

# 地方のための Twitter ローカライズ

## Twitter Localization

### 地方のための Twitter ローカライズ Twitter Localization

1013039 丸山大仁 Hirohito Maruyama

## 1 背景と目的

近年、Web 上でのコミュニケーション手段として、すでに普及が進んでいる電子掲示板システム (BBS) やブログ (Weblog) に次いで、利用者が急増しているものがソーシャルネットワークサービス (SNS) である [1]. SNS は、人と人とのつながりを促進・サポートする、コミュニティ型の Web サイトである [2]. さらに、SNS の利用者数は多く、そのなかでも、2006 年に米国でリリースされた Twitter は、2008 年 4 月 23 日には日本語サイト「Twitter Japan」のサービスも開始され [3], 2014 年 12 月で、LINE や Facebook について 3 番目に利用者数が多い [4]. また、Twitter は世界中で 2 億 8400 万人のユーザーが利用し、国内でも 2000 万人近くが利用しているサービスである [5]. ユーザーは 140 文字までの短文から構成されるツイートと呼ばれる投稿ができる. 現在の Twitter はモバイルベースのユーザー率が全体の 80% [5] を占め、利用者は各地を移動しながらいつでも Twitter を利用することができる. そのため Twitter では位置情報を利用してツイートを投稿したり、情報を検索したりすることが可能である.

Twitter における Facebook など他の SNS と異なる特徴として、以下の 2 点を挙げる事ができる. 1 点目は、リアルタイム性が高いことである. 投稿件数は、1 秒あたり 5700 件あり、トレンド分析や口コミ分析に利用されることがある [6]. 2 点目は、10 代から 20 代の若い世代の利用者が非常に多いことである.

本プロジェクト内で、Twitter を用いる場合に感じる不満な部分を話し合った際に、「アプリメーカー」や「ツイートプロファイリング」のような連携サービスを用いた分析では不明瞭な点が多いということや、Twitter で提供される検索サービスを用いて情報を入手しようとした際に他の情報検索システムを利用しなければならず手間がかかるということが挙げられた. 話し合いで挙げられた不満点をまとめると以下の 2 つに分類された.

- 情報検索の手間が多く不満
  - 1 つのアプリケーションに簡潔化できないか
- 既存の性格診断アプリに対する不満
  - 既存のものにない、独自の性格分析を実現できないか

- より多くのツイートを用いて特徴を得ることができないか

本プロジェクトではこのような部分の改善に関する提案・調査・意見交換を行い、「Twitter の特徴を最大限に活かした新しいサービスを提供」を目的として活動することとした. また、本プロジェクトは上記の 2 点の問題に対する改善を行うために以下の 2 つの班にわかれプロジェクト活動を進めてきた.

- 分析班
  - ツイートから性格を推測し、函館の観光地をお勧めする Web アプリケーション
- 検索班
  - カテゴリから任意の場所の飲食店に関するツイートを検索し地図に表示する Web アプリケーション

結成後、両班ともに前期には意見交換・調査を重点的に行い、Web サイト設計に必要な知識習得と、その得た知識を使った Web アプリケーションの実装を行った. 後期には、実装されたアプリケーションに対してレビューを行い、そこででてきた改善点に対する実装を繰り返し行った.

## 2 課題設定のプロセスと課題設定

本プロジェクトでは、プロジェクトが始まると同時に教員から以下の 2 点のアドバイスをいただいた.

- 実現性は小さい物でよく、まずは案を考える
- 案が決定した段階で、次に実現性が高いもので何の勉強が必要になるかを考える

このアドバイスをもとに、前期ではどのようなものを作成するかを決定するため、メンバー全員で Twitter をどのように利用して、どのような利用のされ方があるのかについて意見交換・調査を行った. 調査と意見交換において、論文や書籍、Web からの情報収集をメンバー 1 人あたり数個提示して、収集した内容を独自のサービスを考える材料とした. 本プロジェクトメンバーがもつ、Twitter に対する不満点についてブレインストーミングを行い、そこで挙げられた不満点に対して、先行事例で解決されているか

調査をした。しかし、ブレインストーミングや調査だけでは、具体的な案はメンバー内で明確に提示されることがなかった。そこで、視点を変え、Twitter から離れ自分たちがどんなサービスを望み、どんなことをしたいのかということに着目し、もう一度メンバー内でブレインストーミングや調査、意見交換を徹底した。さらに、自分たちの行いたいサービスをリストアップした案を、KJ 法をもとにグループ化した。グループ化された案が

- 機械学習を使ったサービス
- 位置情報を使ったサービス

に執着した。上記のサービスごとに 2 つのグループに分かれて、上のサービスから順に分析班・検索班と名付け作業を進めた。両班に活動を行ってもらった際にいくつか注意したこととして、情報共有を効率的に進めるために、Google Drive や LINE などのサービスを使って徹底的に意見交換し、調査したものを情報共有した。また、プロジェクトの時間構成やその日事の活動内容の確認するため、調査や開発に関する不明慮な点や理解したことを共有するためにレジュメを使った。そして、プロジェクト時間中にメンバーの議事録の記入を徹底させた。これは過去において活動した内容を確認するためであったり、プロジェクト活動中に誰がどのタスクを行っているかを確認できるように行った。両班には、具体的な提供したいサービスを考える際に、具体的な対象を絞るために「7W2H」を使った。さらに、提供したいサービスに対する仕様書の作成を行った。「7W2H」が必要になった経緯として、サービスを展開していくうえで提供したい情報の分類を明確化したり、メンバー内の考えを 1 つのものに集約できると考えた結果「7W2H」を実践した。次に仕様書では、仕様書の知識習得を行いメンバー内で共有をし、仕様書に書くべき必要な項目を決定した。また、プロジェクト全体の仕様書と各班の分析班の仕様書、検索班の仕様書の計 3 部の仕様書を作成した。プロジェクト全体の仕様書には、プロジェクトの概要とプロジェクトの背景や目的、プロジェクトの基本方針、達成目標、対象、開発環境に関するものを記述した。また分析班・検索班の仕様書では、問題提起をしたうえで問題提起に対する解決案、今後の展望を含めた現状で構想している画面遷移図の内容を含めたものを記述した。また仕様書と同時進行で、提供したいサービスの構成や構成の役割分担が具体的に可視化できるようになるのが必要だと考え、「WBS 図」を作成した。

仕様書や「WBS 図」をもとに、分析班では Twitter を使って機械学習を活かした性格分析に焦点を絞り、検索班では Twitter を使って位置情報を活かした飲食店検索に焦点を絞り両班ともに活動を進めた。

## 2.1 分析班のシステム

### 2.1.1 システム作成

ツイートプロファイリングのような、既存の Twitter 連携アプリによって得られる出力結果は正確に解析されていないため、そのユーザーがもつ傾向や特徴を調べることは困難である。ここで分析班で議論するなかで以下の意見や問題点が挙げられた。

- ツイートプロファイリングでは、ツイートの分析にはユーザーの過去 500 件のツイートを利用し出力結果は画像つきツイートとして投稿することが可能であるが、こうして得られた出力結果は制度が低い。
- アプリメーカーによって作成された Twitter 解析アプリは、表示結果がランダムである場合や毎回異なることがあり、利用するユーザーに正確な解析結果を提供することが不可能である。

そこで分析班ではこれらの問題点を改善することにより、ユーザーのツイートをより正確に解析して性格といった特徴を抽出し、そのユーザーのプロフィール帳を作成する Twitter 連携アプリを開発することになった。

### 2.1.2 システム実装

システムを実装していくうえで、まず初めに Web アプリケーションの画面に埋め込むものとして以下の 4 つを決定した。

- 性格診断を行ったアカウント名
- 結果画面に折れ線グラフで性格診断の結果を表示
- その性格についての説明
- 共有ツイートできるボタン

また、Web ページの全体構想として以下の 4 つを作成することを決定した。

- ホーム画面
- Twitter アカウントの認証・ログイン画面
- 性格タイプ診断選択画面
- 診断結果画面

以上の内容を実装することを目標として分析班の活動を進めた。

性格の診断手順としては、まず

1. ユーザーは分析班が作成した Web ページのホーム画面にアクセスを行い、診断ボタンを押すと、自分の Twitter アカウントを認証・ログインを行う画面にアクセスされる。
2. ユーザーが認証・ログインを完了すると、診断タイプ選択画面にアクセスされ、現在、未来、集団のどれか 1 つの診断タイプを選択する。
3. 診断タイプを選択すると、TwitterAPI を用いてユーザーの

最新のツイートから 3200 件のツイートを取得し、それを形態素解析にかける。形態素解析されたツイートをエゴグラムの特徴辞書と比較して性格診断を行う。

4. その診断結果とそれによって作られたエゴグラムの折れ線グラフをプロフィール帳形式で Web ページに表示する。また、これらの出力結果をフォロワー間で共有することができる共有ツイート機能もある。

ただし、共有ツイート機能では、分析班が作成した Web ページのアドレスと作成されたプロフィール帳がツイートされるものを想定している。

「現在から未来診断」では、ユーザーのツイート 3200 件から未来の性格を推測するものである。手法としては、まず 3200 件のツイートを TwitterAPI を用いて取得し、時系列順に並べてから 30 等分する。そして、30 等分したものをそれぞれでエゴグラムを作成し、回帰分析を行うことでツイートから未来の性格を推測している。

「集団診断」では、フォロワーのまとまりを考慮して性格を診断していたが、ユーザーと 1 番性格の近いフォロワーを特定して診断するのも面白いという意見があり、これを気の合う人の診断として開発を行った。気の合う人の診断の手順は以下のとおりである。

1. PHP において、TwitterAPI を利用し、相互フォロワーの最新のツイート 100 件ずつ配列に格納する。それらフォロワーのツイートはそれぞれテキストファイルに保存されているので、それらをいっぺんにまとめるようにパスを設定し、Processing に送る。
2. Processing のなかで渡されたツイート群を人数ごとに区切り、点数化を繰り返す。そして、全員分の各要素の点数をまとめて PHP に返す。
3. PHP で返ってきた点数を人数分に区切って配列に格納する。そして、ユーザー自身の各要素の点数との比較を繰り返し行い、気の合う人を算出する。

また後期から追加した内容として「函館の観光地をお勧めしてくれる機能」を追加した。この機能は、ユーザーのツイートを解析し、ランダムフォレストという機械学習のアルゴリズムによってユーザーに函館の観光地を金森赤レンガ倉庫群、函館山、北島三郎記念館、旧函館市公会堂、トラピスチヌ修道院の 5 つの中から選んで出力する機能である。

フロントエンド作成部分では、HTML、CSS、JavaScript を中心に作成を行った。これは「画面遷移図」をもとにレイアウトを作成した。

### 2.1.3 結果

分析班の提供する Web アプリケーションは、以下の

- より多くのツイート数で解析
- 診断手法にエゴグラムを採用
- 現在から未来の性格の推移を推測して診断する、相互フォロワー全体の性格を診断するなど、独自の性格診断を採用
- ローカライズをあらわすものとして、機械学習を利用し函館の観光地をユーザーにお勧めする機能

の機能を実現可能としている。

ホーム画面にある「診断する」のボタンを押すと、アプリケーションと Twitter の連携のためにユーザー自身の TwitterID とパスワードの入力画面に移り、入力を終わると診断タイプの選択画面に移動する。この画面では、通常診断と集団診断が選択でき、通常診断を選択すると、現在の性格診断の結果と現在から未来の推移を表した診断結果の 2 つが表示される。ここでは、現在から未来診断の結果のしたにはスライダーバーがあり、動かすことで現在から未来への性格の移り変わりを 10 段階でみる事ができる。

診断タイプの選択画面で集団診断を選択すると、相互フォロワー全体の性格の診断結果が表示される。この診断では、相互フォロワーとユーザーに気の合う人の TwitterID も表示され、通常診断と集団診断の結果にはユーザーにお勧めの観光地が表示される。そこでは、結果の表を画像として Twitter に共有ツイートすることが可能である。

分析班は、ツイートから性格を推測し、同時に函館の観光地をお勧めする Web アプリケーションを作成することを目的として活動し、現段階でこの目的を達成した。

## 2.2 検索班のシステム

### 2.2.1 システム作成

検索班内で、話し合いを進めていくなかで、Twitter で情報取得をする際、Twitter 検索機能の問題が挙げられた。詳しい情報を検索するためには複数のサービスとの併用が必要となり、Twitter だけでは詳しい情報を得ることが難しいということである。ここで検索班では、グルメに関する情報を検索するということに焦点を当て話を進めたところ、詳細な情報を得るためには、Web での検索、地図アプリ、Twitter の起動など複数のサービスをまたぐ必要があるという問題が上がった。そこで検索班は複数のサービスをまたぐ手間をなくし、1 つのアプリケーションでグルメに関するツイートを得ることはできないかと考えた。

検索班では、任意の場所の飲食店に関する情報を検索し、地図に表示する Web アプリケーションを提案することとなった。

### 2.2.2 システム実装

検索班で具体的な構成を考えるにあたり作成した「WBS 図」「フローチャート」「画面遷移図」の 3 つをもとに、必要な機能や知識を習得し、飲食店に関する情報を入手するための方法としてホットベッパ Web サービスを利用し、ツイート表示をするため

に Twitter 社が提供する API を使用することとした。そして、これらを利用するために必要なプログラミング言語として PHP を使用することとなった。また、フロントエンドを構築するために HTML, CSS, jQuery, JavaScript を使って作業を進めることとなった。検索班では環境構築として、各自に PC に XAMPP を導入した。今後のサービスの拡張を行う可能性を考え、MySQL なども同時に扱う考慮をし、Apache 単体ではなく XAMPP での導入を行った。

作成物の進捗管理をするために Git を用いることにした。そこで、Git の使用方法を理解するため、Git の講習会を学生が中心となって開催した。そのなかで、それぞれの環境に GitGUI を導入し、そこから Git の管理を行っていくことにした。また後期から、作成物に地図を用いること、スマートフォン対応をすることなどの理由から、操作をする際に何度もページ更新をしてしまうとユーザーの使い心地が悪いと考え、ページ更新をせずに済む非同期通信を用いることにした。そこで、Ajax を通してそれぞれ下記の API にアクセスをし、情報を取得することにした。中間発表時の構成では、現在地周辺のみを検索可能とする予定であったが、開発をしていくなかで、利便性を考え任意の場所でも検索可能にすることにした。任意の場所への地図の移動、また店舗の検索をするためにはその地点の緯度、経度が必要である。しかし、ユーザーに直接緯度、経度を入力してもらうのは難しいため、指定したい地名や住所、ランドマークから緯度、経度を取得できる GeocodingAPI を利用した。これを利用することにより、検索時の条件に住所などを入力すると、それを緯度、経度に変換してから他の API にアクセスすることが可能となった。

お店に関する情報を高評価や低評価、その他で表示できるような機能を追加した。ここでは、飲食店に関するポジティブなワード、例えば「うまい」を positive\_array、ネガティブなワード、例えば「まずい」を negative\_array として配列を作りそれを PHP で実装した。

検索班では、地図 Web アプリケーションを提供するので OpenLayers と呼ばれるライブラリを使用した。さらに、お店に関する情報が表示されるピンを表示するために Leaflet と呼ばれるライブラリを使った。

### 2.2.3 結果

検索班の提供する Web アプリケーションはウェブブラウザ上で動作するもので、インターネット上に公開されているウェブサイトへアクセスすることにより利用でき、パソコンとスマートフォンの双方で利用できるようにした。このウェブアプリケーションでは、

1. 任意の場所に関するキーワード（例：函館駅、五稜郭駅）を入力

2. 該当地点を指定し、予め用意されている複数のカテゴリをもとに飲食店を検索
3. 検索結果が地図上にピン表示され、各ピンを選択することにより飲食店の営業時間と予算に関する情報が表示され、ツイート検索を実行

することで検索班のシステムを利用できる。ただし、ツイートは判定に基づき高評価と低評価に分類され、それぞれのツイートのみを抽出して表示できるようになっている。

検索班の Web アプリケーションでは、これらの「飲食店検索」「地図上で表示」「ツイートの検索」という工程を 1 つのウェブサイト上で表示可能にすることを目的として活動し、現段階でこの目的を達成した。

## 3 展望

本プロジェクトの展望として、2月に行われる秋葉原発表会に向けて改善仕様書の作成、作成した段階での開発に直ぐ着手することを進め、教員だけではなくさまざまな人のレビューを受けたいと考えている。改善仕様書は、分析班・検索班に仕様書を作成してもらおう。内容として、それぞれの班の目的や今あるサービスの画面遷移図、改善できる物の画面遷移図、画面の仕様、改善できる物の画面の仕様、追加の部分のスケジューリングなどである。また今後の課題として仕様書の明確化、開発するうえでのスケジュールをより具体的にすること、レビューに対する修正力など至らない部分の改善などが課題である。

## 参考文献

- [1] 大向 一輝, 情報処理, vol.47, no.8, pp.993-1000, September. 2006.
- [2] IT用語辞典, “SNS 【 Social Networking Service 】 ソーシャルネットワーキングサービス”, <http://e-words.jp/w/SNS.html> [Accessed: Dec. 18, 2015].
- [3] “若者における SNS 利用行動およびリスク認知の検討—LINE と Twitter を中心に—”, 荻野正美, 2014.
- [4] nielsen, “スマートフォン利用者の 92% が SNS を利用 ~ ニールセン, SNS の最新利用動向を発表 ~”, [http://www.netratings.co.jp/news\\_release/2015/01/Newsrelease20150127.html](http://www.netratings.co.jp/news_release/2015/01/Newsrelease20150127.html). [Accessed: Jul. 24, 2015].
- [5] Twitter, Inc., “Twitter Reports Third Quarter 2014 Results”, <https://investor.twitterinc.com/releasedetail.cfm?ReleaseID=878170>. [Accessed: Jul. 24, 2015].
- [6] “Twitter を情報源とした発話ロボットシステムの開発 Development of Speech Function for Communication Robot using Twitter”, 藤原 裕樹, 山下 晃弘, 2014.
- [7] Twitter をはじめよう!, “連携ソフトやサービスで Twitter をさらに活用しよう”, <http://www.greenspace.info/twitter/step4.html#services> [Accessed: Dec. 18, 2015].