

卒業論文 2017年度（平成29年度）

Twitterを用いた都道府県別での名産品評価手法 の構築

慶應義塾大学 環境情報学部

大刀祢 楓斗

徳田・村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・
武田 合同研究プロジェクト

2017年1月

Twitterを用いた都道府県別での名産品評価手法の構築

論文要旨

近年、地域ごとに名産品を元にした地域振興が注目されている。多くの自治体で地域の資産である農・水産業により得られる食材や、それらを用いた食製品などを生産・販売し、地域振興を目指して取り組みがなされている。例えば静岡県では緑茶の一大生産地であることを活かしたブランド開発に注力している。高知県馬路村では名産品である柚子を用いた食製品の販売で通信販売により多大な地域振興を遂げている。これらは自治体が地元の名産として推進したものが、実際に知名度を得た成功例である。しかしこれらは有名な成功例であり、自治体ごとの名産品全てが同様に知名度をあげられているわけではない。名産品の宣伝は自治体の広報などの人に依存するところが大きく、実際の世間の認知度には差異がある。このため、地元の特色を反映した品としては優れていても思うように知名度をあげられていないものが存在すると考えられる。現地の地方行政に関わっている人や、実際に訪れて名産品を見て、食べて感じ取った人でないと実情を把握することは難しい。それを表面化させ、地方自治体の支援をする手法を確立することが本研究の目的である。そのような地方名産品の状況を探るためには人々の居場所と心情に関する情報が即座に取得できる必要がある。その手法の一つとして様々な既存研究が存在する SNS 分析を挙げられる。位置情報は SNS での投稿にも付与でき、短時間で高頻度に投稿できるため、現実での出来事について位置情報付きの投稿を行う人・場面が増えているのである。そのため本研究では地域別の名産品の実情を SNS を通じて知ることが出来ると考え、投稿のテキスト分析によって評価を行った。

キーワード

位置情報, 地方, 名産品 マイクロブログ, WordNet, 形態素解析, 極性分析

**Construction of evaluation method by prefecture by
Twitter**

Summary

These are examples that the municipality promoted as local specialty products actually succeeded. However, these are famous success examples, not all famous products are similarly famous. The promotion of special products is largely dependent on people, there are variations in the real awareness of the world. For this reason, it is thought that there are things that are not well known. It is difficult to understand the fact that of specialty products except people who visited, looked and ate and people who are involved in the local administration. It is the purpose of this research to make it surface and to establish a method to support local governments. In order to investigate the situation of such local products, it is necessary to immediately obtain information on people's place and feelings. One of the methods is analysis of SNS which has various existing studies exist. Position information can also be given to posts on SNS and it can be posted frequently in a short time, so people and scenes who post postings with location information on real events are on the rise. Therefore, in this research, I thought that I could know the actual condition of regional specialties through SNS, I verified and evaluated it.

Keywords

geotag, region, local specialty, microblog, WordNet, Polarity analysis

Bachelor of Arts in Environment and Information Studies
Keio University

Futo Otone

目次

第1章	序論	1
1.1	本研究の背景	1
1.2	地方名産品の成功例	1
1.2.1	高知県・ゆず	2
1.2.2	宮崎県・マンゴー	2
1.2.3	兵庫県・神戸ビーフ	3
1.2.4	北海道・夕張メロン	3
1.2.5	栃木県・宇都宮餃子	3
1.3	地方名産品の抱える問題	5
1.4	本研究の目的	5
1.5	本論文の構成	5
第2章	SNS 分析の既存研究	6
2.1	SNS・位置情報を用いた既存研究	6
2.1.1	Twitter Catches the Flu	6
2.1.2	Twitter mood predicts the stock market	7
2.1.3	Visual Analysis of Ego-centric Information Diffusion Patterns in Social Media	7
2.1.4	SNS data Visualization for analyzing spatial-temporal distribution of social anxiety	8
2.1.5	Earthquake shakes Twitter users: real-time event detection by social sensors	9
2.2	SNS・位置情報を用いたサービス	10
2.2.1	Foursquare	10
2.2.2	PokemonGO	10
2.2.3	Twimp	11
2.2.4	TPS	11
2.3	関連研究のまとめと提案する手法	11
第3章	設計	13
3.1	提案手法の説明	13
3.2	WordNet	14
3.3	形態素解析	15
3.4	極性分析	16

第4章 実装	17
4.1 概要	17
4.2 実装環境	17
4.3 Tweet Scraper	17
4.4 Reverse Geocoder	18
4.5 Analyzer	20
4.6 用いる WordNet	21
第5章 評価	24
5.1 評価結果	24
5.2 考察	26
第6章 結論	27
6.1 まとめ	27
6.2 今後の課題と展望	27
謝辞	28
参考文献	29

図目次

1.1	高知県内の種類別栽培面積のグラフ [1]	2
1.2	実際の餃子マップ [13]	4
2.1	株価平均と5つの感情の相関をグラフにした図 [4]	7
2.2	不安に関する投稿をプロットし,それを可視化した図 [6]	8
2.3	地震に関する Tweet をプロットし,震源地を予測した図 [7]	9
2.4	実際に街でポケモンに出会う様子 [8]	10
2.5	富岡八幡宮を例に検索を行った場合の図 [9]	11
3.1	システムフローチャート図	14
3.2	WordNet の例 [11]	15
3.3	形態素解析の例	16
3.4	極性辞書 (ポジティブ) の例 [12]	16
3.5	極性辞書 (ネガティブ) の例 [12]	16
4.1	Yahoo!リバーズジオコーダ API のリクエストパラメータ一覧 [20]	19

表目次

3.1	取得ツイートの形式	13
4.1	実装環境	17
4.2	WordNet に含む語一覧 1	21
4.3	WordNet に含む語一覧 2	22
4.4	WordNet に含む語一覧 3	23
5.1	評価表 1	24
5.2	評価表 2	25

第1章 序論

1.1 本研究の背景

近年,日本各地において観光振興や地域活性化のために自治体や地域団体がその地域において多く収穫・生産されるものを食製品に加工したり,独自の名産品として売り出す事例が多数存在する.また,「B級グルメ」として地域独特の名物や郷土料理を開発・販売し地域振興を行っている.たとえば富士宮市の「富士宮焼きそば」や青森県八戸市の「せんべい汁」,栃木県の「宇都宮餃子」,岩手県盛岡市の「盛岡じゃじゃ麺」などが挙げられる.そして地域振興の一環としてこのようなB級グルメを集めた祭典が年に一回毎年地域を変えつつ開催されている.このように地域名産品による地方を盛り上げようとする取り組みはさらなる広がりを見せている.2006年には「商標法」の一部である「地域団体商標制度」が創設された.これは地域の名称や商品名を組み合わせた名称を組合を通じて商標登録が受けられるという制度であり,地域ブランドを保護しながら更なる地域振興を促進することが目的である.これに登録されている実例としては「大間まぐろ」,「比内地鶏」,「松阪牛」,「長崎カステラ」などが存在する.このように国の支援もありながら,地方振興としてその地方特有の食材を用いた名産品による取り組みは盛り上がりを見せていることが言える.しかしそのような取り組みはすべての事例において成功しているわけではない.中には宣伝方法やその売り出し方から,あまり知名度が思うようにあげられていないものも存在すると考えられる.さらにそのような地域名産品に関する実態は表面化しづらく,すべての地域に対してアンケート調査をすることも現実的ではない.そこで個人から直接発信された意見等を交換する場であり,特定個人の嗜好や当人の実際の意識などに関する情報が多く含まれているマイクロブログに着目した.マイクロブログは情報の更新頻度が高くユーザ自身の行動情報や思考を頻繁に発信する媒体として注目されており,人間の行動分析や心情分析の研究分野で利用されている(選挙の趨勢を予測するもの[15],インフルエンザの流行を予測するもの[3],株価を予測するもの[4]など).テキスト自体の分析手法には形態素解析や分散表現が用いられる場合が多く,様々な研究分野のベースラインの手法と比べて成果が出ている事例が多く存在する.そのため,多数存在する地域の名産品をマイクロブログを通じてその実情を検証できるのではないかと考えた.

1.2 地方名産品の成功例

本研究では日本各地の名産品に対して地方自治体や地域団体が名産品として宣伝し,推薦しているものには名産品ごとに知名度の大きな差異があることに着目した.本節ではその中

でも成功したものの例について紹介する。

1.2.1 高知県・ゆず

以前よりゆず栽培が盛んであり、その販売を模索している中で、ゆずの青果を販売するよりもゆず酢等の加工品の開発に着手し、脳教職員を中心としたマーケティング活動を開始した。販売に際して単にゆずを商品として売るだけではなく、村のイメージをマーケティングに用いて、一商品であったゆずのジュースに「ごっくん馬路村」と名付けることで促進を行った。また、無農薬であることや田舎のゆっくりとしたイメージを付けつつテレビCMも打ち出した。結果、ゆずの生産量は1975年の95tから2006年の700tに増加、売上は1976年の約3000万円から2006年の33億4000万円に増加した。今では地方創生の成功例として県外から多数の視察団体が訪れるほどになっている。

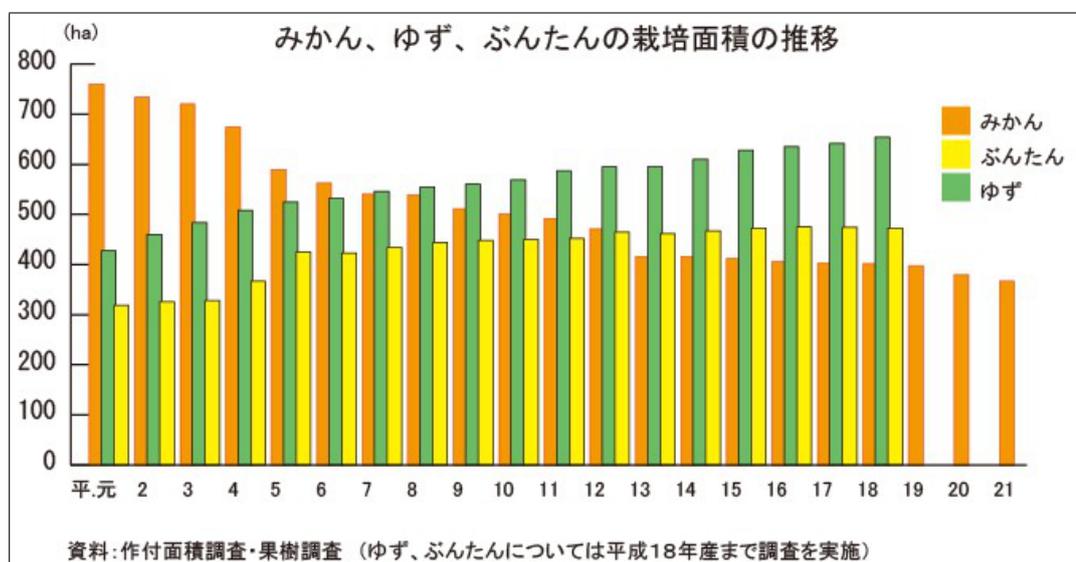


図 1.1: 高知県内の種類別栽培面積のグラフ [1]

1.2.2 宮崎県・マンゴー

1993年に宮崎県で県内のマンゴー栽培農家を主な構成員とした「宮崎県果樹振興協議会亜熱帯果樹部会」が発足した。これが主体となって生み出した「太陽のタマゴ」と銘打った「みやざき完熟マンゴー」が売り出されることになった。当初は県内の販売がほとんどであったが、2007年、東国原英夫氏の県知事就任によりテレビでのPR活動の機会が増え全国区の商品となった。結果、1998年の年間生産量100tから2014年には年間生産量1100tに増加している。

1.2.3 兵庫県・神戸ビーフ

1983年に兵庫県内では肉を生産、流通、消費する人達が「神戸牛流通推進協議会」を結成した。これにより、かねてから一定の知名度はあったものの明確な定義のなかった神戸ビーフの種類や品質が定められた。神戸ビーフとは但馬牛の中でも優れた品質を持つものとされ、神戸牛流通推進協議会の会員が兵庫県内の牛舎で肥育した牛であることが決定された。商標権を取得し、販売ルートを指定点に限定して、毎月の巡回調査で改善指導等を行うことにより管理を行うものとした。結果、子牛の販売価格が2001年4月の37万円から2007年4月には53万円に増加した。[1]

1.2.4 北海道・夕張メロン

元々市内で「夕張キングメロン」の品種で生産されていた赤肉系のメロンであり、「夕張市農業共同組合」が管理の下、同市内で生産されたメロンのみを厳正な検査にかけて合格したものを「夕張メロン」のブランドとし販売を行った。1982年には大手宅配業者、フットワークと提携して産地直送化することにより、欠点とされていた日持ちの短さを克服して全国出荷が可能になった。札幌市中央卸売市場の初競りでは2016年には二玉300万円の値がつくほどであった。[1]

1.2.5 栃木県・宇都宮餃子

宇都宮餃子はもともとは地元の食文化として戦後の頃から市民に親しまれてきた食品である。これを市役所が町おこしとして利用できないかと考え、多数の餃子専門店に働きかけ、市内の餃子店一覧を示した「餃子マップ」の作成や、「宇都宮餃子会」の発足に至るまでになった。以後、多数のテレビ取材やそれに乗じた「宇都宮餃子祭り」の開催などによって観光客の増加が見られた。平成13年には「宇都宮餃子会」が法人化され、産地の管理や登録商標の管理などを行っている。現在では宇都宮餃子は栃木県の大観光名物となり、一般の認知度は非常に高いものとなっている。[1]

1.3 地方名産品の抱える問題

前節では成功例について見たが、そのような事例の成功は当然名産品自体が優れたものであるためだが、宣伝広報に依るところも影響している。地方自治体・団体が推進していくという以上は大規模なコマーシャルや広告を出すことが難しく、宣伝活動は自治体の広報などの人に大きく依存してしまう。したがって、もの自体が優れていたとしても知名度が上がらず、大衆の目に止まらないものがあることを問題視される事態が生じる。本研究ではこの問題を主題に添え、SNS 解析を用いた提案手法を検証する。

1.4 本研究の目的

地方自治体・団体による名産品事業は増えつつあり、地方団体商標制度の創設により商標の登録も可能になった。このため地域振興が盛り上がりを見せていることは事実だが、必ずしもすべての名産品が一部の大成功例と同様の結果を得られるわけではない。名産品自体が市場での発展を約束するものではないからである。市場での発展のためには大衆の不認知を認知の段階に変化させることが必要である。本研究の主な目的は問題として取り上げた、「知名度を思うようにあげられていない地方の名産品」を表面化させ、地方自治体・団体の広報活動のサポートとなる手法を確立することとする。

1.5 本論文の構成

本論文は全6章で構成される。第1章で着目する問題について述べ、背景と目的、地域名産品の成功事例について説明する。この背景についてさらに説明するために第2章で位置情報やソーシャルネットワークサービスを用いた既存研究の紹介、及びそれに関連する技術について論じる。第3章では本研究で用いる提案手法についての説明を行う。第4章では提案手法の実装について実際のコードを説明しながら見ていくものとする。第5章では3章で見た手法を実際に検証した上での結果を評価する手法の説明とその考察を論じ、第6章では本研究で提案した手法を評価した結果から得られる結論と今後への課題を述べる。

第2章 SNS分析の既存研究

この章では本研究の前提となる概念である位置情報とマイクロブログ等のソーシャルネットワークサービスについて論じ、これらの関連研究と実際のサービスの実用例について示す。これによってSNS分析の手法としての有用さを示し、本研究の立ち位置を確認する。また、その概要を実際の例を紹介しつつ見ていくものとする。

2.1 SNS・位置情報を用いた既存研究

SNSとはSocial Network Serviceの略称であり、TwitterやWeiboをはじめとするコミュニケーションを主としたサービスであり、多数のユーザ数とその社会性を利用して多くの研究が行われている。本節ではそれに関連する研究を紹介し、本研究の立ち位置を確認する。

2.1.1 Twitter Catches the Flu

Twitterからインフルエンザに関するツイートを抽出し、SVM(support vector machine)を用いた分類器でその投稿者がインフルエンザにかかっているかを判定した。判定に際して、確実に感染している陽性発言と語句を含んでいるだけで感染していることを確実に指しているわけではない陰性発言に分類する。ここから分類器で陰性発言を取り除いた。陽性発言を単位時間ごとに収集すると現実のインフルエンザの患者数との相関関係が0.89であった。これはGoogleが公式に発表している精度とほぼ同等のものである。しかし過熱報道時には低い精度を示し、ニュースによるバイアスを受けていることが判明した。[3]

2.1.2 Twitter mood predicts the stock market

Twitter から感情に関わるツイートを抽出し、それを用いて株価を予測するものである。予測に際して、感情を (calm, alert, sure, vital, kind, happy) に分類しその時系列データを得てダウ平均との相関を見た。行動経済学の観点から感情は行動に先んじるという推論を立て、アメリカ大統領選挙時と感謝祭のケースを例に実際に検証を行った。calm が最も株価予想精度が高かったという結果を得ている。[4]

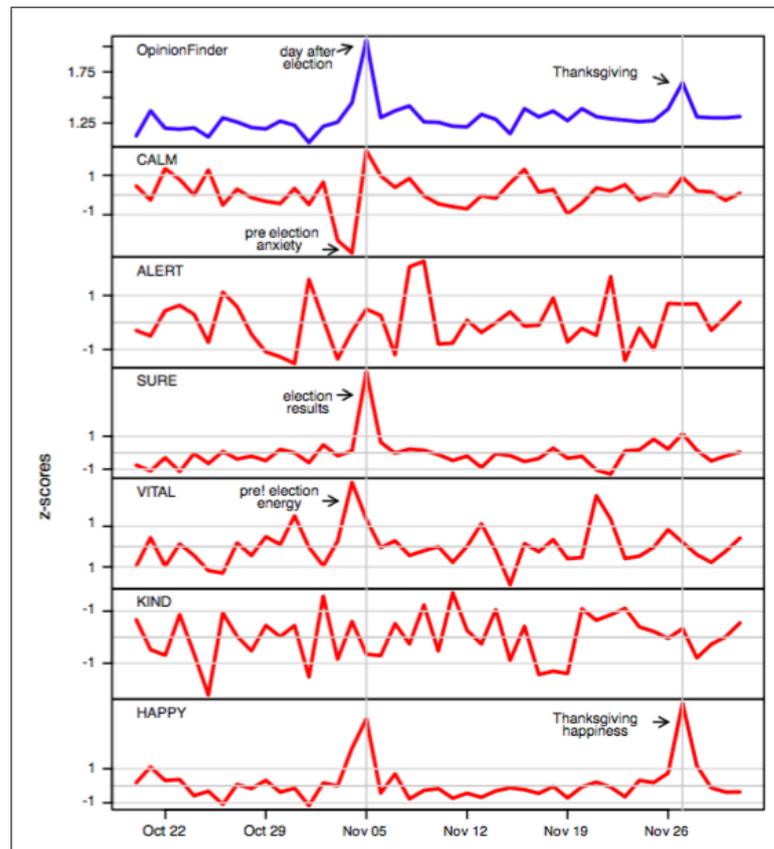


図 2.1: 株価平均と 5 つの感情の相関をグラフにした図 [4]

2.1.3 Visual Analysis of Ego-centric Information Diffusion Patterns in Social Media

weibo からテキスト・タイムスタンプ・id を抽出し、情報がどのように普及していくのかを検証した。検証に際して、情報の普及開始点となる影響力のあるアカウントをオピニオンリーダーとしてコミュニティごとに指定しその波及を見た。少規模のコミュニティではオピニオンリーダーはサービスアカウントであること、大規模なコミュニティに属し、他のコミュニティには影響をあまり持たないオピニオンリーダーはビジネスマンであることなどが判明した。[5]

2.1.4 SNS data Visualization for analyzing spatial-temporal distribution of social anxiety

Twitter から位置情報を含むツイートを抽出し,Machine Learning を用いた不安分類器を用いてツイートの不安度を検証し,それを web アプリケーション上で地図とワードクラウドとして可視化する. 検証に際して,分類機の構築のため不安度の確率を導き出し,適した閾値を見つけ出した. 構築した分類機の精度は 85 パーセントであった. アンケートなどでは書けない本心を把握するための世論調査としての側面も持った研究であった. 不安度数と位置情報とともに表現したことは新しく,社会学や政治学など様々な分野での応用が期待されている.[6]

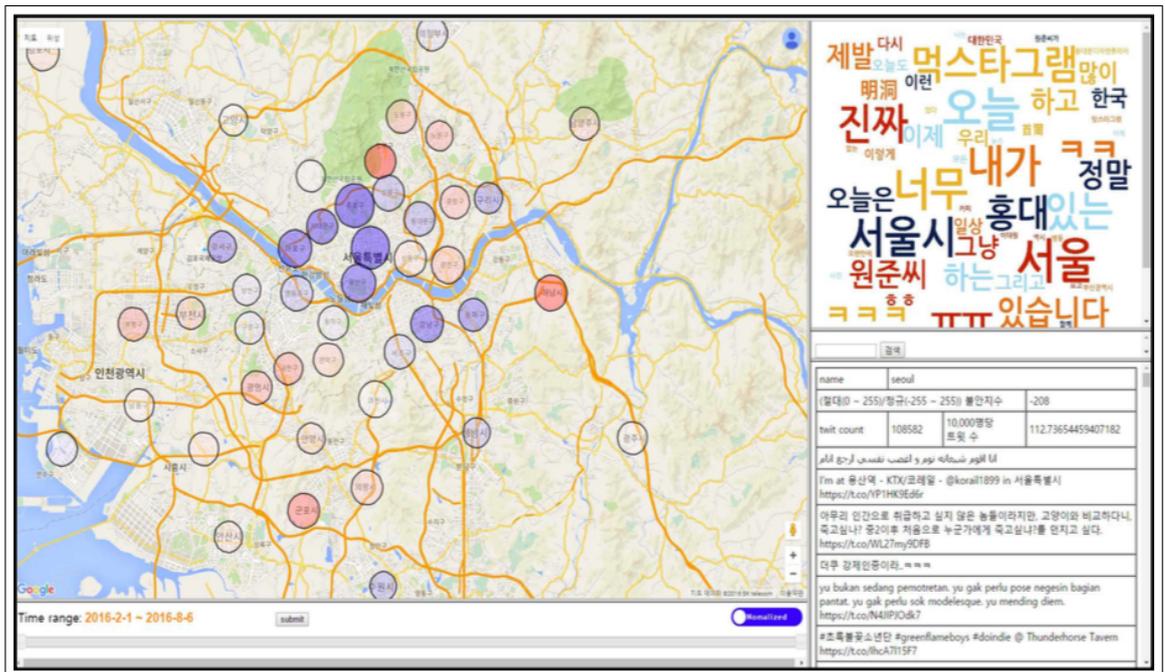


図 2.2: 不安に関する投稿をプロットし,それを可視化した図 [6]

2.1.5 Earthquake shakes Twitter users: real-time event detection by social sensors

Twitter をリアルタイムセンサーとして監視し, 自身や台風などの災害が起こったことを検知する手法を検証した. 気象庁などの機関の発表よりも早く情報を得られ, SNS の即時性が改めて確認できる研究である.[7]

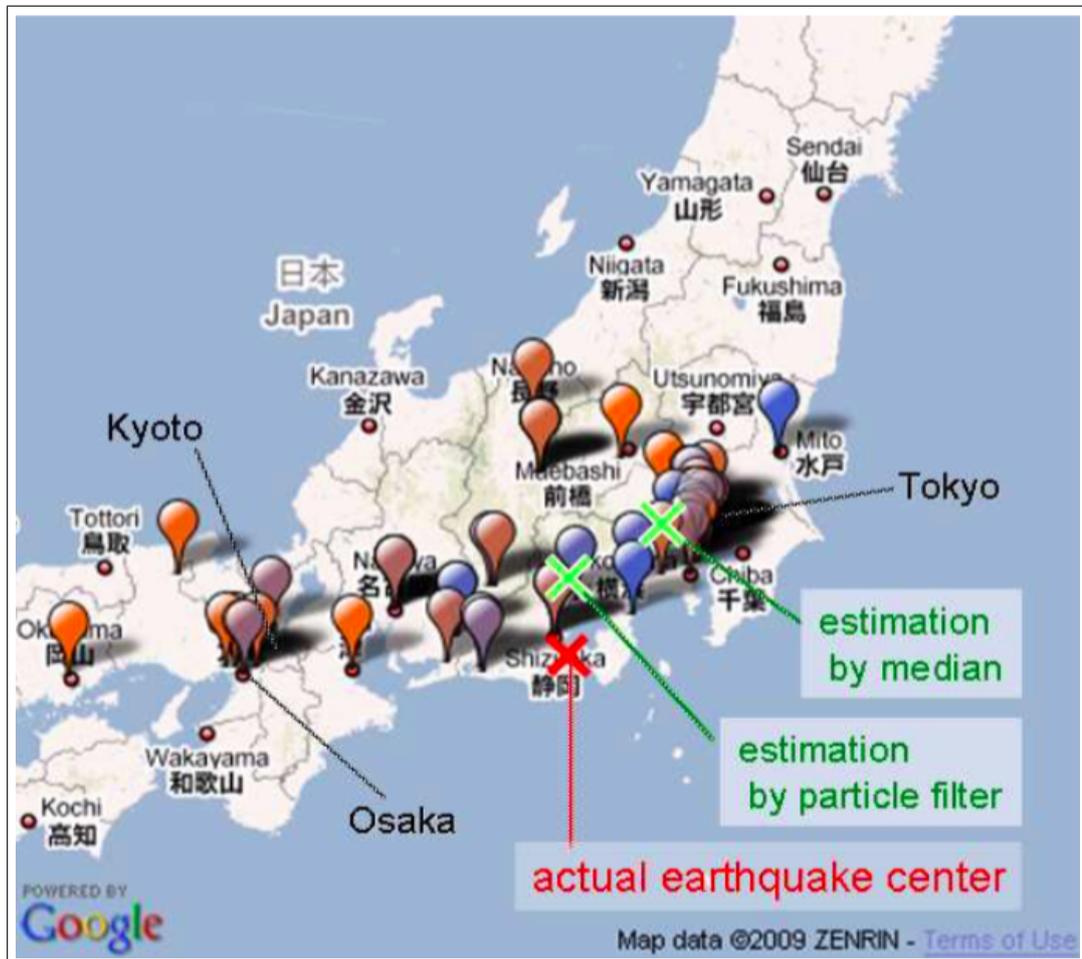


図 2.3: 地震に関する Tweet をプロットし, 震源地を予測した図 [7]

2.2 SNS・位置情報を用いたサービス

今節では前節に引き続き,SNS・位置情報の使用事例について見ていく.前節では研究事例について見たが,今節では実際のサービスやアプリケーションについて見ていく.

2.2.1 Foursquare

Foursquare(フォースクエア)とは携帯端末のGPSから取得できる位置情報を用いたソーシャル・ネットワーキングサービス・位置ゲームである.現実の店舗や施設などに実際に訪れたり場所を登録することで様々な種類のバッジを取得したり,メイヤーの称号を得ることが出来る.このゲームを利用してポイントを集めることで実際の店舗での割引クーポンが使えるなどの連携もなされている.日本ではローソンとの提携も行っている.

2.2.2 PokemonGO

Pokemon GOとはスマートフォン向けの位置情報ゲームアプリである.スマートフォンのGPS機能を使用しながらユーザが移動することでポケットモンスターのキャラを捕獲・育成するものである.実際の店舗や施設に訪れることで様々なアイテム補給を行えるポケストップという拠点もあり,マップ上のあらゆる場所に配置されている.

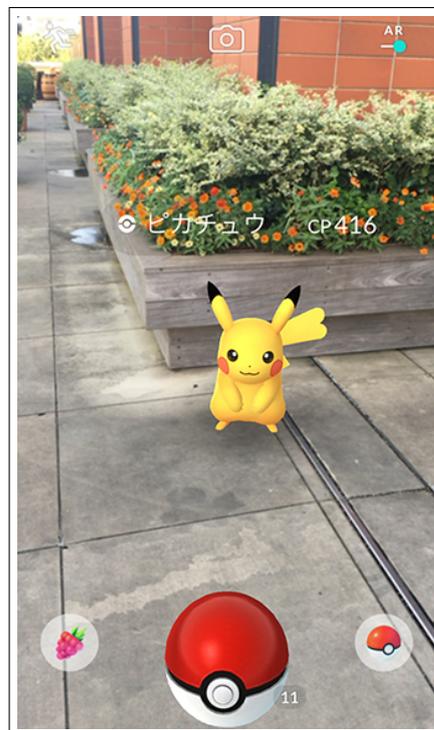


図 2.4: 実際に街でポケモンに出会う様子 [8]

2.2.3 Twimp

twimpとはキーワードを入力することでそのキーワードを用いたツイート回数や表示回数, ユニークユーザー等を調べることができる web アプリケーションである. 実際にツイートした人のフォロワーからどれだけキーワードがタイムライン上で表示されたのか, bot はどれくらいなのかを調べることができ, その結果を図解にすることでどのユーザーで情報の拡散が見られるのか, 話題のキーマンであったのかを一目で判別できるようになっている.[9]

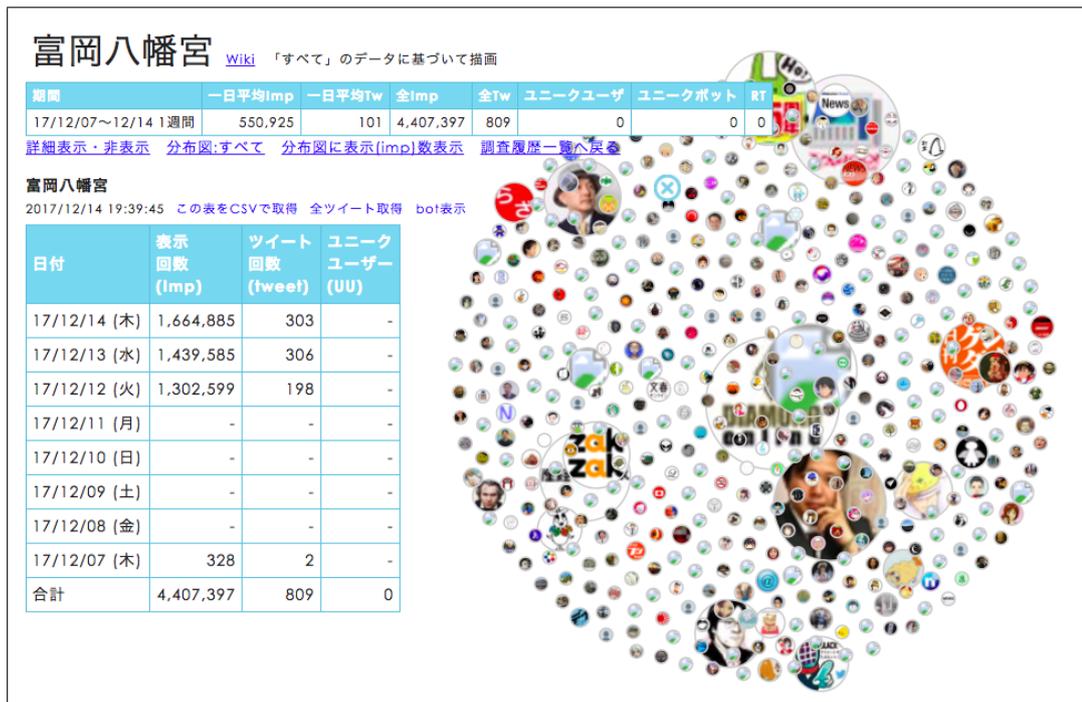


図 2.5: 富岡八幡宮を例に検索を行った場合の図 [9]

2.2.4 TPS

TPSとはTweets-Per-Secondの略称であり, 秒間あたりのツイート数を表す. この数値が高いと, その瞬間に多くの人達があるイベントや出来事についてツイートをしたということを示す. 2013年8月にテレビ放映された『天空の城ラピュタ』のラストシーンでは最大で秒間14万3199ツイートというTPSを達成した. これは世界のTPSランキングで1位の記録であり, このランキングの上位は殆どを日本が占めるという結果になっている.

2.3 関連研究のまとめと提案する手法

関連研究と関連サービスによってソーシャルネットワークサービス, 位置情報を使用した事例について紹介した. 本研究で提案する手法はTwitterの位置情報付きの投稿を収集し, 地

方単位別に名産品に関するツイートを分析することで名産品ごとの評価を得るものである。

位置情報を用いることでユーザー同士に場所を共有させたり、そこから様々な知見を得られることを本章で述べた。本研究は1章に述べた地方の問題を踏まえた上でTwitterを用いた評価手法を提案する。

第3章 設計

本章では検証に際して、本研究で用いる実験手法とその設計について述べる。また、本論文で用いる手法の主要技術である WordNet や形態素解析、極性分析についても述べるものとする。

3.1 提案手法の説明

Twitter から位置情報が付与されたツイートを収集し、csv 形式で保存する。収集したツイートの中から地方各地の名産物に関する単語を含むものを取得する。抽出したツイートのテキストに対して極性分析を行い、その結果を名産物ごとの評価点とする。どの地方の名産物かは抽出したツイートに付与されている位置情報から判別するものとする。

ツイートの取得には Twitter API を利用した。取得するツイートの形式は、ID、timestamp、テキスト、位置情報 (緯度経度) である。

表 3.1: 取得ツイートの形式

1	2	3	4
ID	timestamp	text	位置情報 (経度, 緯度)

取得したツイートの中からテキストと位置情報を取り出し、検証に用いる。位置情報の (緯度, 経度) から日本国内の都道府県の場所を割り出し、特定した都道府県における名産品を選び出す。その名産品の名称に対して WordNet を構築し、その名産品が下位に含む言葉とそれ自身の単語を含むツイートを元のデータセットから抽出する。取り出したツイートのテキストに対して極性分析を行いその文章がポジティブなものかネガティブなものかを数値で算出する。この数値がポジティブに寄っている場合、その名産品は当該県の名産として評価が高いものとみなす。逆に数値がネガティブに寄っている場合はその名産品は当該県の名産として評価が低いものとみなすことにする。以上の操作を全都道府県別・各名産品三種類に対して行って、国内都道府県名産品の実態を調査する。また、各都道府県別の名産品は農林水産省の『特産農作物の生産実績調査』、独立行政法人国立青少年振興機構が提供する『特産品 map』、価格.com の提供する『特産品マップ』を参考に筆者が総合的に判断して決定した。

以下に前述した手法の処理の流れを示す。

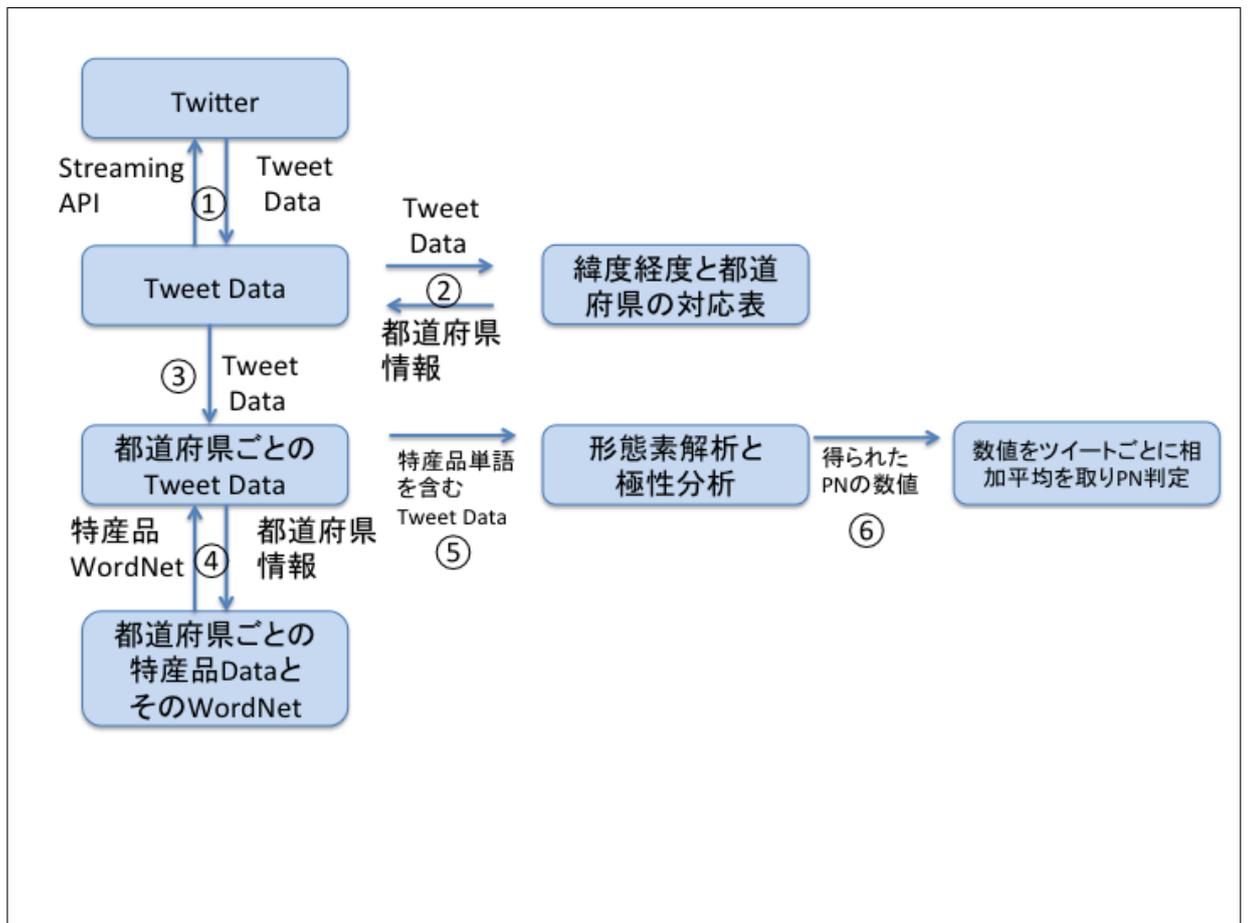


図 3.1: システムフローチャート図

3.2 WordNet

WordNet とは英語の概念辞書のことであり、自然言語処理研究においては有用な概念であるため、さまざまな研究で利用されている。一般の利用者を対象として Web 上の類義語辞書サービスとして日本語 WordNet[16] も提供されている。これは独立行政法人通信研究機構 (NICT) により作成されたものであり、2006 年から大規模かつ誰でも使用可能な日本語意味辞書の作成を目指して開発が行われている。

言葉が上位・下位の関係によって定義されており、他の同義語のグループとの関係も示されている。例えば「犬」は上位概念の「哺乳類」に属し、「哺乳類」は更に上位概念である「動物」という概念に属する。本研究の実験では対象の名産品に対して WordNet を構築し、それに含まれる言葉を含んでいるツイート抽出して分析を行う。つまり名産品の分析対象ツイートを増やすために用いるものとする。

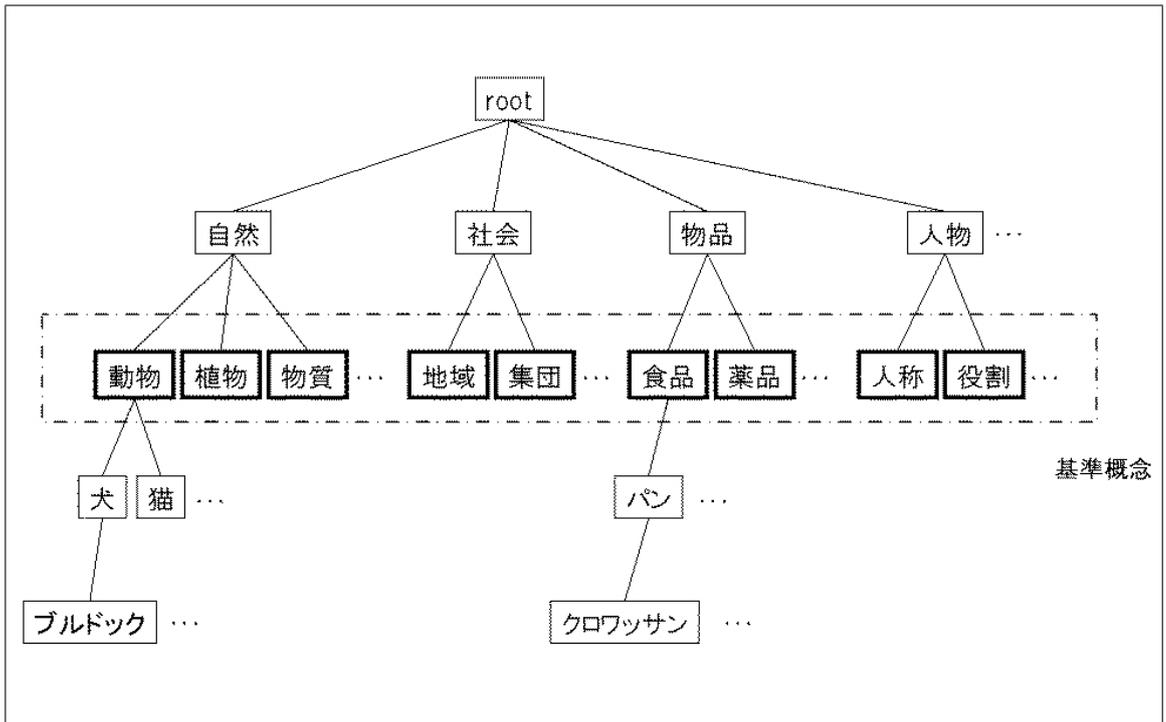


図 3.2: WordNet の例 [11]

3.3 形態素解析

形態素解析とは特に文法的な情報の記載がされていない自然言語のテキストデータから、対象言語の文法や品詞等の情報にもとづき、形態素(言葉単体で意味を持つ最小単位)に分割し、それぞれの形態素へ情報を付与したり、品詞を判別する作業のことである。主要な形態素解析ソフトとしては『MeCab』[14]、『CaboCha』[18]、『JUMAN』[19]などが挙げられる。本研究では京都大学情報学研究科・日本電信電話株式会社コミュニケーション科学基礎研究所共同研究ユニットプロジェクトを通じて開発されたオープンソース形態素解析エンジン『MeCab』を使用する。事前に専門用語や近年の流行語などを辞書として記憶させることも可能であり、辞書の参照によって意味のある語句で区切り、更に精度を上げて解析が行われる。『MeCab』は係り受け解析も行うことができ、汎用的な解析ソフトといえる。実験に際して、解析を行ったテキストの形態素の基本形を抽出してそれが極性辞書の中に存在するかどうかを走査し、存在すれば対応するPN値について相加平均を取る処理をツイートごとに行う。本研究の実験では主に言葉の基本形を取得するために形態素解析を行う。これは実際のツイート中に現れる未然形や連用形のように活用された形だと極性辞書をサーチすることができないためである。形態素解析処理の例として以下に例文「私は慶應義塾大学の湘南藤沢キャンパスで情報技術について学んでいます。」を解析にかけた結果を下図に示す。

私は慶應義塾大学の湘南藤沢キャンパスで情報技術について学んでいます。	
私	名詞,代名詞,一般,*,*,*私,ワタシ,ワタシ
は	助詞,係助詞,*,*,*は,ハ,ワ
慶應義塾	名詞,固有名詞,組織,*,*,*慶應義塾,ケイオウギジユク,ケイオーギジユク
大学	名詞,一般,*,*,*大学,ダイガク,ダイガク
の	助詞,連体化,*,*,*の,ノ,ノ
湘南	名詞,固有名詞,地域,一般,*,*湘南,シヨウナン,シヨーンナン
藤沢	名詞,固有名詞,地域,一般,*,*藤沢,フジサワ,フジサワ
キャンパス	名詞,一般,*,*,*キャンパス,キャンパス,キャンパス
で	助詞,格助詞,一般,*,*,*で,デ,デ
情報	名詞,一般,*,*,*情報,ジヨウホウ,ジヨーホー
技術	名詞,一般,*,*,*技術,ギジユツ,ギジユツ
について	助詞,格助詞,連語,*,*,*について,ニツイテ,ニツイテ
学ん	動詞,自立,*,*五段・バ行,連用夕接続,学ぶ,マナン,マナン
で	助詞,接続助詞,*,*,*で,デ,デ
い	動詞,非自立,*,*一段,連用形,いる,イ,イ
ます	助動詞,*,*,*特殊・マス,基本形,ます,マス,マス
。	記号,句点,*,*,*。 ,。 ,。

図 3.3: 形態素解析の例

3.4 極性分析

極性分析とは人の発言が前向き(ポジティブ)か後ろ向き(ネガティブ)かを判定することである。判定には文章等に含まれる感情評価に関する表現や言葉を抽出して全体としての感情を解析する。Twitter が提供している検索オプションにも実装されており、ポジティブ・ネガティブといった条件設定が可能である。本研究の極性分析では東北大学の乾・岡崎研究室の提供する日本語評価極性辞書 [11] を用いるものとする。この極性辞書は単語に対応する極性情報が-1 を下限,+1 を上限として割り当てられており,-1 に近いほどネガティブ,+1 に近いほどポジティブということになっている。

良い:よい:形容詞:0.999995
喜ぶ:よるこぶ:動詞:0.999979
褒める:ほめる:動詞:0.999979
めでたい:めでたい:形容詞:0.999645
賢い:かしこい:形容詞:0.999486
善い:いい:形容詞:0.999314
適す:てきす:動詞:0.999295
天晴:あっぱれ:名詞:0.999267
祝う:いわう:動詞:0.999122
功績:こうせき:名詞:0.999104
賞:しょう:名詞:0.998943
嬉しい:うれしい:形容詞:0.998871
喜び:よるこび:名詞:0.998861
才知:さいち:名詞:0.998771
徳:とく:名詞:0.998745

図 3.4: 極性辞書(ポジティブ)の例 [12]

称美:しょうび:名詞:0.996923
激賞:げきしょう:名詞:0.996923
熱讀:ねっさん:名詞:0.996916
一品:いっぴん:名詞:0.996884
怜:れい:名詞:0.996876
愉悦:ゆえつ:名詞:0.996868
奮い立つ:ふるいたつ:動詞:0.996844
速攻:そっこう:名詞:0.996844
百鍊:ひやくれん:名詞:0.996839
淵:えん:名詞:0.996838
好個:こうこ:名詞:0.996817
似つかわしい:につかわしい:形容詞:0.996801
喜悅:きえつ:名詞:0.996789
欣然:きんぜん:副詞:0.996789
欣欣然:きんきんぜん:名詞:0.996789

図 3.5: 極性辞書(ネガティブ)の例 [12]

第4章 実装

今章では前章で述べた設計の具体的な実装について述べる。

4.1 概要

本手法で必要となる実装物の要件は3つあり、「ツイートの収集・保存」、「都道府県の判定」、「テキスト分析によるネガポジの判定」であり、以下 `Twitter Scraper`, `Reverse Geocoder`, `Analyzer` とする。なお、名産物特定とそれに付随する `WordNet` の構築は国立研究開発法人情報通信研究機構 (NICT) の提供する『日本語 `WordNet`』を用いることとする。

4.2 実装環境

表 4.1: 実装環境

OS	使用言語
MacOS.ver10.11.6	Python3

4.3 Tweet Scraper

実験に際して大量の位置情報が付与されたツイートデータが必要になる。`Tweet Scraper` はツイートの収集と保存が主な要件である。ツイートの保存は `csv` 形式でファイルに出力するものとし、収集には `Twitter` の `Streaming API` を用いるものとする。

ソースコード 4.1: `Tweet Scraper`

```
1 # coding:utf-8
2 import os
3 import sys
4 import time
5 import json
6 import tweepy
7 import csv
8
9 consumer_key = "コンシューマーキー"
10 consumer_secret = "コンシューマーシークレット"
```

```

11 access_token      = "アクセストークン"
12 access_token_secret = "アクセストークンシークレット"
13
14 row = "__id__: {}, __timestamp__: {}, __text__: {}, __coordinates__: {}" #出力形式を指定
15
16 class StdOutListener(tweepy.StreamListener):
17     def on_data(self, data):
18         try:
19             tweet = json.loads(data) #jsonを辞書形式で取得
20             id = tweet["id"]
21             timestamp = time.time() #時間の取得
22             text = tweet["text"].encode("utf-8").replace("\n", "\u") #テキストの取得
23             coordinates = tweet["coordinates"] #位置情報の取得
24             if tweet["user"]["lang"] == "ja": #日本語にツイートを限定
25                 if coordinates is not None:
26                     with open("twitter.csv", "a") as f:
27                         writer = csv.writer(f, lineterminator="\n")
28                         writer.writerow([id, timestamp, text, coordinates["coordinates"]]) #csvに出力
29         except:
30             pass
31
32         return True
33
34     def on_error(self, status):
35         # disable error handler
36         # print status
37         pass
38
39
40 if __name__ == "__main__":
41     listner = StdOutListener()
42     auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
43     auth.set_access_token(access_token, access_token_secret)
44     stream = tweepy.Stream(auth, listner)
45     stream.sample()

```

4.4 Reverse Geocoder

集めたツイートデータを都道府県別に分類する必要がある。そのため Reverse Geocoder ではツイートデータに付与された位置情報 (緯度, 経度) を入力として受け取り, ツイートごとの都道府県の名称を判定する。一般に各種情報に対して位置情報 (緯度, 経度) を付与することを『ジオコーディング』と呼ぶが, 本実装は逆に位置情報 (緯度, 経度) から情報 (都道府県) を得るため, 『リバースジオコーディング』と呼ぶ。位置情報と都道府県の対応には Yahoo! の提供するリバースジオコード API を用いた。

パラメータ	値	説明
appid (必須)	string	アプリケーションID。詳細は こちら をご覧ください。
lat (必須)	-	緯度。
lon (必須)	-	経度。
datum	string	緯度経度の測地系： <ul style="list-style-type: none"> • <i>wgs</i> - 世界測地系 (デフォルト) • <i>tky</i> - 日本測地系
output	string	YDFの出力形式： <ul style="list-style-type: none"> • <i>xml</i> - XML形式 (デフォルト) • <i>json</i> - JSON (JSONP) 形式
callback	string	JSON (JSONP) として出力する際のコールバック関数名。

図 4.1: Yahoo!リバーズジオコーダ API のリクエストパラメーター一覧 [20]

ソースコード 4.2: Reverse Geocoder

```

1 # coding:utf-8
2 import urllib.request
3 import csv
4 import re
5 import json
6 import sys
7 import codecs
8
9 app_id = "アプリケーションID"
10 base_url = "https://map.yahooapis.jp/geoapi/V1/reverseGeoCoder" #リクエストURL
11 output = "json" #出力データをjsonに設定
12 lon = "0" #位置情報の初期化
13 lat = "0" #位置情報の初期化
14 url = base_url + "?lat=" + lat + "&lon=" + lon + '&output=' + output + "&appid=" +
15     app_id
16 with open('ツイートデータファイル', "r") as fp:
17     reader = csv.reader(fp)
18     lst1 = list(reader) #ツイートデータをリストで取得
19
20     with open('出力ファイル', "w") as fp2:
21         i = 0
22         lst2 = []
23         for i in range(0, 10000):
24             print(i)
25             lon2 = re.search('[0-9]+\.[0-9]+', lst1[i][1]) #ツイートデータからlonを抽出
26             lat2 = re.search('([0-9]+\.[0-9]+)$', lst1[i][1]) #ツイートデータからlatを抽出
27
28             try:
29                 lon2_num = lon2.group(0) #取得した経度
30                 lat2_num = lat2.group(0) #取得した緯度
31                 url = base_url + "?lat=" + str(lat2_num) + "&lon=" + str(lon2_num) + '&
32                     output=' + output + "&appid=" + app_id
33                 res = urllib.request.urlopen(url).read()
34                 json_data = json.loads(res.decode('utf-8'))
35
36                 try:
37                     pref = json_data['Feature'][0]['Property']['AddressElement'][0]['

```

```

        Name'] #県名を取
        得
37         lst1[i].append(pref)
38         fp2.write(str(lst1[i]) + "\n")
39     except:
40         lst1[i].append("not_found")
41         fp2.write(str(lst1[i]) + "\n")
42
43     except:
44         lst1[i].append("not_found")
45         fp2.write(str(lst1[i]) + "\n")

```

4.5 Analyzer

Analyzer はツイートのテキストを入力に受け取り、形態素解析・極性分析を行い、得られた数値をツイートごとに相加平均を取って都道府県の名産物に対して評価指標を付与する。形態素解析には MeCab[14] を用い、極性分析では日本語の極性辞書として東北大学の乾・岡崎研究室が提供している日本語評価極性辞書 [11] を用いるものとする。

ソースコード 4.3: Analyzer

```

1 #coding: utf-8
2 import re
3 import csv
4 import time
5 import MeCab
6 import random
7
8 with open('ツイートデータ', 'r') as fp:
9     reader = csv.reader(fp)
10    tw_df = list(reader) #ツイートデータを取得
11
12 with open('極性辞書データ', 'r') as fp2:
13    reader2 = csv.reader(fp2)
14    pn_df = list(reader2) #辞書から極性語を取得
15
16 i=0
17 tw_tx = []
18 m = MeCab.Tagger('')
19 for i in range(0, ツイート数):
20    tw_tx.append(tw_df[i][0]) #ツイートデータのテキストを取得
21
22 def get_diclist(text): #引数のテキストに形態素解析を行い、結果を辞書で返す関数
23    parsed = m.parse(text)
24    lines = parsed.split('\n')
25    lines = lines[0:-2]
26    diclist = []
27    for word in lines:
28        l = re.split('\t|,', word)
29        d = {'Surface':l[0], 'POS1':l[1], 'POS2':l[2], 'BaseForm':l[7]}
30        diclist.append(d)
31    return(diclist)
32
33 word_list=[]
34 pn_list=[]
35 i=0
36 for i in range(0, 極性辞書データ数):
37    word_list.append(pn_df[i][0])

```

```

38 i=0
39 for i in range(0, 極性辞書データ数):
40     pn_list.append(pn_df[i][3])
41 pn_dict = dict(zip(word_list, pn_list)) #極性辞書の極性語と対応する極性値を辞書で作成
42
43 def add_pnvalue(diclist_old): #形態素解析により得られた言葉の基本形を極性辞書からサーチする関数
44     diclist_new=[]
45     for word in diclist_old:
46         base = word['BaseForm']
47         if base in pn_dict:
48             pn = float(pn_dict[base])
49         else:
50             pn = 'notfound'
51         word['PN'] = pn
52         diclist_new.append(word)
53     return(diclist_new)
54
55 def get_pnmean(diclist): #ツイートごとに極性値の相加平均をとる関数
56     pn_list = []
57     for word in diclist:
58         pn = word['PN']
59         if pn != 'notfound':
60             pn_list.append(pn)
61     if len(pn_list) > 0:
62         pnmean = float(sum(pn_list)/len(pn_list))
63     else:
64         pnmean = 0
65     return(pnmean)
66
67 i=0
68 pnmeans_list = []
69 with open('出力ファイル', 'w') as fp3:
70     for i in range(0, データ数):
71         dl_old = get_diclist(tw_tx[i])
72         dl_new = add_pnvalue(dl_old)
73         pnmean = get_pnmean(dl_new)
74         pnmeans_list.append(pnmean)
75     j=0
76     for j in range(0, データ数):
77         fp3.write(str(pnmeans_list[j]) + ":" + str(tw_tx[j]) + "\n")

```

4.6 用いる WordNet

本節では実験に実際に用いた WordNet を記載する。なお、以下に記載する WordNet は全て国立研究開発法人情報通信研究機構 (NICT) の提供する『日本語 WordNet』によるものである。

表 4.2: WordNet に含む語一覧 1

言葉	WordNet に含む語
さつまいも	サツマイモ, 薩摩芋
明太子	めんたいこ, メンタイコ
さつまあげ	さつま揚げ, さつまあげ

表 4.3: WordNet に含む語一覧 2

言葉	WordNet に含む語
メロン	メロン, 瓜
ジャガイモ	じゃが, 馬鈴薯, ジャガ芋, じゃが芋, ジャガタラ芋, etc
米	供米, 舍利, 白米, 神米, 米, 八木, 禾穀, ご飯, 米穀, 稲孫, 飯米, etc
りんご	リンゴ, 林檎, りんご
にんにく	ニンニク, 蒜, ガーリック, 葫
さくらんぼ	桜んぼう, 桜ん坊, 桜桃, 桜んぼ, サクランボ, サクランボウ, チェリー
わさび	わさび
牛肉	牛肉, 牛, ビーフ
そば	ソバ, 蕎麦
牛タン	牛タン, 牛舌
ずんだ餅	ずんだ
きりたんぼ	キリタンポ
鶏肉	若鳥, チキン, ニワトリ, 食鶏, 鶏, 若鶏, 鳥肉
枝豆	えだまめ
桃	桃, ピーチ, モモ
ラーメン	らーめん
栗	栗, くり, クリ
あんこう	鮫鯨
いちご	苺, ストロベリー, イチゴ
餃子	ぎょうざ, ギョウザ
こんにゃく	蒟蒻, コンニャク
梅	プラム, 梅, 李, 酸桃
ネギ	葱, ねぎ
お茶	茶, お茶, 茶の葉, 茶葉, 御茶
小松菜	こまつな, コマツナ
煎餅	せんべい, センベイ
落花生	ピーナッツ, ピーナット, ラッカセイ, ナンキンマメ, 落花生
梨	セイヨウナシ, ナシ, ありの実, 洋梨, 梨の実, 梨, ペア, ペヤ
スイカ	すいか, 西瓜
豚肉	ポーク, 豚, 豚肉
かぼす	カボス, 香母酢
マンゴー	マンゴー
いよかん	いよかん
キウイ	キウイ, キウイフルーツ, キーウィ
れんこん	蓮根, レンコン

表 4.4: WordNet に含む語一覧 3

言葉	WordNet に含む語
マグロ	鮪, マグロ
みかん	紅ミカン, 紅みかん, ベニミカン, ミカン, 紅蜜柑, 柑子, オレンジ
里芋	里芋, タロイモ, さといも
鮭	サケ, 鮭, サモン, シャケ, サーモン, 秋味, 年魚
イカ	いか, 烏賊
鱒	ます, マス
うどん	うどん, 饅頭
さざえ	サザエ
ふぐ	河豚, フグ
かぶら	かぶら
カニ	カニ, 蟹, かに
豆腐	とうふ, トウフ
ぶどう	グレープ, 葡萄, ブドウ
すもも	すもも
アユ	あゆ, 鮎
柿	カキ, 柿
うなぎ	うなぎ, ウナギ, 鰻
えび	鰨, シュリンプ, 海老, 蛸, 蝦
はまぐり	ハマグリ, 蛤
さとうきび	サトウキビ
なら漬け	なら漬け
お好み焼き	お好み焼き
たこ焼き	たこ焼き, たこやき
茄子	茄子, ナス, なす
タコ	蛸, たこ
どじょう	どじょう
牡蠣	牡蠣, カキ
マスカット	マスカット
松茸	まつたけ, マツタケ
クワイ	クワイ
カステラ	カステラ, スポンジケーキ
すだち	すだち
びわ	ビワ
ゴーヤ	ゴーヤ

第5章 評価

本章では3章に見た設計と4章に見た実装を用いた実験の評価結果について見る。なお、p/nとは positive/negative で PN 値のことであり、小数点第6位を四捨五入した値を記載する。N/Aは該当するツイートデータが得られなかったものとする。

5.1 評価結果

表 5.1: 評価表 1

都道府県	名産 1	p/n	名産 2	p/n	名産 3	p/n
北海道	メロン	-0.07173	じゃがいも	N/A	米	-0.50303
青森県	りんご	N/A	にんにく	N/A	さくらんぼ	N/A
岩手県	わさび	N/A	牛肉	-0.32323	わんこそば	N/A
宮城県	牛タン	N/A	米	-0.32784	ずんだ餅	-0.50223
秋田県	きりたんぼ	N/A	鶏肉	N/A	米	N/A
山形県	さくらんぼ	N/A	牛肉	N/A	枝豆	N/A
福島県	桃	N/A	ラーメン	-0.44614	りんご	N/A
茨城県	栗	-0.345767	あんこう	N/A	メロン	N/A
栃木県	いちご	N/A	餃子	-0.5109	ラーメン	-0.36509
群馬県	こんにゃく	N/A	梅	-0.66471	ネギ	N/A
埼玉県	お茶	0	小松菜	N/A	煎餅	N/A
千葉県	落花生	N/A	梨	N/A	スイカ	N/A
東京都	豚肉	-0.64511	雷おこし	N/A	東京ばな奈	N/A
神奈川県	まぐろ	-0.37989	みかん	N/A	お茶	-0.58886
新潟県	米	-0.56708	里芋	N/A	鮭	N/A
富山県	いか	N/A	鱒	-0.53098	うどん	N/A
石川県	さざえ	N/A	ふぐ	N/A	かぶら	N/A
福井県	蟹	-0.41957	豆腐	N/A	ふぐ	N/A
山梨県	ぶどう	N/A	もも	N/A	すもも	N/A
長野県	そば	N/A	りんご	N/A	西洋梨	N/A

表 5.2: 評価表 2

都道府県	名産 1	p/n	名産 2	p/n	名産 3	p/n
岐阜県	鮎	N/A	牛肉	N/A	柿	N/A
静岡県	お茶	-0.59312	わさび	N/A	うなぎ	N/A
愛知県	鶏肉	-0.57287	きしめん	-0.89626	ひつまぶし	N/A
三重県	えび	N/A	牛肉	N/A	貝	N/A
滋賀県	牛肉	N/A	鮎	N/A	湖魚	N/A
京都府	お茶	-0.59138	京野菜	-0.49056	蟹	N/A
奈良県	なら漬け	N/A	柿	N/A	鮎	N/A
大阪府	お好み焼き	-0.53887	たこ焼き	-0.54326	茄子	N/A
兵庫県	蟹	-0.53128	牛肉	-0.57122	タコ	N/A
和歌山県	みかん	N/A	柿	N/A	梅	N/A
鳥取県	梨	N/A	蟹	N/A	スイカ	N/A
島根県	はまぐり	N/A	どじょう	N/A	牡蠣	N/A
岡山県	オリーブ	N/A	マスカット	N/A	松茸	N/A
広島県	牡蠣	N/A	クワイ	N/A	オレンジ	N/A
山口県	ふぐ	-0.2671	みかん	N/A	ウニ	N/A
徳島県	すだち	N/A	さつまいも	N/A	鶏肉	N/A
香川県	うどん	N/A	オリーブ	N/A	牡蠣	N/A
愛媛県	みかん	-0.6741	いよかん	N/A	キウイ	N/A
高知県	ゆず	N/A	カツオ	N/A	サバ	N/A
福岡県	明太子	N/A	いちご	N/A	ラーメン	-0.53443
佐賀県	牡蠣	N/A	牛肉	N/A	蟹	N/A
長崎県	カステラ	N/A	びわ	N/A	ちゃんぽん	-0.74047
熊本県	れんこん	N/A	スイカ	N/A	みかん	N/A
大分県	ぶどう	N/A	ふぐ	N/A	かぼす	N/A
宮崎県	マンゴー	N/A	牛肉	N/A	鮎	N/A
鹿児島県	さつまあげ	N/A	さとうきび	N/A	鶏肉	N/A
沖縄県	ゴーヤ	N/A	さとうきび	N/A	沖縄そば	N/A

5.2 考察

WordNet に入っている言葉を含むツイートを抽出すると、分析に使えるデータは大幅に少なくなってしまう。そのため分析するデータがそもそも存在しないという名産品が多く存在してしまった。特に中国地方、四国地方、九州地方に関してはデータが非常に少なく、名産品に関するツイート以前に位置情報付きのツイートデータが不足していた。

また、多くの都道府県において共通して、餃子やたこ焼きやきしめんなどの地域の特色を強く反映しているものは位置情報つきツイートが多く存在していた。これは旅行者などがその地域を訪れたことをアピールしやすいためであると考えられる。また、ラーメンに関するツイートデータも他の食材・食品に比べて多く存在していた。

名産品の WordNet は名産品の対象を広く取って完全一致のキーワードマッチを緩和し、分析に用いることのできるツイートデータを増やすという当初の目的を果たすことはできていた。しかし、東京ばな奈や雷おこしなどの固有の商品名になるとほとんど機能しないという問題も確認された。比較的多くの言葉を WordNet に含む「米」や、一文字である「牛」などは他に比べて対象データを多く取ることができた。今回は独立行政法人通信研究機構 (NICT) の WordNet を用いたが、ツイートデータには食品名 (肉, 魚) よりも、料理名 (ステーキ, 刺身) などが含まれていることが多く、WordNet を完全に活かすことができなかつたといえる。そのためツイートデータを最大限に活用するために当該食品の料理名なども手広く含んだ独自の WordNet の必要性が感じられた。

第6章 結論

本章では本研究のまとめと今後の課題について述べる。

6.1 まとめ

本研究は地方が抱える名産品の知名度の格差という問題とソーシャルネットワークサービスが持つ情報の即時性や位置情報を付与できるという点に着目し、地方名産品の評価は当該地方に住んでいたり、実際にその地方へ行った人の名産品に関する SNS への投稿内容から調べることができるのではないかとこの動機により行われた。ソーシャルネットワークサービスの登場とその普及により、スマートフォンやパーソナルコンピュータから投稿を行うことが簡単になり、人の心情や考えをテキストから知ることがかつてより容易になりつつある。さらに Twitter をはじめとするマイクロブログの位置情報付与機能は GPS を搭載するスマートフォンとの親和性が非常に高く、空間的制約を取り払ったといえる。このため、位置情報を用いた研究やサービスが興隆し、その利用価値は増大している。本研究では特にマイクロブログのテキストとそれに付与される位置情報に着目し、投稿内容に対して形態素解析と極性分析を行うことにより、地方名産品の評価指標を得ることを試みた。その結果、全てではないが、幾つかの名産品に対して指標を得ることができた。現在、日本国の地方活性化は急務とされる課題であり、政府・地方自治体問わず様々な政策を実施しているが未だ解決には至っていない。地方名産品は都会に集中しすぎた人口を一時的に観光や旅行で呼び戻す一つの手段であるとも言える。しかしそれを周知させる方法は地方自治体の宣伝効果に委ねられており、消費者側が能動的に評価する方法を持ち得ていない。地方の名産品を知ることの際して、消費者が受動的な宣伝広告に頼らず評価する方法の一つとして本研究の手法を提案した。

6.2 今後の課題と展望

本研究では Twitter の投稿内容の極性分析を行った。実験に際してはデータセットの大きさが充分でなかったために、名産物の粒度を都道府県の単位で行った。これは当初私が考えていた市町村単位ではそれぞれの充分なデータが抽出できなかったからである。しかし、都道府県単位でもデータ量の格差があり、すべての名産品について分析を行うことはできなかった。今後の課題は更にデータセットの量を増やし細かい粒度での名産物に対する評価指標を得ることである。

謝辞

本論文の作成ならびに研究活動のご指導を頂きました慶應義塾大学徳田・村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・武田合同研究プロジェクトの皆様方に感謝を申し上げます。

また普段から私をご指導してくださった ARCH グループの先輩方, 刺激を与えてくれた ARCH グループの同期生皆様, 私と同じ ARCH グループでないにも関わらずご指導してくださった大戸浩司氏には多大なる感謝をしております。皆様との出会いを与えてくださった慶應義塾大学という場にも感謝しています。

私を地元石川から送り出し, 関東の私立大学に通わせ卒業まで様々な面で支援してくださった両親と祖父には本当に感謝しています, ありがとうございます。以上を以て謝辞とさせていただきます。本当にありがとうございました。

参考文献

- [1] 農林水産省大臣官房企画評価課知的財産戦略チーム：農林水産者・地域食品における地域ブランド化の先進的取組事例集 農林水産省, 2007
- [2] 濱田恵三：地域ブランドによる観光まちづくりの一考察 流通科学大学論集, 流通・経営編, 第22巻第2号, 2010
- [3] 荒牧英治, 増川佐知子, 森田瑞樹：Twitter Catches the Flu: 事実性判定を用いたインフルエンザ流行予測, 情報処理学会研究報告, 2009.
- [4] Johan Bollen, Huina Mao, Xiao-Jun Zeng：Twitter mood predicts the stock market, 2010.
- [5] Siming Chen, SHUai Chen, Zhenhuang Wang Yadong Wu, Jie Liang, Xiaoru Yuan, Nan Cao D-Map: Visual Analysis of Ego-centric Information Diffusion Patterns in Social Media
- [6] Joo Hong Lee, Jae Min Kim, Yong Suk Choi：SNS data Visualization for analyzing spatial-temporal distribution of social anxiety
- [7] T. Sakaki, M. Okazaki, and Y. Matsuo.：Earthquake shakes Twitter users: real-time event detection by social sensors. In Proceedings of the 19th international conference on World wide web, pages 851860. ACM, 2010.
- [8] Pokemon Go の遊び方 <http://www.pokemongo.jp/howto/play/>
- [9] Twimp <http://twimp.nu/about.php>
- [10] DeepAge [https://deepage.net/bigdata/machine learning/2016/09/02/word2vec power of word vector.html](https://deepage.net/bigdata/machine%20learning/2016/09/02/word2vec%20power%20of%20word%20vector.html)
- [11] 日本語評価極性辞書 <http://www.cl.ecei.tohoku.ac.jp/>
- [12] 中国四国農政局 http://www.maff.go.jp/chushi/nousei/kochi/gaiyo/g_nogyo07.html
- [13] 宇都宮餃子マップ <http://www.utsunomiya-cvb.org/pamphlet>
- [14] Mecab <http://taku910.github.io/mecab/>
- [15] 那須野薫, 奥山晶二郎, 中西鏡子, 松尾豊：Twitter における候補者の選挙地盤に注目した国政選挙の当事者予測 情報処理学会 vol.56 No.10 ,2044-2053(Oct,2015)
- [16] 日本語 WordNet 国立研究開発法人情報通信研究機構 (NICT)

参考文献

- [17] 平尾拓也, 鈴木孝彦, 宮田光樹, 廣川左千男 : 意味辞書 WordNet の不整合抽出, 情報処理学会研究報告, 2012
- [18] Cabocha, <http://taku910.github.io/cabocha/>
- [19] JUMAN, 京都大学黒橋・河原研究室, <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>
- [20] Yahoo!リバーズジオコード API, <https://developer.yahoo.co.jp/webapi/map/openlocalplatform/v1/reversegeocoder.html>