

Instagram で多くの人に閲覧される投稿の傾向

-自然言語処理を用いた投稿の自動生成-

2019SS070 杉浦 健太

指導教員: 三浦 英俊

1 はじめに

2000年代でのスマートフォンの普及により利用者が急増している SNS(ソーシャルネットワーキングサービス)は、これまで個人間での利用が多かった。例えば思い出のシェアや共通の趣味を持った人との新しい繋がりなどが主な例である。

それが近年、多くの企業がこぞって公式のアカウントを持ち、自社のサービスのアピールや採用のための発信を始めるようになった。数ある SNS の中でも、その傾向は Instagram で顕著に見られる。この変化は、デジタルマーケティングの重要性が増してきたことが主な原因である。

そんな Instagram について企業が懸念するのは「SNS の波に乗り遅れ、将来的に競争に負けてしまうのではないか。」というものである。

本研究はそんな SNS(特に Instagram)において悩める企業の後押しとなるように以下のことを目的としていく。

- おすすめに表示され、多くの人から認知を得ることのできる投稿の傾向を掴む。
- おすすめに表示される投稿の投稿文を自動的に生成し、企業のリソース効率化を後押しする。

2 データの取得

本研究で必要となるのは、参考とする Instagram のアカウントの数値である。ただしこれを 1 つずつ目視してはあまりに膨大な作業手間ある。そこで利用したものが「Instagram API」である。Instagram API を用いて主に下記の情報をスプレッドシートに出力した。具体的には図 1 のようなものである。

指定したアカウントのフォロワー数/フォロワー数/投稿数/直近の投稿画像/直近の投稿のキャプション/直近の投稿の URL/直近の投稿の投稿日/直近の投稿のいいね数/直近の投稿のコメント数

ID	アカウント名	ユーザー名	アカウントURL	メディアタイプ	メディアID	投稿URL	投稿日	いいね数	コメント数
1	美容系アカウント	美容系アカウント	美容系アカウント	VIDEO		https://www.instagram.com/p/CD1234567890/	2023-05-12 10:00:00	458	7
2	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567891/	2023-05-12 10:00:00	981	0
3	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567892/	2023-05-12 10:00:00	1234	3
4	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567893/	2023-05-12 10:00:00	567	1
5	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567894/	2023-05-12 10:00:00	1987	2
6	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567895/	2023-05-12 10:00:00	145	0
7	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567896/	2023-05-12 10:00:00	1231	2
8	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567897/	2023-05-12 10:00:00	118	0
9	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567898/	2023-05-12 11:27	97	3
10	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567899/	2023-05-12 10:00:00	875	10
11	美容系アカウント	美容系アカウント	美容系アカウント	CAROUSEL_ALBUM		https://www.instagram.com/p/CD1234567900/	2023-05-12 10:00:00	800	15

図1 競合調査シート

3 テキストマイニングの目的

テキストマイニングとは、文章を単語ごとに分け、その単語の出現回数を調べるものである。その中でも、それぞれの単語の関連性や品詞を認識して分類してくれるものなど、目的に応じて様々な用途で使用されている。[1]

ここで「おすすめに表示される Instagram アカウントの投稿文」と「おすすめに表示されない Instagram の投稿文」に分けてテキストマイニングを行うことで、おすすめに表示されるために挿入するべきワードや言い回しを洗い出すことができると期待した。

4 投稿についての一般的な考え方

一般的に Instagram では以下のようなアルゴリズムになっていると言われている。

- 見ている人の投稿への滞在時間が長いほど評価が高くなり、おすすめに掲載されやすい。
- いいねやコメントが集まると評価が高くなり、おすすめに掲載されやすい。
- 投稿する画像の枚数は多ければ多いほど滞在時間が伸びておすすめに掲載されやすい。

5 テキストマイニングの手順

テキストマイニングは既存の仕組みを使用すれば結果を出力することが可能だが、仕組みを理解するために文献を参考にしながら環境構築をした。以下の手順でテキストマイニングを行う。

1. 文章を読み込み、文章を形態素ごとに分解する。
2. 形態素の中の機能語を取り除く。
3. 取り除いた後の形態素の出現回数をカウントする。
4. 形態素の品詞ごとにグループをわけ、色を振り分ける。
5. Wordcloud を用いて頻出回数が多いものは大きく、少ないものは小さく表示させる。

6 テキストマイニングの結果

図 2: おすすめに表示される Instagram のアカウント 5 つの直近 25 投稿の投稿文をテキストマイニングをした結果である。

図 3: おすすめに表示されない Instagram のアカウント 5 つの直近 25 投稿の投稿文をテキストマイニングをした結果である。

図 2、図 3 のいずれも、美容系に特化したアカウントでのみをピックアップしている。

