

公立はこだて未来大学 2024 年度 システム情報科学実習
グループ報告書

Future University Hakodate 2024 Systems Information Science Practice
Group Report

プロジェクト名
コスメ×データサイエンス
Project Name
Cosmetics × Data Science

グループ名
クチコミ評価班
Group Name
Review Rating Group

プロジェクト番号/**Project No.**
13

プロジェクトリーダー/**Project Leader**
橋本海羽 Miu Hashimoto

グループリーダー/**Group Leader**
立石達人 Tatsuto Tateishi

グループメンバ/**Group Member**
秋山和奏 Wakana Akiyama
浅田新輝 Shinki Asada
木村安寿 Anju Kimura

指導教員
佐藤直行 富永敦子 島内宏和 花田光彦

Advisor
Naoyuki Sato Atsuko Tominaga Hirokazu Shimauchi Hanada Mitsuhiko

提出日
2025年1月21日
Date of Submission
January 21, 2025

概要

コスメ購入の際、クチコミは非常に参考になる情報源である。しかし、クチコミが本当に有用かどうかは全てを読まないと判断できない。@cosmeのクチコミ機能には「Like」機能があり、参考になったクチコミに対してユーザーがLikeを送ることができる。本研究では、Like数の多いクチコミの有用性を検証し、クチコミの有用性を自動判別するモデルの作成を目的とした。研究では2023年1月～6月のミルククレンジングに関する1087件のクチコミデータを分析対象とし、具体的な使用感や効果の説明、肌質やメイクの状況の明確な記述、PREP法による文章構成などを基準に、クチコミを「有用」「有用でない」「中間」に分類した。

テキスト分析ソフトウェアKH Coderを用いて行ったクロス集計と対応分析の結果、有用なクチコミはLike数との相関が認められ、Like数が有用性判断の指標となることが示された。また、有用なクチコミでは「毛穴」「カバーマーク」「洗浄」「肌荒れ」といった、商品の具体的な使用感や効果を示す単語が頻出することが判明し、これらの単語を含むクチコミは実際の使用体験に基づく情報を提供している可能性が高いことが示唆された。

文脈を考慮した自然言語処理が可能なBERTを用いて、有用なクチコミ200件と有用でないクチコミ200件を学習データとして判別モデルを構築した結果、81.25%の精度を達成した。今後の課題として、「中間」に分類されたクチコミのより詳細な分類や、データの収集・前処理の改善、誤分類事例の分析によるモデルの解釈性向上などが挙げられる。これらの改善により、クチコミの有用性をより高精度に判別できるモデルの開発が期待される。

目次

1. はじめに
 - 1.1. 背景
 - 1.2. 目的
2. 関連研究
 - 2.1. 有用な商品レビューの特徴と抽出について
 - 2.2. レビュー記事からの評判情報抽出の一手法の提案
 - 2.3. BERTを用いた単文の感情極性推定手法の提案とその有効性
3. 手法
 - 3.1. 使用したデータについて
 - 3.1.1. データの概要
 - 3.1.2. データのラベル付けについて
 - 3.2. KH Coderを用いた分析
 - 3.2.1. KH Coderについての説明
 - 3.2.2. KH Coderで行ったデータの前処理
 - 3.2.3. 分析手法（クロス集計と対応分析の説明）
 - 3.3. BERTを用いたモデル作成
 - 3.3.1. BERTについての説明
 - 3.3.2. モデルの作成方法
4. 結果
 - 4.1. KH Coderを用いた分析
 - 4.1.1. クロス集計の結果
 - 4.1.2. 対応分析の結果
 - 4.2. BERTを用いたモデル作成
 - 4.2.1. モデルの精度について
5. 考察
 - 5.1. KH Coderによる分析で得られた有用なクチコミの特徴の考察
 - 5.2. BERTで構築したモデルの考察
6. まとめ
 - 6.1. 全体をまとめた説明
 - 6.2. 今後の課題
7. 参考文献

1. はじめに

1.1. 背景

コスメ購入の際、クチコミは非常に参考になる情報源である。特に、自分と似た肌質やニーズを持つ他のユーザーのクチコミから、商品の実際の使用感や効果について具体的な情報を得ることができる。しかし、クチコミが本当に有用かどうかは全てを読んだうえで判断する必要がある。一部のクチコミには誤解や偏見が含まれていたり、個々の使用状況に大きく依存していたりする。特に商品やブランドを意図的によく見せる目的で投稿されているものは、誤った印象を与える可能性がある。そのため、ユーザーにとって参考にならないものもある。さらに、多くのクチコミを比較する際にはどれが最も信用できる情報かを見極める必要がある。

1.2. 目的

「@cosme」のクチコミには「Like」という欄がある。これは商品を選ぶ際に「気になった」「参考になった」クチコミに対してLikeを送る。[1]

「@cosme」のクチコミは2023年12月21日時点で2000万件を突破するほど、大量にコスメに関するクチコミが書かれている。しかし、クチコミの中でも、Like数が多いものとそうでないものがある。Like数が多いほど、クチコミはユーザーから見て参考になるのだろうか。本研究の目的は二点ある。一点目はLike数が多いクチコミが有用であるか確かめること、二点目はクチコミが有用かどうか判断するモデルを作成することである。

2. 関連研究

クチコミ評価班では、以下の関連研究を参考にした。

2.1. 有用な商品レビューの特徴と抽出について

草刈ら(2018)[2]は、楽天市場商品レビューデータから、役に立たないレビューを除去し、役に立つレビューを抽出することを目的として研究を行った。Health、Game、Computer の3ジャンルの楽天レビューデータについて、商品未購入者のレビューを「役に立たないレビュー」と仮定し、その特徴に当てはまったレビューを取り除いた。その後SVM (サポートベクターマシン)を利用して学習を行い、「役に立たないレビュー」を検出することを試みた。また、ダイエットジャンルの高評価レビューを目視で確認し、内容が褒めてばかりのレビューを役に立たないレビューの学習例として活用した。その結果、未購入者の高評価レビューの正答率

はHealthが0.9518、Gameが0.8839、Computerが0.9520と役に立たないレビューの学習正例として有用であることを示した。また、購入者の高評価レビューの中で褒めてばかりのレビューは、役に立たないレビューの学習例として有望であることが判明した。このことから、有用なレビューを抽出するには、未購入者のレビューや褒めてばかりの高評価レビューは取り除くべきだと分かった。

2.2. レビュー記事からの評判情報抽出の一手法の提案

梅村ら(2010)[3]は、WEB掲示板などの膨大なテキストデータから、製品に関するクチコミを効率的に抽出する手法を研究した。梅村らの手法では、「評価項目」と「評価表現」という2つの要素を自動的に同定し、それらを含む文をクチコミとして抽出することを目的としている。

評価項目とは、ユーザが評価の対象として頻繁に言及する項目を指す。一方、評価表現は、評価項目に対する意見や感想を表現したものである。例えば「A社のカメラの解像度が素晴らしいです」というクチコミでは、「解像度」が評価項目に、「が素晴らしいです」が評価表現に該当する。この手法では、エントロピーに基づき評価項目や評価表現に重みづけを行い、得られた重みをもとに新たな評価項目と評価表現を自動的に獲得する。その後、獲得した要素を用いてクチコミを自動的に抽出する仕組みを構築した。

また、梅村らはSakaiら[4]とPantelら[5]の2つの従来手法と提案手法との比較を行った。Sakaiらの手法では、企業の業績に関する記事から、業績の要因を表す「因果表現」を共通頻出表現と手がかり表現の2つに分割し、少数の手がかり表現からエントロピーに基づいて双方を自動的に抽出する。手がかり表現とは「が好調」「が不振」といった表現のことである。

Pantelらの手法では、少量のシードインスタンスを入力とし、パターンの生成、パターンの評価と選定、インスタンスの抽出を行う。この3段階を繰り返すことにより、意味関係を自動的に抽出する。

従来手法と提案手法の比較では、有意な差は確認されなかったものの、評価項目や評価表現の数を増やすことで、精度や再現率がわずかに向上することが示された。一方で、有意な差が得られなかった要因として、書き手によって評価表現の記述方法が多様であり、同じ内容でも異なる表現が存在することが挙げられる。この影響により、再現率が低下する可能性が示唆されている。

これらの結果から、本グループにおいて有用なクチコミを抽出するためには、類似する意味の評価表現をあらかじめ統一したり、どのような表現が有用であるかを明確に定義することが必要であることが分かった。

2.3. BERTを用いた単文の感情極性推定手法の提案とその有効性

中澤ら(2020)[6]は、ニュース記事の単文から感情極性(ポジティブ・ネガティブ)を推定する新しい手法を提案した。従来の感情極性分析では、Word2Vecを用いた手法が主流であったが、本研究ではBERTを活用した新たなアプローチを試みている。感情極性の推定には、Laboro.AI社が開発した日本語BERTモデルを採用した。このモデルは、Livedoorニュースのコーパスを用いて事前学習されており、ニュース記事の分析に適している。モデルの学習データとして、読売新聞から無作為に抽出した単文500個について、30名の被験者による7段階評価(とてもネガティブ～とてもポジティブ)のアンケートを実施した。本研究では、提案手法の有効性を検証するため、2つの既存手法との比較を行った。1つ目は構文解析を用いた手法で、文の構造から感情極性を推定する。2つ目は形態素解析を用いた手法で、動詞や修飾語の極性値に基づいて推定を行う。比較実験の結果、BERTを用いた手法では、ネガティブな感情をより強く判定する傾向が見られた。この原因として、ポジティブな教師データが相対的に少なかったことが指摘されている。一方、既存手法では、構文解析を用いた手法がポジティブな値を得にくい傾向があり、形態素解析を用いた手法ではポジティブな評価は適切に捉えられたものの、ネガティブ値が過度に大きくなるという課題が明らかになった。これらの結果から、今後の改善点として、教師データの拡充やネガティブデータのダウンサンプリング、重要単語やニュースジャンルといった特徴量の追加、さらには各手法を組み合わせたアンサンブル学習の検討が提案されている。全体として、BERTを用いた感情極性推定の可能性が示されつつも、まだ改善の余地があることが示唆された。

3. 手法

3.1. 使用したデータについて

3.1.1. データの概要

データは@cosmeに投稿された2023年1月～6月の1087件、10月の8件のミルククレンジングのクチコミ計1095件を使用した。使用するデータをミルククレンジングにした理由として、スキンケア商品はテスターが少なく、長期間使用しないと効果が分かりにくいいため、クチコミがより重視されると考えたためである。また、本手法ではクチコミに手動でラベル付けを行い教師データを作成する必要がある。そのため、スキンケア商品の中で比較的クチコミ件数が少なく、クチコミ全件に目を通しやすいミルククレンジングを選択した。

3.1.2 データのラベル付けについて

クチコミが有用かどうかは個人ごとに解釈が異なる。グループ間でその解釈を一致させるため、有用なクチコミの定義を決めることにした。この定義はミルククレンジングのクチコミを班全員で確認した上で、特に有用であると感じたクチコミの特徴や内容を基に策定した。

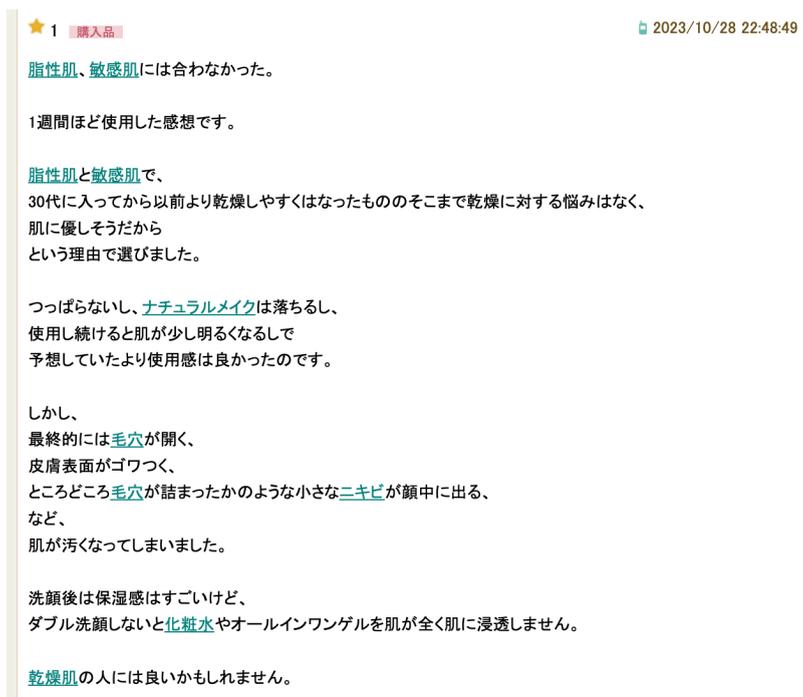


図3.1 特に良いと感じたクチコミ

以下が有用なクチコミの定義となる。

定義一覧

- ・具体的な使用感や効果を詳細に説明している。
- ・自身の肌質やメイクの状況を明確に伝えている。
- ・具体的な事例や比較を用いて説明している。
- ・商品の用途や効果的な使い方を提案している。
- ・具体的な結果や変化を報告している。
- ・他のユーザーへのアドバイスや提案が含まれている。
- ・文章の構成がPREP法(結論→理由→具体例→結論)となっている。

この定義に基づき、定義が当てはまるクチコミを有用、当てはまらないものを有用でない、それ以外を中間とし、分析を行った。分類したクチコミの件数としては有用が200件、有用でないが124

件、中間が771件となった。この分類作業は班の4人で協力して行い、各メンバーがクチコミを確認し、意見を出し合いながら進めた。この分類により、有用とされるクチコミの特徴を明確にし、今後の分析やモデル作成に役立つ基盤を構築した。

3.2. KH Coderを用いた分析

3.2.1. KH Coderについて

KH Coderはテキストデータの内容の分析を行うことができるソフトウェアである。

KH Coderでは、「特徴語」や「語と語の関係」「語と数値等の変数との関係」などを分析することができる。

3.2.2. KH Coderで行ったデータの前処理

KH Coderでは単語の強制抽出や使用しない単語を指定する処理を行った。KH Coderでは単語を自動抽出する際に、「リピート」が「リ」と「ピート」に分かれて抽出されることがある。このように1つの単語が2つ以上に分かれての抽出を防ぐため、強制抽出の機能を用いて「リ」と「ピート」を1つの単語として抽出した。

強制抽出を行った単語としては、「スキンケア」「リピート」「コスパ」「保湿」「化粧水」「マスカラ」「美白」「色落ち」「カバーマーク」の計9単語である。

使用しない単語の指定では、化粧品のカチコミ全般に出現する「使う」「思う」「使用」「感じる」「感じ」「笑」の計6単語を指定し、分析対象から除いた。

※無料版のKH Coderだと、強制抽出と使用しない語をそれぞれ1語までしか指定できない。有料版であればこのような制限がない。

※本研究では有料版のversion3.0.0を使用した。

3.2.3. 分析手法

2023年1月～6月のミルククレンジングのデータを使い、KH Coderを利用してクロス集計と対応分析を行った。

Like数と有用なクチコミの関係を分析するためにクロス集計を行った。有用なクチコミでよく使われている単語を抽出するために、対応分析を行った。クロス集計は、クチコミデータをLike数の有無とラベル付けした得られた「有用、有用ではない、中間」を組み合わせた計6グループに分けた。それぞれのグループに当てはまったクチコミの件数を表でまとめた。その後、6つのグループに対し、カイ二乗検定を行い、Like数との関係があるのか判定した。

対応分析とは、単語間の関係をグラフとして可視化するものである。詳細は4.1で結果の図を用いて説明する。

有用なクチコミに頻繁に出現する単語、有用かどうか関係なく頻繁に出現する単語等を可視化した。この結果から有用なクチコミに頻繁に出現する単語を抽出した。

3.3. BERTを用いたモデル作成

3.3.1 BERTについての説明

BERTは、Googleが開発した自然言語処理のための深層学習モデルである。BERTは、言語理解タスクを実行することができ、文の意味や文脈を理解することができる。BERTの最も大きな特徴は、双方向性を持っていることである。つまり、文の前後を考慮した文脈理解ができる。これにより、高度な自然言語処理タスクを行うことが可能になる。BERTは、文章の要約、文書分類、質問応答、言語翻訳など、様々な自然言語処理タスクに応用されている。BERTは、大規模なコーパス(文章の集まり)から事前に学習されたモデルであり、学習済みのモデルを転移学習することで、より小規模なデータセットでも高い精度で処理することができる。[7]

BERTを用いた理由として、自分たちでモデルを1から作成するより、元々あるモデルを使用する方が良いと判断したためである。

3.3.2. モデルの作成方法

東北大学の自然言語処理グループが公開している日本語BERTモデルの第3版(v3)を使用した。このモデルは、日本語Wikipedia(2023年1月2日時点)とCC-100データセットの日本語部分を用いて事前学習されている。BERT-base相当の構成となっており、具体的にはTransformer Encoderを12層重ね、隠れベクトルの次元数が768、attention headの数が12、中間層の次元数が3072といったパラメータを持つ。これにより、日本語の自然言語処理タスクにおいて高い性能を発揮することが期待できる。

有用なクチコミ200件、有用でないと中間合わせて200件のクチコミを使用してモデルを作成した。まず、クチコミデータ400件を訓練データとテストデータに分割し、分割時と学習時の再現性を確保するために、乱数シードを設定した。

モデルの準備として、東北大学の日本語BERTモデルのトークナイザと、シーケンス分類用のモデルを読み込んだ。モデルのドロップアウト率は0.5に設定した。データセットを扱うために、ReviewDatasetクラスを定義し、テキストデータをトークナイズしてモデルに入力可能な形式に変換した。

データは訓練用と検証用に分割し、それぞれのデータセットをDataLoaderでバッチ化した。モデルとクラスの重みをGPUに移動し、クラスの重みを考慮した損失関数を定義した。オプティマイザとしてAdamWを使用し、線形の学習率スケジューラを設定した。

学習は20エポックで行い、各エポックでモデルを訓練し、検証データで評価を行った。検証精度が向上した場合、モデルを保存し、早期終了の条件も設定した。各エポック後には、検証データの混同行列を作成し、可視化した。最終的に、訓練損失、検証精度、AUCスコアの推移をプロットし、学習の進行状況を視覚的に確認した。

以上より、日本語のクチコミデータに対する高精度なテキスト分類モデルを構築することができた。

4. 結果

4.1. KH Coderを用いた分析

4.1.1. クロス集計の結果

自由度1かつ5%水準のカイ二乗値が5.02である。クロス集計の結果より、有用なクチコミのカイ二乗値が5.156、中間のクチコミのカイ二乗値が0.468、有用ではないクチコミのカイ二乗値が2.726であった。したがって、有用と分類されたクチコミが5%水準でLike数と関係があることが示された。中間・有用ではないと分類されたクチコミは有意な関係が見いだせなかった。

クチコミの有用さとLike数に関するクロス集計の結果を表4.1に示す。

表4.1 クロス集計の結果

	有用	中間	有用でない	ケース数
Like数0	92 (15.23%)	434 (71.85%)	78 (12.91%)	604
Like数1以上	100 (20.70%)	337 (69.77%)	46 (9.52%)	483
合計	192 (17.66%)	771 (70.93%)	124 (11.41%)	1087
カイ二乗値	5.156*	0.468	2.726	

4.1.2. 対応分析の結果

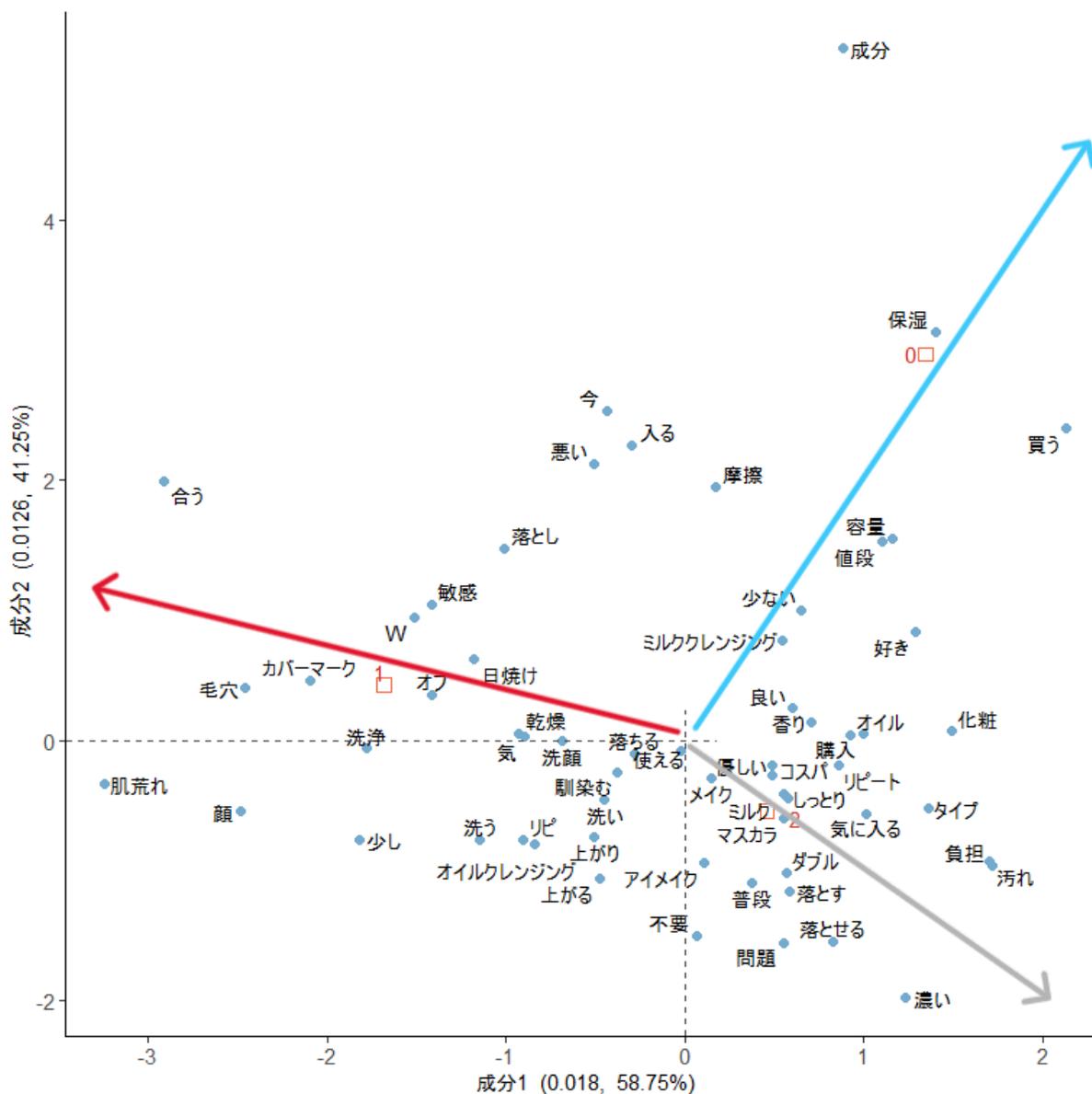


図4.1 対応分析の結果

図4.1中の赤の矢印の方向が、有用と分類されたクチコミにおいて出現しやすい単語が分布する方向である。水色の矢印の方向が、有用ではないと分類されたクチコミにおいて出現しやすい単語が分布する方向である。灰色の矢印の方向が、中間と分類されたクチコミにおいて出現しやすい単語が分布する方向である。原点(0,0)から離れている単語ほど特定の分類のクチコミに偏って出現しやすい。

「合う、毛穴、カバーマーク、洗浄、肌荒れ、W、敏感」は赤の矢印周辺に分布し、原点から離れているため、有用なクチコミの特徴語であることがわかった。

「保湿、買う、容量、値段」は水色の矢印周辺に分布し、原点から離れているため、有用ではないクチコミの特徴語であることがわかった。

4.2. BERTを用いたモデル作成

4.2.1. モデルの精度について

今回のモデルでは、正解率(Accuracy) は約 81% と高い水準を示した。しかし、クラス間のデータ分布やビジネス上の目的によっては、Accuracy だけでは十分にモデル性能を判断できない場合がある。そこで、Precision(約 0.82)、Recall(約 0.81)、F1-score(約 0.81)、AUC(約 0.88)といった複数の指標も計測し、総合的に評価を行った。結果として、全体的にバランス良く高いスコアが得られていることから、本モデルが有用であると判断している。これにより、Accuracy の高さのみならず、誤判定の少なさ(Precision)や見逃しの少なさ(Recall)、バランスの良さ(F1-score)、閾値に対する頑健性(AUC)といった観点からも、本モデルの有用性を裏づけている。

	precision	recall	f1-score	support
ネガティブ	0.86	0.75	0.80	40
ポジティブ	0.78	0.88	0.82	40
accuracy			0.81	80
macro avg	0.82	0.81	0.81	80
weighted avg	0.82	0.81	0.81	80

図4.2 モデルの精度

5. 考察

5.1. KH Coderによる分析で得られた有用なクチコミの特徴の考察

有用なクチコミのみLike数と関係があることから、Like数の有無は有用なクチコミを判定する一つの指標になると考えられる。

有用なクチコミにより頻繁に出現する単語「合う、毛穴、カバーマーク、洗浄、肌荒れ、W、敏感」は、商品の使用感や効果を述べる場面でよく使用されている。例えば、「敏感肌ですが今の所肌荒れ等はしていません。このまま使用して肌荒れしなければずっと使っていこうと思います。」のク

チコミがこれに該当する。このことから、これらの単語がよく出現する有用なクチコミは、実際に商品を使用しないとわからない効果や使用感が述べられている傾向にあると思われる。

本研究の特徴語を抽出する手法は、草刈ら(2018)[2]で用いた高評価レビューの特徴語を抽出する手法よりも、実際に商品を使用しないとわからない効果や使用感などのユーザーが本当にほしい情報が述べられているクチコミの特徴語を抽出しやすくなると思われる。

5.2. BERTで構築したモデルの考察

構築したモデルは日本語のクチコミデータに対して高い分類性能を示しており、実際に活用しても有用であると考えられる。特に、「有用」および「有用ではない」の両クラスでバランスの取れた適合率と再現率を達成している点は、モデルの汎用性を示唆している。しかし、極端に文章が短いクチコミは、内容が有用であっても、有用ではないと判定される傾向が見られた。この傾向が見られた理由として、クチコミを教師データとして使用する際、クチコミに含まれる1文ごとにラベル付けを行うのではなく、1つのクチコミごとにラベル付けを行ったことが考えられる。1文ごとではなく、1つのクチコミごとにラベル付けを行うと、教師データ1つあたりの文章が長くなり、予測の際に文章が短いクチコミは有用ではないと判断されやすくなったと考えられる。また、「文章の構成がPREP法となっている」という定義に従ってラベル付けを行うと、自然と長い文章のクチコミが有用であるとラベル付けされやすくなることも、1つの理由として考えられる。この傾向を改善するためには、1文ずつラベル付けを行うことが有効だと考えられる。

6. まとめ

6.1. 全体をまとめた説明

クチコミ評価班では、Like数が多いクチコミが有用であるか確かめること、クチコミが有用かどうか判定するモデルを作成することを目標に活動を行った。KH Coderでは、Like数が多いクチコミが有用であるか確かめるため、クロス集計と対応分析を用いて分析を行った。クロス集計の結果から、有用なクチコミとLike数との間には関係があることが分かった。また、対応分析の結果から、有用なクチコミには、実際に製品を使用しないと分からない、効果や使用感が書かれる傾向があることが分かった。クチコミが有用かどうか判定するモデルの作成には、BERTを使用した。その結果、有用であるクチコミを判別するモデルの精度は8割以上となった。

6.2. 今後の課題

全体の課題として、有用、有用でない以外のラベル付けをより詳細に行える余地があると考えられる。今回のラベル付けでは中間としたクチコミを、有用寄り、もしくは有用ではない寄りに分けることができれば、ラベル付けを行う際により高い精度でクチコミを分類できると思われる。

BERTを用いたモデル生成は、現状ではまだまだ改善の余地が残されていると言える。検証データにおける制度が81.25%という結果は、決して低い数字ではないが、さらなる性能向上が求められていることは明らかである。

そのためには、まずモデルの学習に用いるデータの質と量を見直す必要があると考えられる。より多くのデータを収集し、多様な表現に対応できるようモデルを訓練することで、制度の工場が期待できる。さらに、前処理の改善も重要である。テキストデータに含まれるノイズや不要な情報を適切に排除し、モデルが学習しやすいデータを整えることで、より効果的に学習を進められる。クチコミに「である調」や「ですます調」などが混在したり、一部絵文字や顔文字が完璧に除去できていなかった可能性が高い。それらを前処理で除去することで、より精度の高いモデルが作成できると考える。また、モデルの解釈性を高めることも重要な課題だ。モデルがどのようなロジックで判断し、結果を導き出しているのかを理解することで、モデルの改善点を明確化することが出来る。さらに、BERTの特性を活かした改善も検討する余地がある。BERTは、文脈を理解する能力に長けたモデルだが、今回はその能力を十分に引き出せなかった可能性が高い。ファインチューニングを行った際、サンプルが長文に偏っていたことで、文章が短いだけで低い評価を入力したテキストに与えていたケースが多く散見されていた。これらの課題を解決することで、BERTによるモデルの精度を向上させ、より信頼性の高いシステムを実現できると考えられる。

7. 参考文献

[1]株式会社アイスタイル.「Like」ボタンとは. @cosme. 2022.

<https://istyle.collasq.com/index.php?action=faq&cat=3&id=57&artlang=ja>, (参照 2025-01-08).

[2]草刈祐子, 林驍, 鈴木孝彦, 伊東栄典, 廣川佐千男. 有用な商品レビューの特徴と抽出について. 情報処理学会研究報告 火の国シンポジウム B3-4. 2018.

[3]梅村真史, 酒井浩之, 増山繁. レビュー記事からの評判情報抽出の一手法の提案. 言語処理学会第16回年次大会発表論文集. 2010, p.166-169.

[4]Patrick Pantel and Marco Pennacchiotti. Espresso: Lever-aging generic patterns for automatically harvesting semantic relations. In Proceedings of the 21st International

Conference on Computational Linguistics and 44th Annual Meeting of the ACL, pp. 113-120, 2006.

[5]Hiroyuki Sakai and Shigeru Masuyama. Cause information extraction from financial articles concerning business performance. IEICE Trans. Information and Systems, Vol. E91-D, No. 4, pp. 959-968, 2008.

[6]中澤政孝, 亀井且右, 前田陽一郎, クーパーエリック. BERTを用いた単文の感情極性推定手法の提案とその有効性. 日本知能情報ファジィ学会 ファジィシステム シンポジウム 講演論文集(FSS2020オンライン). pp. 177-180, 2020.

[7]NTT DOCOMO. BERTとは何か: 自然言語処理モデルにおけるChatGPTとの違いと特徴. AI suite. 2023. <https://aisuite.jp/column/bert/>, (参照 2025-01-08).