

画像キャプション生成モデル用いた 店舗評価システムの開発

塩崎 功也[†] 岩佐 英彦[†]

[†] 近畿大学工業高等専門学校 総合システム工学科 制御情報コース

1. はじめに

近年、画像キャプション生成の研究が盛んに行われている[1]。本稿では、SNSデータを用いて画像キャプション生成を行えるかを検討し、投稿評価モジュールとして店舗評価システムに導入することで、投稿の数と質に基づく店舗評価システムの開発を目指す。

2. 深層学習を用いた画像キャプション生成

2.1 画像キャプション生成モデル 本モデルは画像をCNNへ、単語列をLSTMに入力し、双方の特徴量から辞書登録されている単語出現確率を出力する。そこから最大値の単語を出力し、この処理を回帰的に文末まで行うことで文章を生成する。投稿評価モジュールとしてInstagramの投稿データを評価する為、訓練データにはInstagramの投稿データを用いる。モデルの訓練は交差エントロピー誤差関数を用いて誤差を算出し、各層へ逆伝播させることでパラメータの最適化を行った。

2.2 結果 開発したモデルに図1の画像を入力した結果を表1に示す。結果、文意を把握することが困難な文章が生成された。しかし画像の特徴を捉えた単語や#が付く単語が文末に生成されていることから、本稿のSNSデータに対する処理は一定の効果が示された。



図1. 入力画像

表1. 生成文

1:素晴らしい落ち着く神社疲れ入り嵐山自分道場所#森
2:今日求めソース秋焼き友達カレーライスたこ焼きごはん#麵#麵#珈琲
3:紅葉紅葉なっ秋下谷神社道場所町落ち着く#紅葉#神社

3. 店舗評価システム

3.1. システム概要 システムの概要を図2に示す。利用者はブラウザを通して、情報(現在地、検索範囲半径、キーワード)を送信する。その情報に基づき、システム側で外部データを取得し、評価順位を提供する。

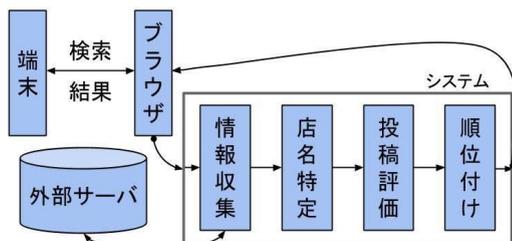


図2. システム概要図

3.2 投稿点数化手法 投稿の点数化手法を数式(3)に示す。数式(1)は画像テンソル I と単語列 W から2章のモデルが出力する v 種類の単語出現確率である。数式(2)は順序集合 A と自然数 x から集合要素 Ax を返す。数式(3)は投稿画像 I と投稿文 $\{W_1, W_2, \dots, W_n\}$ から投稿点数を返す。

$$M(I, W) = \{P_1, P_2, \dots, P_v\} \quad (1) \quad ID(A, x) = Ax \quad (2)$$

$$S(I, W) = \frac{1}{n} \sum_{i=0}^n ID(M(I, \{W_x \in W | 0 \leq x < i\}), W_i) \quad (3)$$

4. 店舗評価システム実行結果

3.1節で述べた情報に(近大高専, 5000, カフェ)を入力した結果を図3に示す。図中左の結果は投稿数による評価順位である。図中右の結果は3.2節の手法で投稿の数と質に基づいた点数化を行った評価順位である。



図3. 検索結果

投稿数が少ない店舗であっても投稿文に店舗の高評価に結びつく単語が多く含まれる場合、評価点が向上し順位が変動することを確認できた。

5. 今後の課題

小規模な実験により、投稿画像に対する投稿文に店舗の高評価に結びつく単語が多く含まれる場合には店舗の評価点が向上することが確認できたので、画像キャプション生成モデルに用いる学習データを増加させて大規模な実験を行うことが今後の課題である。

参考文献

[1] 小林豊ほか, "深層学習による日本語キャプション生成システムの開発", 人工知能学会, SIG-AM-17-04, 2017.