

低コストデータによるみかんのおいしさ分析

11571123 宮崎 雅大 (灘本研究室)

あらまし：日本人は昔から数多くの果物を食べてきた。その中でも「みかん」は代表的であり、冬の季節には欠かせない果物となっている。しかし、安定しない収入などの問題により後継者が不足している。そこで、本研究では小規模農園に低コストセンサーを設置することで、みかんのおいしさに影響している気象データを取得し分析する。また、シズルワードに着目し Twitter を用いてみかんのおいしさに関するデータを取得し分析する。

1. はじめに

近年、日本の農業界では深刻な後継者不足に悩まされており、その理由として「重労働」や「安定しない収入」が挙げられる。農業試験場では、気象データ、肥料データ、土中データなどの様々なデータの分析を行っている。しかしながら、農業試験場のデータを使用するにはセンサーに莫大なコストがかかる。また、土地の整備があまりされていない農園では農業試験場のデータが使用できない問題点がある。そこで本研究では、低コストで製作できるセンサーを土地の整備があまりされていない農園、特に小規模農園に設置することで独自に気象データを取得し、おいしい果物を作るための気象条件を分析する。また、本研究では日本人にとって親しみ深く、古くから日本で栽培されているみかんを対象とする。

一方、インターネット上には様々なおいしさに関する情報が溢れている。特に、Twitter¹は誰でも気軽に匿名で発信できることから、食べ物に関する情報が多数ある。そこで、ユーザのみかんに対するおいしさの情報を分析する事を行う。本研究では、おいしさの情報として、シズルワード^[1]に着目する。シズルワードに関する様々な研究^{[2][3][4]}があるが、これまで、我々はこのシズルワードに着目し、様々な研究をしてきた^[5]。その結果、シズルワードとある食べ物との関係性が判明している。そこで、本研究では、Twitter上のみかんとシズルワードとの関係の分析も行う。

2. 気象データとみかんの関係抽出

本研究では、実際に神奈川県小田原市の小規模農園にセンサーを設置した。一般的にみかんの生育に大きな影響を与えるといわれている温度・湿度・(気圧)・照度のデータを取得できるセンサーを製作する。以下、設置したセンサーの概要、そして現状について述べる。

2.1 安価センサーの概要

本研究で使用するセンサーについて述べる。温湿度・気圧センサーは AE-BME280(J3)を使用、照度センサーは TSL2561、通信媒体には TWELITE を使用し上記のセンサーから与えられた数値を Raspberry Pi を通じてインターネットを介してデータを転送する。本研究で使用するセンサー製作にかかる費用は1つ当たり約4,000円と安価に作成することができた。回路図を図1に示す。また、受信側の Raspberry Pi に安定的に電力を供給するためにソーラーパネル2枚、カーバッテリー2基を用いる。

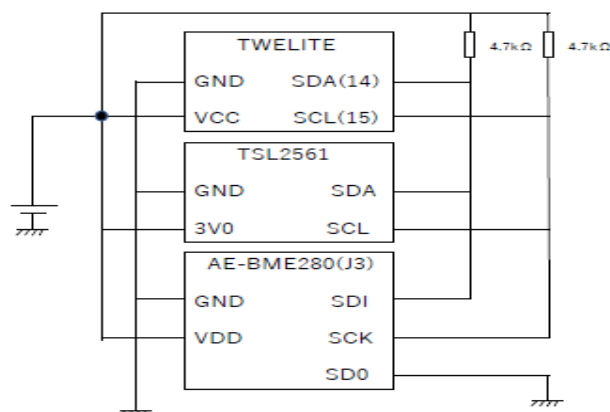


図1 気象センサー回路図

製作したセンサーを神奈川県の小規模農園内の気象が異なる5本の木に設置し数値データを取得する。設置したセンサーの写真を図2に示す。データの取得方法は、図2の小屋に設置した受信機により、TWELITE を用いて各々のセンサーからデータを受信する。5つのセンサーから受信したデータを Raspberry Pi を用いてインターネットを介して研究室の PC に送信する。農園における電源は図2の小屋の上に設置した1畳ほどのソーラーパネルにより供給される。

2.2 現状

大型台風 21 号の影響により受信側の主電力がダウンしたことで気象データ取得が困難になった。その後、復旧を試みたが改善されず、2019

¹Twitter: <https://twitter.com/>

年1月現在データ取得の復旧ができていないのが現状である。



図2 設置したセンサーとソーラーパネル

3. シズルワードとみかんの関係抽出

一般の人々がみかんのおいしさをどのように表現しているかを分析するために、Twitter 上のみかんとシズルワードとの関係を分析する。

3.1 分析手順

分析手順は、まず「みかん」「ミカン」の両方をふくむツイートを Twitter API を用いて収集する。ストップワードに「敬称」や「RT」など食べ物のみかんに関係しないツイートや商用目的のツイートを設定した。次に収集したツイートの中からシズルワードの「味覚系」「食感系」「情報系」の3つの領域毎にシズルワードを含むツイートを抽出する。また、Bot 等の同じ文章を何度も繰り返しツイートしているものに関しては、既出文と比較し類似度がある閾値 α 以上のものを省く。その後、3つの領域毎に特徴を分析、考察を行う。本研究では α を 0.5 とする。

3.2 分析データ

使用するシズルワードは味覚系表現 97 語、食感系表現 130 語、情報系表現 135 語の合計 362 語を用いる。分析対象ツイートは、Twitter API を用いて 2018 年 10 月 1 日から 2019 年 1 月 8 日に「みかん」「ミカン」を検索クエリとし収集した 452,987 ツイートの内、BOT 等を省き、且つシズルワードを含むツイート 9,873 件を用いる。

3.3 分析結果

味覚系表現を含むシズルワード付きツイートの総数は 2,917 件である。その内、上位 5 件は「甘い」が 1,320 件、「サワー」が 291 件、「すっぱい」が 239 件、「ビター」が 84 件、「さっぱり」が 68 件であった。「サワー」「ビター」が含まれる理由としては、「みかんサワー」「ビターオレンジ」といったアルコール飲料のツイートがほとんどであり、本研究の対象としていない加工されたみかんに関するツイートであった。

食感系表現を含むツイートの総数は 2,472 件であり上位 5 件は「ゴロゴロ」が 301 件、「硬い」

が 222 件、「ほっこり」が 199 件、「もちもち」が 101 件、「冷やし」が 79 件である。「ゴロゴロ」「ほっこり」の理由としては「コタツでゴロゴロとみかんを食べる」などのツイートが多数見られみかんを食べる際の状態を表している。一方で「冷やし」というワードは「みかんをキンキンに冷蔵庫で冷やし」等のみかんをおいしく食べる際の方法であり、みかんの直接の味の評価ではなく、望ましくない結果であった。

情報系表現を含むツイートの総数は 8,138 件であり上位 5 件は「かわいい」が 2,182 件、「俺の」が 783 件、「生の」が 565 件、「ナチュラル」が 477 件、「100%」が 400 件である。「かわいい」「俺の」が 2,965 件ありアニメの「みかん」指しているため敬称にかからず抽出された。「ナチュラル」「生の」などは「アンナチュラル」や「中学生の」など表記ゆれが抽出されている部分が多数見受けられた。

4. まとめと今後の課題

本研究では、低コストセンサーを製作し神奈川県小田原市に設置を行った。また、シズルワードに着目して Twitter 上からみかんに関するツイートを取得し、分析を行った。今後の課題として四季によるシズルワード使用率の違い、年毎のシズルワードとみかんの関係性の分析を行うことでよりユーザのみかんに対する好みの推移が表れると考える。

5. 参考文献

- [1] 大橋正房, シズル研究会, “「おいしい」感覚と言葉”, BMFT出版部, 2010 “SIZZLE WORD シズルワードの現在”, BMFT出版部2015
- [2] 高田実央, 佐藤哲司, “レシピタイトルに見られるシズルワードの多義性解消手法の提案”. 一般社団法人. 電子情報通信学会学技報 IEICE Technical Report DE2018-19(2018-12).
- [3] 真辺諒, 長尾哲志, 安藤一秋, “Twitter からの土産に関するレビュー情報の抽出”. 情報処理学会第80回全国大会. 5L-04.
- [4] 笠井信吾, 齊藤史哲, 石津昌平, “共起語ネットワークに基づいたシズルワードの特性分析”, The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018.
- [5] Daisuke Kato, Mai Miyabe, Eiji Aramaki, and Akiyo Nadamoto, “ Sizzle Word Analysis from Multiple InternetMedia. ” Network-Based Information Systems (NBIS) 2015 18th International Conference on, pp. 444-451, 2015.