原康之, 大須賀照彦: Twitter から人間行動属性の自動抽出, 電子情報通信学会技術研究報告, 人工知能と知識処理, Vol.110, No.105, pp.19-23 (2010).
- [15] C. Cortes, and V. Vapnik: Support-Vector Networks, Machine Learning, Vol.20, No.3, pp.273-297 (1995).
- [16] J. Lafferty, A. McCallum, and F. Pereira: Conditional Random Fields: Probabilistic models for segmenting and labeling sequence data, ICML, pp. 282-289 (2001).
- [17] David Allen: Getting Things Done: The Art of Stress-Free Productivity, Penguin Books (2012).
- [18] K. Daisuke, and S. Kurohashi: A Fully-Lexicalized Probabilistic Model for Japanese Syntactic and Case Structure Analysis, In Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL2006), pp. 176-183 (2006).