

料理レシピの分かりやすさ評価指標

岸田 大陸
Riku Kishida

1. はじめに

近年、クックパッドや楽天レシピなどのユーザー投稿型の料理レシピサイトの利用者が増加している。これらのサイトでは、一般のユーザーが自らのレシピを投稿し、他のユーザーと共有することができるため、料理初心者から上級者まで幅広い層にとって便利な情報プラットフォームとなっている。また、特定の料理ジャンルや目的に応じた多様なレシピが手軽に検索できるという点も、利便性の高さに寄与している。しかしその一方で、誰でも投稿できるという開かれた仕組みにより、内容の質にばらつきが生じるという課題も存在する。中には、説明が不十分で手順が曖昧なレシピ、用語が専門的すぎて初心者には理解が難しいレシピなども多く、利用者が正確に調理手順を再現できないケースも見受けられる。このような背景から、レシピの「分かりやすさ」を定量的に評価することができれば、ユーザーにとって有益なフィルタリングの指標となり、実用的な価値が高いと考えられる。具体的には、レシピを閲覧する側のユーザーは、分かりやすさの高いレシピを優先的に選択することが可能となり、より正確で効率的な調理が期待できる。一方、投稿する側のユーザーにとっても、分かりやすさの評価をフィードバックとして受け取ることで、自身のレシピの改善につなげることができると考えられる。従来、文章の分かりやすさに関する研究では、文字種のバランスや文の長さといった形態的要素に着目する傾向が強かった。例えば、計算機マニュアルの分かりやすさの定量的評価方法[1]では、ひらがな・漢字の使用割合や見出しの有無といった形式的な指標を説明変数とし、アンケート結果などから得られた「分かりやすさの印象評価」を目的変数として重回帰分析を行っている。しかし、料理レシピというテキストジャンルにおいては、形態的な特徴だけでなく、調理に固有の語彙や手順表現、曖昧な表現といったレシピ特有の要素が多く含まれている。これらの要素は、一般的な文書とは異なる難しさを持ち、形態的要素のみを用いた従来の評価指標では十分に評価しきれない可能性がある。したがって、料理レシピに特化した分かりやすさ評価指標を検討する必要がある。本研究では、料理レシピのテキストデータにおける分かりやすさを、形態的要素に加えて、曖昧な表現や専門用語といった内容的特徴も考慮した上で、多角的に評価する手法を提案する。これにより、レシピ特有の難しさや独自の表現にも対応可能な、より実用的な評価が実現できると考えている。

2. 関連研究

2.1 計算機マニュアルの分かりやすさ

関連研究として、計算機マニュアルの分かりやすさの定量的評価方法[1]がある。この研究では、形態的要素（ひらがな、漢字の割合、見出しの数など）を説明変数、計算機マニュアルの分かりやすさのアンケート結果を目的変数として重回帰分析を行い、モデル評価式を算出している。

2.2 日本文の読みやすさの評価式

分かりやすさに似た研究として日本文の読みやすさの評価式[2]がある。この研究では、入門書などの文章は分かりやすい、法律関係の文章を分かりにくいなどと定義し、形態的要素を説明変数として回帰式を算出している。説明変数として以下の変数が用いられている。

- 文の平均の長さ(s)
- アルファベット連の平均の長さ(a)
- ひらがな連の平均の長さ(h)
- 漢字連の平均の長さ(c)
- カタカナ連の平均の長さ(k)
- 句点(。)あたりの読点(、)の数(cp)

そして、以下の評価式が算出されている。

$$RS = -0.12s - 1.37a + 7.4h - 23.18c - 5.4k - 4.67cp + 115.79$$

この式から読み取れることとして、例えば漢字連の平均の長さが高くなると、読みやすさが大きく下がることがわかる。一方で、ひらがな連の平均の長さが高くなると、読みやすさは上がることがわかる。こうした結果から、日本語の読みやすさには文字種や構文構造といった形態的な要素が大きく関連していることがわかる。

2.3 BERT

BERT[3]とは Transformer の Encoder を使っている事前学習モデルである。従来の自然言語処理モデルの場合、文章を単一方向からしか処理することができなかった。しかし、文頭、文末の双方向からエンコードすることでこれまでの手法よりも精度が格段に向上している。BERT は事前学習として Masked Language Model (MLM) が行われている。MLM は図のように行われている。まず、入力 の 15% のトークンを [Mask] トークンに置き換える。そして、[Mask] トークンの内 10% をランダムな単語に変換し、さらに 10% をそのままにする。置換された単語を周りの文脈から当てるタスクを解くこ

とで、単語に対応するための文脈情報を学習していく。本研究ではこれを使用する。

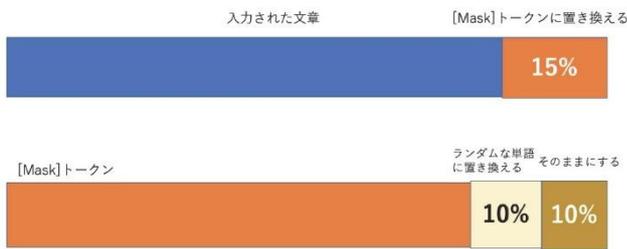


図 2.1: MLM の学習方法

3. 提案手法と目的

従来の文章の分かりやすさの研究では、形態的要素に着目するが多かった。そこで、本研究では、形態的要素以外の観点で、料理レシピのテキストデータから分かりやすさを定量的に評価することを目的とする。

3.1 分かりやすさのラベル付け

Google Forms を用いて、大学 4 年生および大学院生を対象に、料理レシピの分かりやすさに関するアンケート調査を実施した。対象者は合計 20 名である。アンケートでは、事前に選定した 30 個の料理レシピについて、それぞれのレシピがどの程度分かりやすいと感じたかを尋ねた。評価方法としては、1 (分かりにくい) から 5 (分かりやすい) までの 5 段階評価を採用した。調査においては、各被験者が全てのレシピに対して評価を行い、その結果として、各レシピに対する評価スコアの平均値を算出した。この平均スコアは、各レシピに対して「分かりやすさラベル」として用いることができ、後の分析において目的変数として扱うことが可能である。図 2.1 は、実際に被験者に提示されたアンケートフォームの一部を示しており、回答者がどのような形式で評価を行ったかを視覚的に確認できる。一方、図 2.2 では、全 30 レシピに対する評価結果の分布をヒストグラムとして示しており、どの程度の分かりやすさのばらつきが存在するか、また全体として評価が高い傾向にあるか低い傾向にあるかを概観することができる。これらの図は、アンケート結果の妥当性や分布の偏りの有無を確認する上でも有用である。

☆ほうれん草の白和え☆

豆腐は、ペーパーで水切りしておきます。

人参は細切りにして、塩ゆでしておきます。ほうれん草も、熱湯に根のほうから入れ、さっと塩ゆでして冷水に取ります。

ほうれん草をペーパー等でよく水切りして、3cm位に切ります。

ボウルにこれらの材料と豆腐、人参、ほうれん草を加え、なめらかになるまで、よく混ぜ合わせたら出来上がり！

この料理レシピの分かりやすさ

1 2 3 4 5

分かりにくい ○ ○ ○ ○ ○ 分かりやすい

図 3.1: アンケート画面

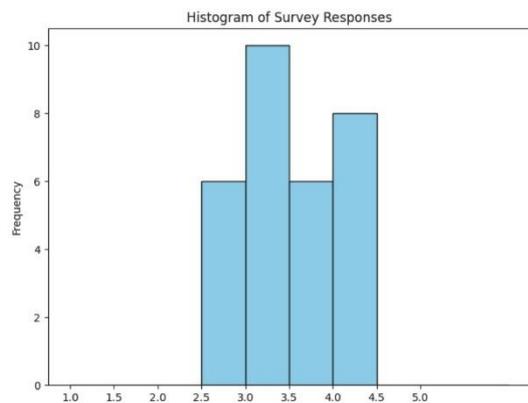


図 3.2: アンケート結果のヒストグラム

3.2 単語の難易度に基づく分かりやすさの評価

料理レシピ内には、一般的な日常会話や文章ではあまり用いられない、専門的な語彙が多数含まれている。たとえば、「湯引き」「ささがき」などの調理特有の用語は、調理に慣れていない初心者にとって意味が理解しづらく、調理の手順を正確に把握する上での障壁となり得る。これらの難解な単語は、レシピの可読性や理解のしやすさに大きな影響を与えるため、文章のわかりやすさを評価する上で重要な要素と考えられる。本研究では、このような単語の難易度に着目し、レシピテキストにおける分かりやすさを定量的に評価する手法を提案する。具体的には、単語の難しさを評価する指標として、大規模言語モデルである BERT[3]を活用する。前述したとおり、BERT は、文中の一部の単語をマスクし、それを文脈から予測する Masked Language Model (MLM) によって事前学習がなされており、文脈に即した語彙の予測性能を持っている。この特性を利用し、対象となる料理レシピのテキストから、名詞を一語ずつ順に[MASK]トークンで置き換えるという処理を行う。たとえば、「じゃがいもを茹でる」という文に対して、「[MASK]を茹でる」という形式に変換し、BERT が[MASK]部分に最も適切と判断する単語を予測するように設定する。その後、BERT が出力する上位 10 件の予測候補の中に、元の単語が含まれていれば、その単語は「正解」

とみなす。図 3.3 は、予測候補に正解の単語が含まれていた例を示しており、BERT が適切な語彙を推定できたケースである。一方、図 3.4 は予測候補の中に正解の単語が含まれず、BERT が予測に失敗したケースを示している。これらの処理を、対象となる料理レシピ内のすべての名詞に対して行い、総当たりに検証を進めた。最終的に、料理レシピごとの正解率を算出した。図 3.5 は正解率のヒストグラムである。この正解率を、該当レシピにおける「単語の難易度に基づく分かりやすさ」として評価指標とする。正解率が高いほど、文脈において一般的な語彙が用いられていることを示しており、結果としてそのレシピが理解しやすいことを意味すると考えられる。なお、日本語の BERT の pre-trained models は図 3.6 の東北大学の乾研究室が公開しているものを使用した。

[MASK]でマーガリンを熱し、両面こんがりと色が着くまで焼きます
 正解: フライパン
 候補: ここ, 鍋, 火, 熱, 工場, これ, フライパン, 窯, 高温, 好み

図 3.3: 正解となる例

包んで[MASK]を整えます。
 正解: 形
 候補: 味, 風味, 生地, 色, 型, バランス, 香り, 状態, 盛り付け, 彩り

図 3.4: 不正解となる例

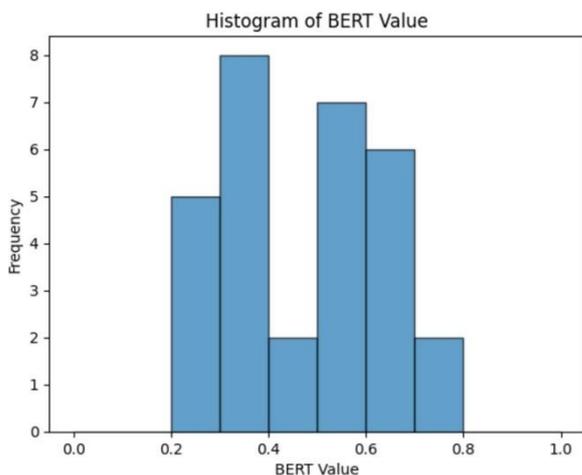


図 3.5: BERT の分かりやすさの評価のヒストグラム

['cl-tohoku/bert-base-japanese-whole-word-masking'](https://github.com/cl-tohoku/bert-base-japanese-whole-word-masking)

図 3.6: 使用した BERT の pre-trained models

3.3 曖昧な表現の数に基づく分かりやすさの評価

料理レシピ内には、レシピ特有の曖昧な表現が多く用いられている。これらの表現は、明確な数値や基準が示されていない場合が多く、人によって解釈が異なる可能性が高い。したがって、読者の理解や調理の再現性に影響を及ぼす要因となり得る。実際に、料理レシピ内の曖昧な表現を補完・明確化することを目的とした研究も行われており[4]、このような曖昧さはレシピの分かりやすさに深く関係していると考えられる。本研究では、レシピに含まれる曖昧な表現の数を指標とし、分かりやすさの定量的評価を試みる。そのためにはまず、料理レシピに現れる曖昧な表現を網羅的に抽出する必要がある。本研究では、独自に曖昧表現の辞書を作成し、曖昧さの自動的な検出を可能にした。曖昧表現の収集にあたっては、実際の料理レシピサイトや、ChatGPT などの言語モデルを活用し、多様な表現を洗い出した。これにより、約 1,000 件の曖昧な表現を収集した。さらに、日本語 WordNet を用いて、これらの表現に対する類義語を検索し、表現のバリエーションを拡張した結果、最終的に約 1,500 件の曖昧な表現を得ることができた。これらの表現は、調理の文脈における多様な曖昧性を反映しており、解析に適したデータとなっている。さらに、収集した曖昧な表現を以下の 4 種類に分けた。

- (1) 量・サイズに関する曖昧さ
調味料の分量や食材の切り方・大きさが具体的に示されておらず、調理者の感覚に委ねられている表現。
例: 「少し入れる」「大きめに切る」
- (2) 時間・タイミングに関する曖昧さ
調理の時間や行程のタイミングが明確でない表現。
例: 「少し待つ」「数分置く」
- (3) 動作に関する曖昧さ
調理の動作が具体的でなく、強さや速さが不明瞭な表現。
例: 「軽く叩く」「さっと混ぜる」
- (4) 状態・質感に関する曖昧さ
調理の仕上がりの状態や触感が主観的で、客観的な基準が不明瞭な表現。
例: 「しんなりするまで」「カリッとするまで」

4. 予備実験

まず、4 種類の曖昧な表現が、それぞれ料理レシピの分かりやすさにどの程度の影響を与えているのかを明らかにするために、これら 4 種類の曖昧な表現の出現数を説明変数として設定し、目的変数としてアンケート調査によって得られた分かりやすさの評価スコアを用いて、重回帰分析を行った。さらに、曖昧な表現のみに加えて、単語の難易度に関する情報も含めた計 5 つの特徴量を説明変数とし、同様に重回帰分析を実施した。この分析の目的は、曖昧さや単語の難易度といった言語的な特徴が、実際に人間の感じる分かりやすさの評価とどの程度関連しているのかを数量的に把握することである。モデルの性能評価には、一個抜き交差検証

(Leave-One-Out Cross Validation) という手法を用い、各データを一つずつ検証用に取り除いて学習・評価を繰り返すことで、過学習を抑えた汎化性能の高い推定を目指した。また、モデルの予測精度を測るための指標として、二乗平均平方根誤差 (Root Mean Square Error, RMSE) を算出し、各モデルの性能を比較した。加えて、比較対象として、関連研究である計算機マニュアルの分かりやすさを定量的に評価した先行研究[1]で用いられている形態的要素、すなわち文の長さやひらがな・漢字の出現頻度などの特徴量を本研究でも説明変数として採用し、同様の手法で重回帰分析を行い、提案手法との性能の差を検証した。これにより、本研究で着目している曖昧さや難易度といった指標が、従来の手法と比べてどの程度分かりやすさの説明に寄与するのかを評価した。

5. 結果

曖昧な表現に関する重回帰分析の結果として、各特徴量に対応する回帰係数の値を表 5.1 に示す。また、重回帰分析を行い、一個抜き交差検証によって RSME を算出した結果を表 5.2 に示す。

表 5.1: 曖昧な表現の回帰係数

曖昧な表現の種類	回帰係数	P 値
量・サイズ	-0.248	0.010
時間・タイミング	-0.192	0.025
動作	-0.022	0.768
状態・質感	0.162	0.162

表 5.2: 重回帰分析の結果

説明変数	RSME
単語の難易度	0.388
単語の難易度 + 曖昧な表現	0.301
形態的要素	0.493

6. 結果の考察

まず、表 1 の結果について考察する。

(1) 量・サイズに関する曖昧さ

有意水準 5% で統計的に有意な負の影響を与えていることが示された。これは、例えば「少し入れる」や「大きめに切る」といった定量的・定規的な基準が明確でない表現が、レシピの実践において混乱を招く可能性を示している。特に、初心者にとっては、具体的な分量や切り

方の情報が不足していると、調理の結果に対する自信を持つことが難しく、再現性が低下する。量やサイズは料理の味や見た目を大きく左右する要素であるため、このカテゴリの曖昧さが分かりやすさを損なう大きな要因となっていると解釈できる。

(2) 時間・タイミングに関する曖昧さ

このカテゴリについても、有意水準 5% で統計的に有意な負の影響が認められた。時間やタイミングに関する表現が不明瞭であると、調理中に「次の工程へ進むべきタイミング」や「適切な加熱時間」などが判断しづらくなり、結果として調理の失敗に繋がる可能性が高くなる。特に、「数分煮る」「しばらく置く」といった曖昧な表現は、材料や調理環境によって意味が異なり得るため、経験の少ないユーザーにとっては判断が難しい。このような曖昧さが分かりやすさを低下させることは直感的にも理解でき、回帰分析の結果とも整合的である。

(3) 動作に関する曖昧さ

回帰係数が小さく、P 値は有意水準 5% を大きく上回っているため、動作に関する曖昧な表現はレシピの分かりやすさに対して統計的に有意な影響を与えていないと判断される。これは、たとえ動作の表現に曖昧さが含まれていたとしても、多くの場合は文脈や一般的な調理経験から意味を補完することが可能であるためと考えられる。たとえば、「さっと炒める」「軽く混ぜる」といった表現は、具体的な時間や強さが示されていないが、読み手は前後の工程や食材の種類からおおよその動作を推測することができる。このように、動作に関する表現の曖昧さは、料理初心者であっても大きな混乱を招きにくく、全体としてレシピの理解に対する影響は限定的であると考えられる。

(4) 状態・質感に関する曖昧さ

最後に、状態や質感に関する曖昧表現については、回帰係数が正の値を示した一方で、P 値は統計的に有意とは言えない値であった。これは、状態や質感の表現が、必ずしもレシピの分かりにくさの原因とはならず、むしろ文脈によっては分かりやすさを補助する役割を果たすことを意味している可能性がある。たとえば「カリッとするまで」や「しんなりするまで」といった表現は、数値では表現しづらい調理の完成状態を直感的に伝えており、視覚や触覚を通じて理解しやすいと感じるユーザーも少なくない。よって、このカテゴリの曖昧さは、他と異なり必ずしもネガティブに作用するとは限らず、状況によってはレシピの理解を助ける効果も持ち得ると考えられる。

次に、表 2 に示された回帰モデルごとの RMSE の比較結果について考察する。RMSE は予測精度を示す指標であり、値が小さいほどモデルによる予測誤差が小さく、

実測値に近い予測ができていないことを意味する。まず、形態的要素（漢字やひらがなの割合、文の長さなど）を説明変数としたモデルの RMSE が最も高くなった点に注目する必要がある。これは、従来の読みやすさ評価で重視されてきた要素が、料理レシピのような実用文における「分かりやすさ」の本質を十分に捉えていない可能性を示唆している。たとえ文体や構成が整っており、文字の種類や比率や文の長さが適切であったとしても、専門的な語彙や曖昧な表現が多く含まれていれば、レシピ全体としての理解のしやすさは大きく損なわれてしまう。そのため、形態的指標のみでは、ユーザーが実際に「わかりやすい」と感じる度合いを的確に評価することは困難であると考えられる。一方で、単語の難易度と曖昧な表現の数を説明変数としたモデルでは、すべてのモデルの中で最も RMSE が低く、予測精度が高い結果が得られた。これは、難解な単語が理解の障壁となり、曖昧な表現が読者によって異なる解釈を生じさせることで、レシピの分かりやすさに強い影響を与えていることを示している。また、この二つの要因はそれぞれ独立に、かつ補完的に作用しており、異なる観点から「理解の難しさ」を構成している。難しい単語は単語レベルでの理解を妨げ、曖昧な表現は文脈依存の読み取りに負荷をかける。これらを同時にモデルに組み込むことで、ユーザーの読みづらさや不安定な解釈のリスクをより精緻に捉えることができ、その結果として分かりやすさの予測精度が向上し、RMSE の低減につながったと考えられる。以上の結果は、料理レシピの分かりやすさを評価するにあたって、形態的要素だけでなく、語彙の難易度や意味的な曖昧さといった、より実践的かつ認知的な観点を取り入れる必要があることを示している。今後、さらに多様な要素を加味したモデルを構築することで、より汎用性が高く、精度の高い分かりやすさ評価が実現できると期待される。

7. 今後の予定

新たな料理レシピの分かりやすさの評価方法として、レシピに記載された手順に基づく評価手法の検討を行っている。料理レシピは一般的に、材料を用いた調理手順が時系列順に並んで記述されており、それらの手順が論理的かつ効率的に構成されていることが、利用者にとっての理解のしやすさ、すなわち分かりやすさに大きな影響を与えられとされる。例えば、無駄な手戻りが生じないように手順が整理されているかどうかは、調理の流れをスムーズに進めるうえで重要な要素である。手順に基づく分かりやすさの評価方法として、二つの方法を検討している。

一つ目の方法は、既存の料理レシピの手順を並び替え、その結果として分かりやすさにどの程度差が生まれるかを調べることである。単純に手順を適当に並び替えれば分かりにくくなることは明白であるため、どのように並び替えるかについては慎重な検討が必要である。料理レシピには、前処理、調理、仕上げ、盛り付けといった大まかな工程があり、それぞれの工程内での手順の並び替えを検討している。

二つ目の方法は、特定の料理（例えばカレーライスなど）に着目し、その料理に関する複数のレシピを収集した上で、各レシピに共通して現れる手順を抽出し、標準的な手順を導出するというものである。この手法ではまず、複数のレシピから頻出する手順パターンを抽出し、出現回数が所定の閾値を超えるものを選定する。次に、頻出パターン間の依存関係に基づいて有向非巡回グラフ (DAG) を構成し、これに対してトポロジカルソートを適用することで、その料理における標準的な調理手順の順序を決定する。トポロジカルソートとは、グラフ理論における手法の一つであり、有向非巡回グラフの各ノード（ここでは手順）を、すべての辺が順序的に整合するような並び順に変換するアルゴリズムである。これによって得られた標準手順は、その料理における「お手本」となる調理順序とみなすことができる。

最後に、各レシピと標準的な順序との間における差異を定量的に評価するため、編集距離や最長共通部分列 (LCS) といった手法を用いて順序のズレを数値化する。このようにして算出された逸脱度に基づき、個別のレシピがどれほど補油準的な流れから外れているかを把握し、それが分かりやすさの評価にどのように関与するかを明らかにすることを目指している。

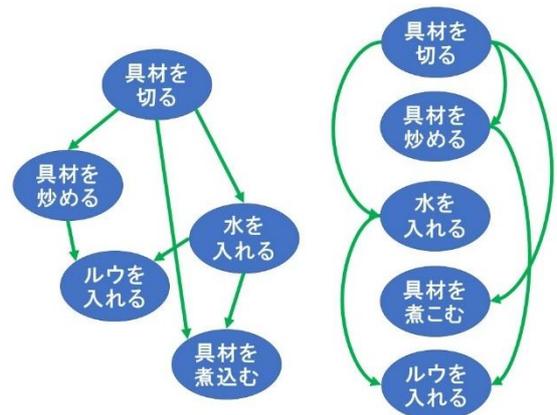


図 7.1: トポロジカルソートの例

参考文献

- [1] 高橋善文・牛島和夫 (1991). 計算機マニュアルの分かりやすさの定量的評価. 情報処理学会論文誌, 460-469.
- [2] 建石由香・小野芳彦・山田尚勇 (1988). 日本文の読みやすさの評価式. 情報処理学会研究報告ヒューマンコンピュータインタラクション (HCI), 1-8.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805.
- [4] 福元颯・松下光範 (2019). 調理手順の類似性に基づく料理レシピの曖昧な表現の補完. 電子情報通信学会 HCG シンポジウム 2019 論文集.