

SNSの投稿を対象とした 鉄道における 改善点分析システム

学籍番号：1421102 氏名：齊藤 昭平

指導教員：鷹野 孝典 准教授

研究背景

鉄道を利用する人にとっては駅の設備，また環境がしっかりと整備されているかは重要である．

【駅の設備の例】

- エレベータやスロープなどのバリアフリー設備，乗り換え案内標識，トイレ，階段の数，通路の幅など

→ **改善点とともに，評価点がSNS上に投稿されることがある．**

関連研究(1)

- **「Twitterにおける言及関係によるクラスタリングを利用したスパムアカウント判定手法の検討」** (菊池, 吉村, 情報処理学会, 2014)
 - Twitterのユーザの投稿を基に、クラスタ係数や普及数を指標したつながりの強弱を用いて、スパムアカウントの判定を行っている.
- **「Twitterにおけるテレビ番組の実況ツイートの可視化」** (菅野, 竹島, 加納, 東京工科大学, 2014)
 - 「#(番組名)」が付加されたツイートを取得し、番組前・中・後の感情を分析し、それを可視化する.
- **「Twitter上のあるユーザの意外な情報抽出方法の提案」** (大原, 灘本, 甲南大学, 2014)
 - Twitterユーザの内容や属性データからTwitterユーザの近況の変化の情報を抽出する手法の提案を行う.

関連研究(2)

- **「Twitterに基づく都市鉄道の運転見合わせの発生状況の分析」** (堀江, 関谷, 金子, 土木学会論文集, 2015)
 - 首都圏で発生した運転見合わせを原因別に分析し, 鉄道会社のTwitterで発表されている運転再開見込み時間と実際に再開した時間との誤差を調査している.
- **「統計的学習モデルを利用した日本語慣用句の意味的曖昧性解消」** (宮田, 竹内, 岡山大学, 2017)
 - 同じ言葉・慣用句でも意味が複数ある場合において, 統計的学習モデル(KNP・W2V)を用いて, 意味性解消手法を提案している.
- **「Twitter 上で共感を生み出すツイートの性質に関する考察」** (大川, 高間, 人工知能学会, 2012)
 - ツイートに対し多くのユーザが共感するケースに着目し, 発生メカニズムを解明を目指している.

研究課題

- 駅ごとのレビューはあるが、その駅の特定の場所ごとに分類されているレビューは存在しない。
- また、規模の大きな駅になればレビューの数が多くなっていき、どこに改善を求めているのかが分かりにくくなってしまう。

→ 具体的なレビューがわからない

研究動機

- 前述の研究課題のスライドにも述べたように、規模の大きな駅になるとレビューの数が多くなるが、レビューの整理がされていなく、どこがどうなのかが分かりづらい。
- 係り受け分析を行うことで、場所などが明確化できる。

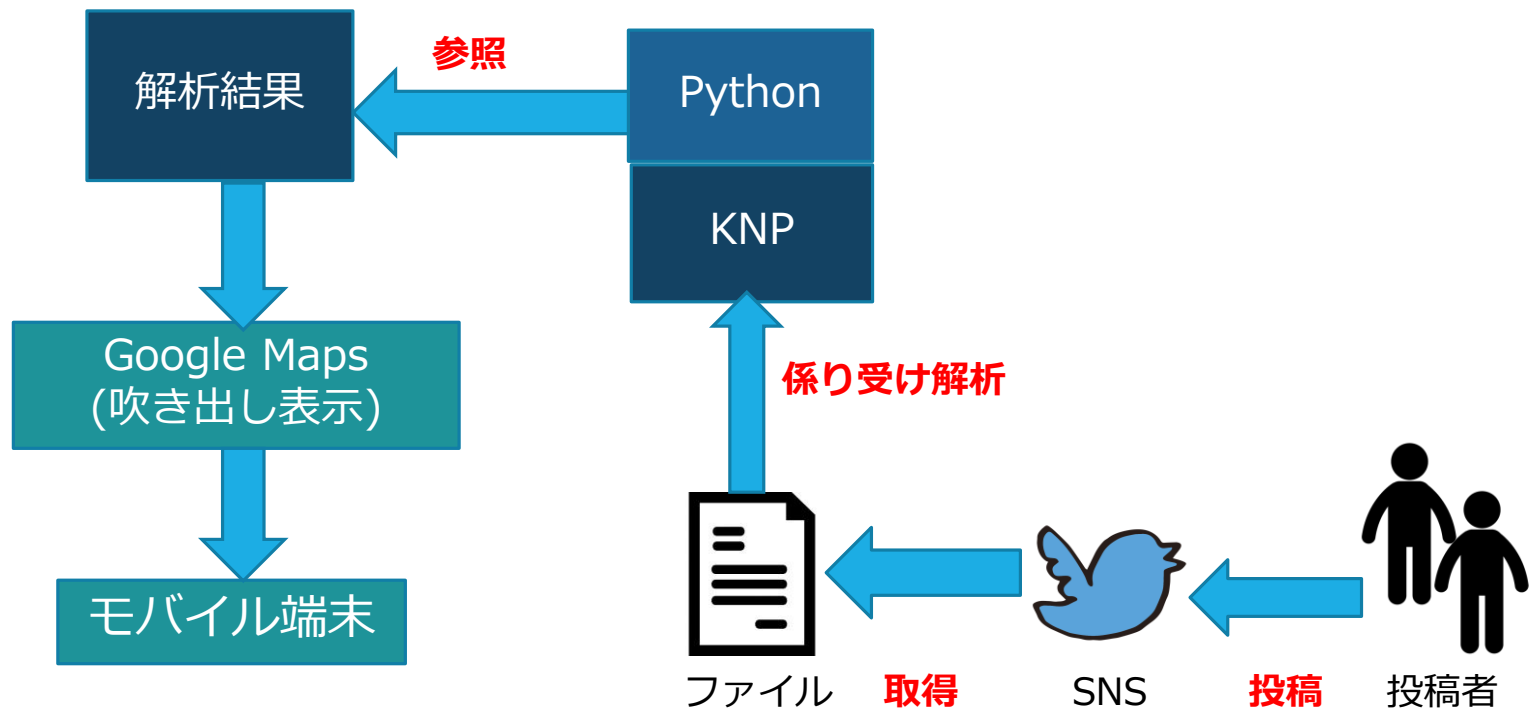
本研究のアプローチ

- 旅行者に向けていきたい.
- 同じ改善点を求めるツイートでも言い回しが異なることがある.
- **取得した投稿を基に、同じ部分での改善点や良い点の投稿数がどれほどあるかを係り受け解析を用いて、分類分けしていく.**
- 不定の推定もしていきたい(特定できなくともある程度把握できればと思っている).



駅の良い点、改善すべき点が駅ごとに把握できるようになる.

提案システム



実装

- 取得した投稿を基に、どの箇所に改善点、また評価点の投稿が集中しているかを**係り受け解析**、**クラスタリング**を用いて分類を行う。
- Google Maps APIを用いてマッピングを行う。

→ ターミナル駅において、どういう部分を改善してほしいというツイート、また良いと思ったところのツイートがどれだけあったかをGoogleマップ上に吹き出しで表示していく。

吹き出しの一例)

〇〇駅の改善点

通路が狭い：30件

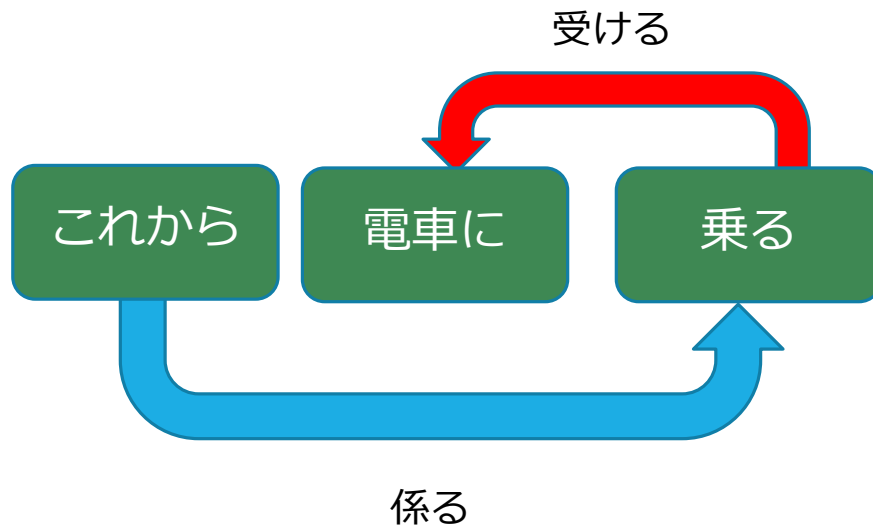
〇〇駅の良い点

案内がわかりやすい

係り受けについて

係り受けとは、2つの文節が意味の上で結びついている状態を指す。

例文「これから電車に乗る」



実装詳細(1/3)

まず、投稿をAPIを用いて投稿を取得する.

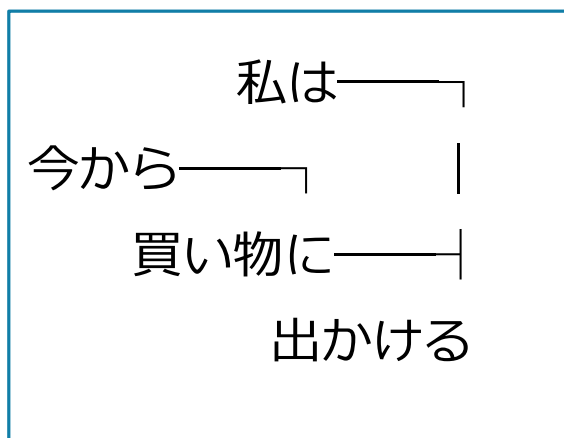
以下、取得した投稿の抜粋を下記に示す.

- 東京駅ラビリンスは、京王線(京葉線)が最下層で難関だと思うの… あと、さすがにどこのロッカーに荷物入れたかは私にもわからない
- 東京駅についた出口がわからないこれは迷子になりそう
- 東京駅全然わからない 乗るところこ

実装詳細(2/3)

係り受け解析ソフト「KNP」を用いて、投稿文を解析する.

例) 「私は今から買い物に出かける。」



出力結果によると、「私は」は「出かける」に係っており、「出かける」は、「買い物に」を受けている.

実装詳細(3/3)

解析結果を基に、吹き出しを表示する。



実験(1/)

「東京駅 わからない」で検索

期間: 10/7 ~ 12/13

取得件数: 554

投稿内容に関しては, スライド11を参照

実験(2/)

例文 「いつも地下への入口がわからないダンジョン東京駅」

何時(いつ)も => 分かる

地下 => 入り口

入り口 => ダンジョン

分かる => ダンジョン

今後の予定

- SNSの投稿を随時取得する。また、今後取得した投稿の内容次第で、単語リストに単語を追加する。
- ファイルをアップロードできるプログラムを作成する(DBに格納した単語が含まれているツイートの件数をカウントさせる)。
- Google maps APIでマッピングをするプログラムを作成する。
- 関連度の算出方法
- 論文の執筆