

品詞に着目した料理レシピタイトルの魅力度分析

高木 七海[†] 久徳 遙矢^{††,†} 道満 恵介^{†††,†} 川西 康友^{††††,†}
平山 高嗣^{†††††,†} 駒水 孝裕[†] 井手 一郎[†]

[†] 名古屋大学 〒464-8601 愛知県名古屋市千種区不老町

^{††} 愛知工科大学 〒443-0047 愛知県蒲郡市西迫町馬乗 50-2

^{†††} 中京大学 〒470-0393 愛知県豊田市貝津町床立 101

^{††††} 理化学研究所 〒619-0288 京都府相楽郡精華町光台 2-2-2

^{†††††} 人間環境大学 〒444-3505 愛知県岡崎市本宿町上三本松 6-2

E-mail: [†] takagin@cs.is.i.nagoya-u.ac.jp, taka-coma@acm.org, ide@i.nagoya-u.ac.jp, ^{††} kyutoku-haruya@aut.ac.jp,
^{†††} kdoman@sist.chukyo-u.ac.jp, ^{††††} yasutomo.kawanishi@riken.jp, ^{†††††} t-hirayama@uhe.ac.jp

あらまし 近年、Web 上のレシピサイトを利用する機会が増え、投稿型レシピサイトで自身が考案した料理レシピを公開する人も増えている。一般に料理レシピの投稿者は、自身の料理レシピを閲覧、調理してもらうことを目的としている。そのため、閲覧者にとって魅力的な料理レシピの作成が必要であり、そのタイトルは人の目を惹くための重要な要素であると考えられる。本研究では、人が料理レシピのタイトルから直感的にその料理の美味しさや手軽さなどを総合的に想起して作りたくなる度合いを魅力度と定義する。そのうえで、魅力的なタイトルの自動生成を目的として、既存の多数の料理レシピにおけるタイトルの特性を分析する。具体的には、品詞に着目してタイトルを分析したうえで、タイトルの長さや品詞構成と魅力度の関係を分析した結果について報告する。

キーワード 料理レシピ, タイトル分析, タイトル生成, 品詞解析

Analyzing Attractiveness of Cooking Recipe Titles Based on Parts of Speech

Nanami TAKAGI[†], Haruya KYUTOKU^{††,†}, Keisuke DOMAN^{†††,†}, Yasutomo KAWANISHI^{††††,†},
Takatsugu HIRAYAMA^{†††††,†}, Takahiro KOMAMIZU[†], and Ichiro IDE[†]

[†] Nagoya University Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601 Japan

^{††} Aichi University of Technology 50-2 Manori, Nishihassama-cho, Gamagori, Aichi, 443-0047 Japan

^{†††} Chukyo University 101 Tokodachi, Kaizu-cho, Toyota, Aichi, 470-0393 Japan

^{††††} RIKEN 2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0288 Japan

^{†††††} University of Human Environments 6-2 Kami Sanbonmatsu, Motojuku-cho, Okazaki, Aichi, 444-3505 Japan

E-mail: [†] takagin@cs.is.i.nagoya-u.ac.jp, taka-coma@acm.org, ide@i.nagoya-u.ac.jp, ^{††} kyutoku-haruya@aut.ac.jp,

^{†††} kdoman@sist.chukyo-u.ac.jp, ^{††††} yasutomo.kawanishi@riken.jp, ^{†††††} t-hirayama@uhe.ac.jp

Abstract Recently, occasions to use recipe Websites is increasing, and also the number of users who publish their own cooking recipes on these Websites is increasing. In general, the cooking recipe publishers wish that the published cooking recipes will be viewed and actually cooked by other users. Therefore, it is necessary to create attractive recipes, and their titles are considered as an important element to attract users' attentions. In this study, we define attractiveness as the degree to which people intuitively feel the taste and ease of cooking from the recipe title and want to cook it. From the above background, we analyze the characteristics of titles for many existing cooking recipes for realizing automatic attractive title generation. Specifically, we report the analysis of titles focusing on their parts of speech, and then the analysis of the relationship between their length or part-of-speech structure and the attractiveness of the title.

Key words Cooking recipe, title analysis, title generation, parts-of-speech analysis

1. はじめに

一般の人々が家庭で料理をする際には、料理レシピを参考にすることが多い^(注1)。近年、料理レシピの入手先として Web 上のレシピサイトやレシピアプリを利用する人が増えている。実際に、レシピサイトやレシピアプリを利用している人の割合は、2022 年時点で月に 1 回以上料理をする人のうち 75.2 % を占めると報告されている^(注1)。2013 年時点では本割合が 68.9 % であると報告されていることから^(注2)、年々増加している。以上のように、レシピサイトを利用する機会が増えているが、その中で、投稿型レシピサイトに自身が作成した料理レシピを投稿する機会も増えている。

一般に投稿型レシピサイトでは、投稿者は、自身の料理レシピを閲覧してもらい、また調理してもらうことを目的として、図 1 左上のような料理レシピを投稿する。一方で、閲覧者はキーワード検索をし、このとき、表示された検索結果の一覧から目を惹かれた料理レシピを閲覧する。表示された料理レシピの中でも、タイトルや料理画像から直感的に美味しさや手軽さを感じられるものを閲覧すると考えられる。これらのレシピサイトには、図 1 右上のような閲覧者が料理レシピを閲覧して実際に調理したことをフィードバックするレポート機能^(注3)がある。これにより、投稿者は自身が投稿した料理レシピの人気度を把握できる。投稿者がレポートを多く得るためには、まず閲覧者の目を惹く料理レシピを作成する必要がある。閲覧者の目を惹く料理レシピは、美味しさや手軽さを想起させるものであり、これらの情報を含むか否かで、同種の料理レシピであっても料理レシピに対する印象が大きく変化する。そのため、料理レシピのタイトルは人の目を惹くための重要な要素であるといえる。しかし、一般の投稿者が料理レシピの特徴を的確に捉えた魅力的なタイトルを付与することは必ずしも容易ではない。

これらの背景に基づき、我々は、料理レシピに対する魅力的なタイトルの作成支援が重要であると考えた。そこで、本研究では、そのような支援の一つとして魅力的なタイトルの自動生成を実現するため、既存の料理レシピに付与されたタイトルの品詞構成を分析し、魅力度との関係について明らかにすることを目的とする。図 1 に示すように、タイトルに使用されている品詞の種類や並び、品詞の数、品詞が魅力度に及ぼす影響など、タイトルの特性が判れば、品詞構成のテンプレートを用意することができ、魅力的なタイトル生成が可能になると考えられる。そのため、本報告ではタイトルの特性の中でも品詞特性に注目して、タイトルの品詞構成の分析、タイトルの長さや魅力度の

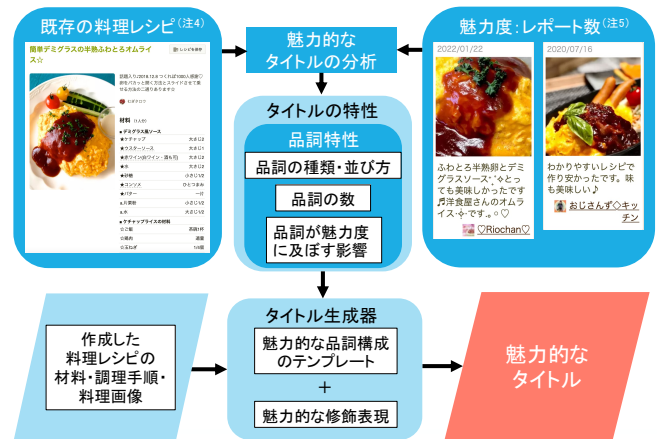


図 1 魅力的なタイトルの生成方法の概要

関係の分析、タイトルを構成する品詞の割合と魅力度の回帰分析を行なう。ここでは、人が料理レシピのタイトルから直感的に美味しさや手軽さなどを総合的に想起して作りたくなる度合いを魅力度と考え、これを閲覧者による各レシピへのレポート数として定義する。以降、2. で関連研究について述べ、3. で分析方法、4. で分析実験について述べる。最後に、5. で本報告をまとめる。

2. 関連研究

2.1 料理レシピにおけるタイトルの生成に関する研究

料理レシピにおけるタイトルの生成に関する研究としては、料理レシピの中身や関連する文書を手がかりに、タイトルを生成したものがある。金内ら [1] は料理レシピの特徴を的確に表したタイトルを生成することを目指し、料理レシピに対するレポートに基づいて、タイトルを生成する手法を提案している。この手法では、手がかり語を指定し、その直前に出現する単語を料理レシピの特徴や目的を表現する情報として抽出する。そして、あらかじめ決められたテンプレートにその情報を埋め込む形で、タイトルを生成する。この研究では、レポートからタイトル生成を行なうことを目的としているのに対して、本研究では、閲覧者の目を惹くタイトルを生成することを目的としている点が異なる。またこの研究では、例えば『『運動会におすすめ』の『お弁当』』のように、『『料理をする目的』の『料理名』』という一つの決まったテンプレートに当てはめる形でタイトルを生成しているのに対して、本研究では、魅力的なタイトルの構造自体を分析し、様々なパターンの可能性を考えるため、より柔軟かつ多様なタイトルを生成できるようになると期待される。

2.2 料理レシピにおけるタイトルの分析に関する研究

料理レシピにおけるタイトルの分析に関する研究として、橘

(注1)：株式会社マルハニチロホールディングス，“～マルハニチロホールディングス、「料理レシピに関する調査」～”，https://www.maruha-nichiro.co.jp/corporate/news_center/news_topics/20200818_research_recipe2020_1.pdf Aug. 2020 [2023/2/21 参照]。

(注2)：株式会社マルハニチロホールディングス，“～マルハニチロホールディングス、「料理レシピに関する調査」～”，https://www.maruha-nichiro.co.jp/corporate/news_center/research/pdf/20130227_recipe_cyouusa.pdf Feb. 2013 [2023/2/21 参照]。

(注3)：クックパッドにおいて「つくれば」、楽天レシピにおいて「つくったよレポート」と呼ばれている。

(注4)：むぎタロウ，“簡単デミグラスの半熟ふわとろオムライス☆”，<https://cookpad.com/recipe/2163455/> [2023/2/21 参照]。

(注5)：クックパッド株式会社，“簡単デミグラスの半熟ふわとろオムライス☆のつくれば”，<https://cookpad.com/recipe/2163455/tsukurepos/> [2023/2/21 参照]。

ら [2] は、タイトル名に含まれる「簡単」、「子供が喜ぶ」、「ヘルシー」といった修飾表現に着目し、それらの根拠（ネーミングコンセプト）を料理レシピから抽出する手法を提案している。この手法では、同種の料理に対するレシピを集約して、その料理における典型的な要素を抽出する。そこから、各料理レシピの要素との差異を抽出することで、ネーミングコンセプトを抽出する。一方、Kusmierczyk ら [3] は料理レシピのタイトルとその料理の栄養素の相関を分析している。その結果、両者に強い相関関係があることを明らかにし、料理レシピのタイトルからその料理の栄養素を予測している。これらのように、料理レシピのタイトルの特徴や、料理レシピの内容とタイトルの整合性に注目した研究は様々に行なわれている。しかし、タイトルの魅力度に着目した研究は、これまで行なわれていない。

2.3 料理画像のキャプションに関する研究

料理画像のキャプションに関する研究がいくつか存在する。西村ら [4] は、画像付き料理レシピの作成を容易にすることを目的とし、調理過程を撮影した画像列を入力として料理レシピを生成する手法を提案している。具体的には、画像列を入力として文章を出力する Visual storytelling のモデルを用いて、入力画像に適した重要語を過不足なく含む表現の情報を活用しながら料理レシピを生成している。名高ら [5] は料理画像から料理名と材料を同時推定する手法を提案している。料理画像には複数の材料が含まれるために、その各々の推定が困難であるが、推定精度を向上させるために Transformer decoder を用いたモデルを提案している。これらの研究では、画像やテキスト情報を使用して料理レシピの特徴を的確に捉えることを目標としているが、魅力度の検討は行なわれていない。

3. タイトルの分析方法

投稿型レシピサイトに投稿された料理レシピのタイトルを以下に例示する。

- (1) 濃厚きのこたっぷりクリームシチュー
- (2) 祖母直伝☆絶品肉じゃが
- (3) 簡単♪美味しいあっさりポトフ

これらはいずれも、料理を表現するために最も重要な料理名と、その料理レシピ固有の特徴を表現するための修飾表現の組み合わせから構成されている。この修飾表現として、いずれの例も、名詞の「濃厚」、「絶品」、「簡単」や、形容詞の「美味しい」、更には副詞の「たっぷり」など、様々な品詞を含んでいる。更に、タイトル例 (2), (3) のように記号を用いているものも多数存在する。実際に、後述のクックパッドデータセット [6] に含まれる約 172 万件の料理レシピのタイトルを調査したところ、記号が使用されているタイトルの割合は 57.5 % と過半数を占めていた。投稿者はこのように料理名だけでなく修飾表現をタイトルに含めることで、料理の美味しさや手軽さを想起しやすくし、料理レシピの魅力度を高めていると考えられる。しかし、修飾表現は多様な品詞を含み、またその並び方も多様である。そこで、我々は、品詞構成を分析することが、タイトルとその魅力度の関係を考える際に重要であると考え、それらの関係を分析することにした。

表 1 説明変数の計算例

タイトル：絶品！魚介のグラタン

品詞構成：名詞＋記号＋名詞＋助詞＋名詞

名詞	記号	助詞	動詞	副詞	形容詞	助動詞	...
0.60	0.20	0.20	0.00	0.00	0.00	0.00	

表 2 レポート数の分類方法

階級	ラベル
1 ~ 3	0
3 ~ 7	1
7 ~ 20	2
20 ~ 55	3
55 ~ 148	4
148 ~ 403	5
403 ~ 1,097	6
1,097 ~ 2,981	7
2,981 ~ 8,103	8
8,103 ~	9

以降、3.1 でタイトルの品詞構成の分析、3.2 でタイトルの長さや魅力度の関係の分析、3.3 でタイトルを構成する品詞の割合と魅力度の回帰分析について詳細に述べる。

3.1 タイトルの品詞構成の分析

本分析では、タイトルに使用されている品詞の種類や並びを明らかにする。まず形態素解析と品詞タグ付けによりタイトルを品詞分解する。本処理には、日本語形態素解析器 MeCab^(注6) を使用する。MeCab は処理エンジン本体の他に辞書を必要とし、一般的に IPA 辞書 [7] が推奨されているが、新語などに対応するために、Web 上の多数の言語資源から得た新語を含む mecab-ipadic-NEologd [8] を追加して辞書を拡張した。以上のように品詞分解を行なった結果から得られる、タイトルの品詞構成に基づいて魅力的なタイトルの傾向を分析する。ここで、「夏＋野菜」や「豚しゃぶ＋冷麺」、「♪＋♪」のような連続する同一品詞を個別に取り扱おうと、品詞構造のパターンが過剰に細分化され、傾向の分析が困難となることが予想される。そこで、「夏＋野菜」は「夏野菜」という一つの名詞として扱うように、連続する同一品詞を結合して分析する。

3.2 タイトルの長さや魅力度の関係の分析

本分析では、魅力的なタイトルに含まれる修飾表現の量を明らかにする。ここで、本節の冒頭で示したタイトル例 (3) では「簡単♪美味しいあっさりポトフ」と表現されており、単に「ポトフ」のような料理名のみからなるタイトルよりも構成する品詞の数が多い。このように、修飾表現が多いほどタイトルが長くなることをふまえて、タイトルを構成する品詞の数をタイトルの長さとして定義し、レポート数との関係を分析する。なお、ここでは 3.1 のような連続する同一品詞の結合は行なわない。

3.3 タイトルを構成する品詞の割合と魅力度の回帰分析

本分析では、各品詞が魅力度に及ぼす影響を明らかにするために、タイトルを構成する品詞の割合を説明変数、レポート数を

(注6)：工藤拓，“MeCab: Yet another part-of-speech and morphological analyzer,” <http://taku910.github.io/mecab/> [2023/2/21 参照].

表3 全ての料理レシピに対するレポート数の統計情報

平均	56.51
標準偏差	56.52
第1四分位数	0
第2四分位数	0
第3四分位数	2
最大値	15,654
最小値	0

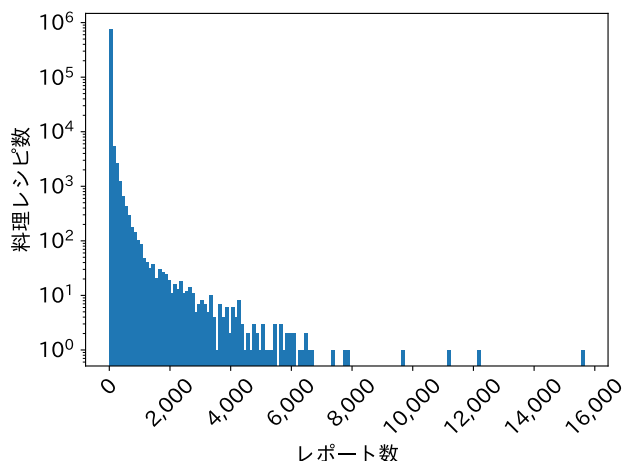


図2 全ての料理レシピにおけるレポート数の頻度分布

目的変数として回帰分析を行なう。説明変数の計算例を表1に示す。タイトルを構成する品詞の出現頻度から品詞の割合を算出して、説明変数とする。目的変数は、表2に示すように、指数関数を用いたビン幅によりレポート数を10階級に分類して、階級のラベルに0~9の数値を付与する。これは、インフルエンサーなどが投稿した、極端にレポート数が多い料理レシピが存在することを考慮して、レポート数の偏りを軽減させるための処理である。回帰分析モデルには、説明変数を段階的に分類していく決定木アルゴリズムを用いる。決定木の有益な点は、得られた回帰モデルに対する各説明変数の寄与度合いを評価する重要度を定量化できる点である。これにより、タイトルを構成する品詞がレポート数に及ぼす影響度合いを明らかにする。また、平均絶対誤差 (MAE ; Mean Absolute Error) を用いて回帰の性能を評価する。MAEは各予測誤差の絶対値の平均から算出される指標である。実測値を y_i ($i = 1, 2, 3, \dots, n$)、モデルから計算した予測値を \hat{y}_i とし、以下の式で定義される。

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

4. タイトルの分析実験

本節では、3.で説明した料理レシピにおけるタイトルの分析実験について報告する。まず4.1で、使用するデータセットの詳細について紹介する。次に4.2で、各分析実験について報告する。そして4.3で、実験結果について考察する。

4.1 データセット

本実験では、国立情報学研究所情報学研究データリポジトリ

表4 相対的に魅力的な料理レシピにおけるタイトルの品詞構成

順位	品詞構成	頻出度
1	名詞+助詞+名詞	6%
2	名詞+記号+名詞	5%
3	名詞	4%
4	名詞+助詞+名詞+助詞+名詞	3%
5	名詞+記号+名詞+助詞+名詞	2%

表5 相対的に魅力的でない料理レシピにおけるタイトルの品詞構成

順位	品詞構成	頻出度
1	名詞	13%
2	名詞+助詞+名詞	7%
3	名詞+記号+名詞	4%
4	名詞+助詞+名詞+助詞+名詞	3%
5	名詞+助詞+名詞+助詞+名詞+動詞	2%

を通じて、クックパッド株式会社から提供されたクックパッドデータセット [6] のうち、投稿された料理レシピ約172万件と、それらに対するレポート(「つくれば」と呼ばれる)のデータを使用した。各料理レシピに対するレポートを数え、各料理レシピに対する「レポート数」とした。このレポートに関する統計情報の詳細を表3に示す。また、それぞれのレポート数に対する料理レシピ数を対数スケールを用いて図2に示す。この表3及び図2において、第2四分位数まで0であったため、レポート数が0件のものを数えたところ、全体の56%を占めていた。また、標準偏差は56.52であり、レポート数の偏りが大きいことが判った。以降の分析では、これらの統計情報をふまえて使用する料理レシピを適宜選択した。

4.2 分析実験

ここでは、まず4.2.1でタイトルの品詞構成の分析、次に4.2.2でタイトルの長さや魅力度の関係の分析、最後に4.2.3でタイトルを構成する品詞の割合と魅力度の回帰分析の実験についてそれぞれ述べる。

4.2.1 タイトルの品詞構成の分析実験

3.1で述べた手順に従って、品詞構成を分析した。4.1で示した統計情報から、レポート数が1桁である料理レシピの割合が非常に大きいことをふまえて、レポート数が100件以上の約1万2千件の料理レシピは相対的に魅力的であると考え、タイトルの品詞構成を分析した。品詞構成のうち、頻出した上位5件を表4に示す。また、魅力的なタイトルの比較対象とするために、レポート数が100件未満の料理レシピを相対的に魅力的でないと考えて、タイトルの品詞構成を分析した。品詞構成のうち、頻出した上位5件を表5に示す。いずれも、上位5パターンだけで全体の2割以上のタイトルを網羅している。

4.2.2 タイトルの長さや魅力度の関係の分析実験

3.2で述べた手順に従って、データセット中の全ての料理レシピ約171万件を対象にして、タイトルを構成する品詞数とレポート数との関係を分析をした。ここでは、タイトルを構成する品詞の数を長さとし、レポート数が多いものを魅力度が高いと考えた。タイトルの長さやレポート数の関係を散布図で表し

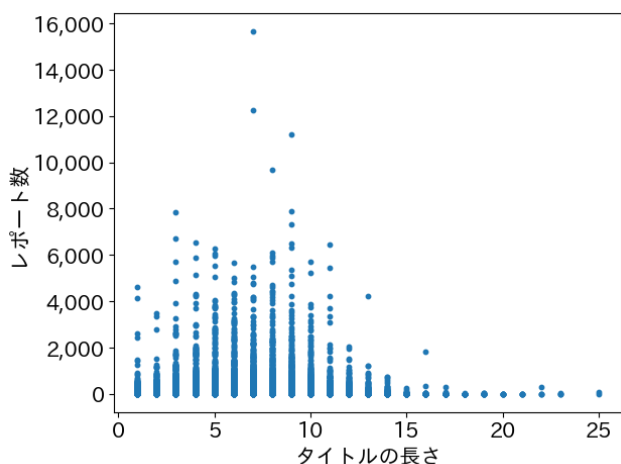


図3 タイトルの長さとレポート数の関係

たものを図3に示す。

4.2.3 タイトルを構成する品詞の割合と魅力度の回帰分析実験

3.3で述べた手順に従って、タイトルを構成する品詞の割合を説明変数、レポート数を目的変数として回帰分析を行なった。説明変数、目的変数は表1と表2でそれぞれ述べた手順に従って算出した。また、4.1で示した統計情報から、データセットの大半を占めるレポート数0件のレシピは回帰の障害になると考え、レポート数が1件以上の料理レシピを対象として分析した。分析対象データの7割を学習データとして使用し、残りの3割をテストデータとした。そして、LightGBM [9]を用いて回帰モデルを作成した。この際、Optuna [10]を用いたパラメータチューニングにより、回帰時のハイパパラメータを決定した。回帰モデルの評価指標としては、3.3で述べたMAEを用いる。以上のようにして回帰分析を行ない、グラフを用いて実測値と予測値の比較を行なった。このとき、実測値により分けられた料理レシピの階級毎の予測値の平均値を用いて描画を行なった。

得られた回帰モデルによる予測値と実測値のグラフを図4に示す。MAEは0.79と比較的小さな値を示したが、図のように予測値と実測値に大きな差がみられ、予測値がグラフ上でほぼ一定となった。そのため、予測値と実測値で縦軸の範囲を変えて描画したグラフを図5に示す。予測値が階級毎の大小関係を正しく示していることから、得られた回帰モデルは実測値の差異と同様の傾向を予測できていることを確認した。また、各説明変数の重要度を図6に示す。名詞、記号、助詞の重要度が非常に高く、形容詞、動詞、副詞の重要度も高かった。一方で、感動詞、接続詞、連体詞の重要度は低かった。

4.3 考察

4.3.1 タイトルの品詞構成の分析実験の考察

レポート数が100件未満の料理レシピにおけるタイトルの品詞構成パターン(表5)と、魅力的なタイトルの品詞構成パターン(表4)を見ると、前者では名詞のみで構成されたタイトルの割合が最も高かったのに対し、後者では名詞+助詞+名詞や記号+名詞+記号のような品詞構成が、名詞のみで構成された

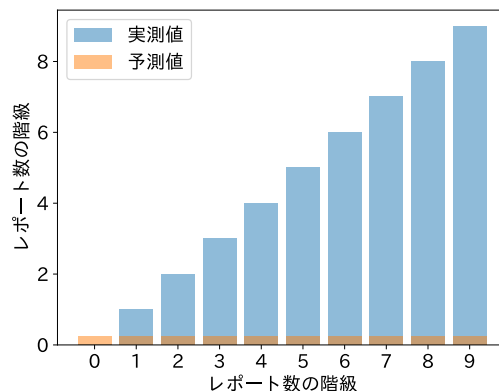


図4 回帰分析結果

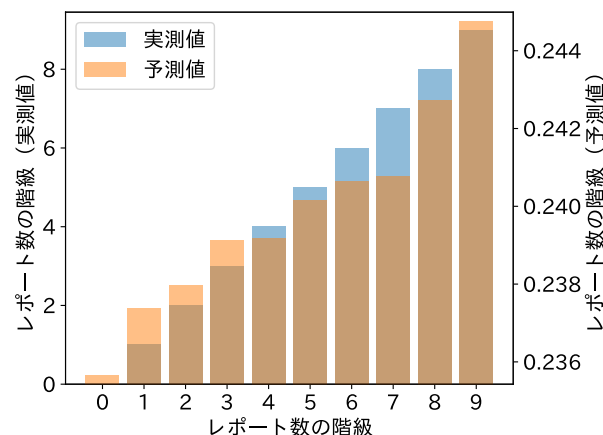


図5 回帰分析結果 (予測値と実測値で縦軸の範囲を変えた場合)

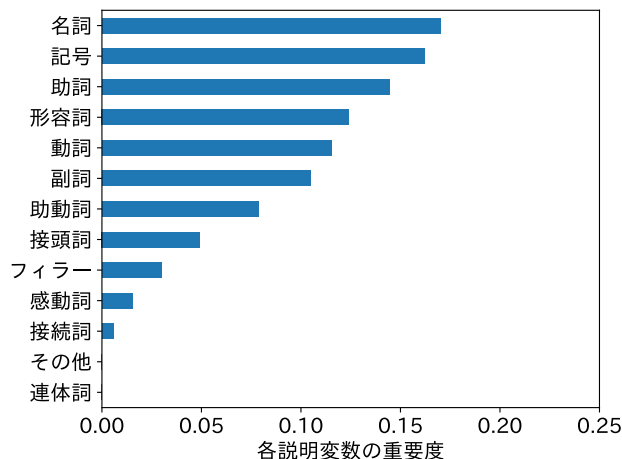


図6 得られた回帰モデルにおける各説明変数の重要度

タイトルよりもわずかに高い割合を示していた。これらの結果から、魅力的なタイトルには料理名を修飾する記号や名詞がより多く含まれる傾向にあることが示唆された。

4.3.2 タイトルの長さや魅力度の関係の分析実験の考察

図3から、品詞数が3~11個の場合にレポート数が多い結果となった。このことから、1つの品詞のみ用いたような単純なタイトルよりも、修飾表現を含むタイトルの方が魅力的であ

表6 実測値の差が n 以上 ($n = 1, 2, 3, 4$) のレシピ間の大小関係の正解率

実測値の差	正解率
1 以上	53.32 %
2 以上	54.87 %
3 以上	56.36 %
4 以上	57.30 %

ると言える。一方、12 個を超えるとレポート数が少なくなる傾向が見られた。この結果から、いくらタイトルに修飾表現が多くても、非常に長いタイトルは複雑な構造となり、魅力的ではなくなると考えられる。つまり、タイトルに含める修飾表現を適切に取捨選択することが魅力的なタイトルを構成する上で重要であることが示唆された。

4.3.3 タイトルを構成する品詞の割合と魅力度の回帰分析 実験の考察

この実験では、説明変数と目的変数の関係を正確に回帰するモデルは構築できなかった。この原因としては、レポート数を10段階に分類しても、依然としてそれぞれの階級の料理レシピ数に大きな偏りがあったことが考えられる。また、レポートはタイトルだけでなく、料理画像や調理手順の良し悪しなど様々な内容に対するフィードバックであるため、説明変数との関係は単純でないと考えられる。回帰に失敗している一方で、MAEの値は小さな値であった。これは、階級ラベル0や1に含まれる料理レシピが大半を占めたため、図4のような回帰結果であってもMAEが小さくなったと考えられる。このように、予測値そのものは実測値と大きな差が見られたが、図5のように相対的な差を可視化すると、実測値が指数的に大きな料理レシピに対して、予測値も相対的には指数的に大きな値を予測できていた。また、このことを確認するため、実測値の階級の差が n 以上 ($n = 1, 2, 3, 4$) のレシピ間の大小関係を予測できているかを正解率としてそれぞれ算出した。その結果を表6に示す。表から、実測値の差が大きいほど、すなわち差異が明確であると考えられる階級間ほど、タイトルの相対的な魅力度を正確に判定できていた。このことから、各階級の上下関係を概ね表現可能な回帰モデルが得られたと言える。従って、各説明変数の重要度はある程度信頼できると考えられる。各説明変数の重要度をみると、レポート数に非常に大きく影響する品詞は名詞、記号、助詞であることが判った。そのため、4.2.1の分析結果と同様に、記号や助詞などを用いた修飾的な表現を考慮することが、魅力的なタイトルを生成する上で重要であることが示唆された。

5. む す び

本報告では、魅力的なタイトルを自動生成するための第一歩として、既存の多数の料理レシピにおけるタイトルの品詞構成を分析し、それに基づいて、タイトルと魅力度との関係を分析した。具体的には、まず、魅力的なタイトルに使用されている品詞の種類や並びを明らかにするため、タイトルの品詞構成を分析した。次に、魅力的なタイトルを構成する品詞の数を明らか

にするため、タイトルの長さや魅力度の関係を分析した。最後に、品詞の割合が魅力度に及ぼす影響を分析するため、タイトルを構成する品詞の割合と魅力度の回帰分析をした。その結果から、魅力的なタイトルを作成するためには、記号や助詞などで料理名を修飾しつつ、かつ簡潔に表現することが重要であることが明らかになった。しかし、タイトルの品詞構成のみからは、魅力度として定義したレポート数を正確に推定することはできなかった。これは、料理レシピの魅力度はタイトルだけでなく、画像や調理手順などから総合的に決まるためであると考えられる。また、タイトルについても、品詞構成のみではなく、用いられている単語の適切さなど、本研究で想定した要因以外の様々な要因があると考えられる。

今後の課題としては、以下が挙げられる。

- 回帰分析の外れ値判定及び説明変数、目的変数の改良
- タイトルと料理カテゴリの関係の分析
- タイトル生成のためのテキスト及び画像の分析

謝辞 本研究の一部はJSPS 科研費 JP20K12038, JP22H00548 の支援による。また本研究では、国立情報学研究所 IDR データセット提供サービスによりクックパッド株式会社から提供を受けた「クックパッドデータセット」[6]を利用した。

文 献

- [1] 金内萌, 難波英嗣, 角谷和俊, “投稿型レシピサイトにおけるレビュー情報に基づく料理タイトル自動生成,” 第8回データ工学と情報マネジメントに関するフォーラム, no.P3-2, Mar. 2016.
- [2] 橘明穂, 若宮翔子, 難波英嗣, 角谷和俊, “料理名の修飾表現の関係性に基づくレシピのネーミングコンセプト抽出,” 電子情報通信学会技術研究報告, DE2013-36, Sept. 2013.
- [3] T. Kusmierczyk and K. Nørnvåg, “Online food recipe title semantics: Combining nutrient facts and topics,” Proceedings of the 25th ACM International Conference on Information and Knowledge Management, pp.2013–2016, May 2016.
- [4] 西村太一, 橋本敦史, 森信介, “重要語に着目した写真列からのレシピの自動生成,” 自然言語処理, vol.27, no.2, pp.257–279, June 2020.
- [5] 名高祐輔, 青野雅樹, “Transformer decoder を用いた料理画像からの料理名と食材の推定,” 第12回データ工学と情報マネジメントに関するフォーラム, no.B8-3, Mar. 2020.
- [6] クックパッド株式会社, “クックパッドデータ,” 国立情報学研究所情報学研究データリポジトリ, Feb. 2015.
- [7] 浅原正幸, 松本裕治, “IPAdic version 2.7.0 ユーザーズマニュアル,” Nov. 2003.
- [8] 佐藤敏紀, 橋本泰一, 奥村学, “単語分かち書き辞書 mecab-ipadic-NEologd の実装と情報検索における効果的な使用方法の検討,” 言語処理学会第23回年次大会, no.B6-1, Mar. 2017.
- [9] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, “LightGBM: A highly efficient gradient boosting decision tree,” Proceedings of the 31st International Conference on Neural Information Processing Systems, pp.3149–3157, Dec. 2017.
- [10] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A next-generation hyperparameter optimization framework,” Computing Research Repository arXiv Preprint, arXiv:1907.10902, July 2019.