

修士学位論文

題 目

アンサンブル学習を利用した
深層言語モデルによる
マルチラベル分類手法の改良

主 査 森 直樹 教授

副 査 藤本 典幸 教授

副 査 黄瀬 浩一 教授

令和 3 年（ 2022 年）度 修了

(No. 2210104025) 楠本 祐暉

大阪府立大学大学院工学研究科
電気・情報系専攻 知能情報工学分野

目次

1	はじめに	1
2	関連研究	2
2.1	マルチラベル分類 (Multi Label Classification, MLC)	2
2.2	Sentiment Analysis (SA)	3
2.3	Aspect Based Sentiment Analysis (ABSA)	4
2.3.1	三浦らが取り組む ABSA	4
3	楽天トラベルレビュー	6
3.1	楽天トラベルレビュー: アスペクト・センチメントタグ付きコーパス . .	6
3.2	データのクレンジング	6
	参考文献	10

図目次

2.1	SA の分析フロー図 (文献 ^[1] Fig 1 より)	3
2.2	アスペクトベースセンチメント分析ネットワークモデル (文献 ^[2] Fig 1 より)	5
3.1	各アスペクトどうしの相関関係	8

表 目 次

3.1 楽天トラベルレビュー：アスペクトセンチメントタグ付きコーパスのデータ具体例	7
3.2 ニュートラルデータと 1 つのカテゴリにポジネガが両方付与されたデータの例	9

1 はじめに

近年では, SNS や EC サイト (electronic commerce site) の発達により日々大量のテキストデータが生成されている. それらを有効に活用するために, さまざまなシーンでテキスト分類タスクの需要が高まっている. テキスト分類タスクとは, 事前に定義されたカテゴリまたはラベルをテキストに割り当てる処理である. 特にビジネスシーンにおいては, 作業の省人化とデータの有効活用を目的として, 各テキストを自動的にカテゴリに分類することと, その精度を向上させることは大きな課題となっている. テキスト分類では 1 つのテキストに 1 つのラベルのみを割り当てる場合もあるが, データを有効活用するためにはテキストには 2 つ以上のラベルを割り当てる必要がある. このようにテキストに複数のラベルを割り当てるタスクはマルチラベル分類 (Multi Label Classification, MLC) と呼ばれ, 自然言語処理 (Natural Language Processing, NLP) の分野では感情分析におけるタスクとして注目を集めている. その中でもアスペクトベースセンチメント分析 (Aspect-Based Sentiment Analysis, ABSA)^[3] は文書から様々な観点を抽出し, その観点到感情極性を付与する研究である. アスペクト情報とは本研究の目的は NLP におけるテキスト分類の問題に従い, MLC の分類精度を向上させることに定めた. 分類対象となるテキストデータはアスペクトベースで作成された日本語文章のカスタマーレビューであり, 単一または複数のラベルが付与されている.

2 関連研究

本章では、本研究の提案手法の研究について述べる。

2.1 マルチラベル分類 (Multi Label Classification, MLC)

機械学習における重要な問題としてマルチラベル分類 (Multi Label Classification, MLC) がある。MLC とは、1 つの対象に対して複数の事前に定義されたカテゴリまたはラベルを割り当てる分類問題である。この他の分類問題として、対象物を 2 値の集合に分類する 2 値分類や、2 値以上の集合に分類するマルチクラス分類があり、これらは MLC と区別してシングルラベル分類 (Single Label Classification, SLC) と呼ばれる。MLC の具体的な応用先として、テキスト分類、感情分類、画像分類などがある。MLC を解く手法については大きく 2 つのアプローチに分けられることが吉村ら^[4]によって報告されている。1 つ目は分類器を拡張することで MLC をそのまま扱う手法であり、決定木のアルゴリズムを MLC に適用できるように拡張した Quinlan^[5]の手法などがある。2 つ目は MLC を SLC に変形させてシングルラベル分類器を用いてラベル推定をする手法であり、その中でも Binary Relevance (BR) 法^[6]と Label Powersets (LP) 法^[7]が主流である。BR 法とは各ラベルごとに「1 対残り (One-vs-Rest)」で学習をして 2 値分類器を作成することで、MLC を SLC に落とし込んで解く手法である。BR 法の関連手法として全クラスから任意の 2 クラスを選択して「1 対 1 (One-vs-One)」の学習をして 2 値分類器を作成するペアワイズ法 (Pairwise Method)^[8]がある。LP 法は候補ラベルとして考えられる全ての組み合わせ酒豪の中から 1 つの組み合わせを選び出すように分類器を学習して、ラベルを予測する手法である。吉田ら^[4]は、BR 法の利点は計算量の少なさである一方で、ラベル間の相関関係を利用した予測が出来ない欠点があると述べている。対して、LP 法はラベル間の相関関係を利用した予測が出来る利点があるが、計算量がラベル数に対して指数関数出来に増大する欠点がある。つまり、BR 法と LP 法の間にはトレードオフ関係があるとされている。3 章に示す図より、ほとんどのアスペクトどうしの相関は強くないことが確認できる。したがって本研究では相関関係を考慮しない BR 法と、その関連手法であるペアワ

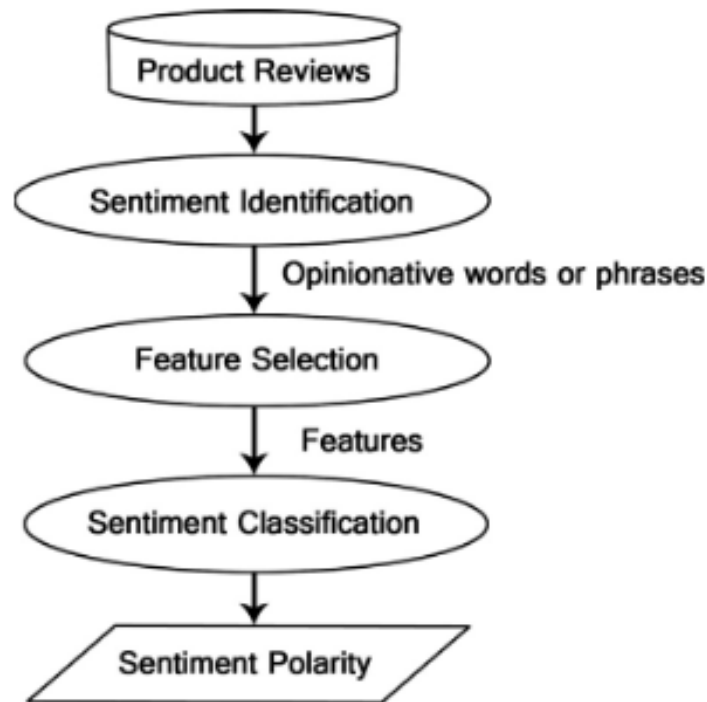


図 2.1: SA の分析フロー図 (文献^[1] Fig 1 より)

イズ法を採択した.

2.2 Sentiment Analysis (SA)

Sentiment Analysis (SA)^[1] は、あるエンティティに対する人々の意見、態度、感情を計算機的に研究する分野である. Liu らによるとエンティティ^[9]とは、文章中における製品, サービス, 組織, 個人, 問題, イベント, トピックなどを指しており, SA における文書分析には重要な指標になるとされている. 図 1 に Medhat ら^[1] による, 商品レビューの場合での SA の一般的な分析フロー図を示す. SA の分析フローは, 生成されたレビュー文章の感情識別をして, エンティティとなる言葉やフレーズに関連した特徴量を選択する. そして, それらの特徴量から文章がどのような感情を表しているかを特定して, その極性を分類することが目標となる. Medhat らによると, 例では商品レビューが採択されているが, SA は製品レビューだけでなく、株式市場^{[10][11]}、ニュース記事^[12]、政治討論^[13]にも適用することが報告されている.

2.3 Aspect Based Sentiment Analysis (ABSA)

Aspect-Based Sentiment Analysis (ABSA)^[3] は近年の SA においてアスペクト情報に基づいた感情分析をする研究分野である. Liu^[3] は従来の SA におけるテキスト分類では、文書のエンティティに関する感情の詳細が得られないという問題があったため、アスペクト情報を利用する必要があると提唱した. アスペクト情報とは文章におけるアトリビュート（肯定的または否定的）と文章内容のエンティティで構成される. アスペクト情報を利用するということは、言語構造（文書、段落、文、節、句）を見るのではなく、フレーズとそれに関する感情を利用するということである.

アスペクト情報に基づいて文章を分析する例を示す. 例えば、「サービスはあまり良くないが、私はこのレストランが大好きだ」という文章がある. この文章はレストランというエンティティについては肯定的だが、サービスというエンティティについては否定的なのである. このように全体としてはやや肯定的と感じる文章でも、アスペクトベースに分析をすることでより詳細な解釈が可能となる.

ASBA における重要な研究ワークショップとして、SemEval (Semantic Evaluation)¹ がある. SemEval は NLP の国際的な研究ワークショップで、意味解析の現状を改善し、自然言語の意味論においてますます困難になっている様々な問題において、高品質のアノテーションデータセットの作成を支援することを使命としている. ABSA はテキスト分類におけるアプローチの必要性が注目を集め、2014 年の SemEval2014^[14] で初めて導入されたタスクであり、その後も継続的にさまざまなアスペクトベースのデータセットによる識別性能の評価がなされている.

2.3.1 三浦らが取り組む ABSA

本研究では三浦らのモデルの構造に倣って MLC モデルを構築した. そこで、?? 章で詳述する本研究の新規性と比較するために三浦らのモデル概要図を示して研究の概要とデータの流れを説明する. 図 2.2 に先行研究である三浦

¹<https://semeval.github.io/>

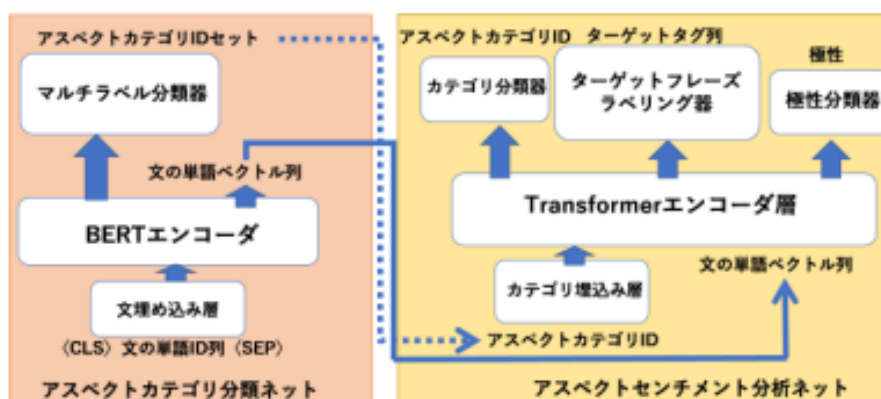


図 2.2: アスペクトベースセンチメント分析ネットワークモデル (文献^[2] Fig 1 より)

らが提案したアスペクトベースセンチメント分析ネットワークモデル^[2]を示す。図 2.2 のアスペクトカテゴリ分類ネットでは、文埋め込み層で文書トークンを生成して、BERT エンコーダでそれぞれのトークンの分散表現を得る。そして、それらの分散表現の先頭トークンである CLS トークンをマルチラベル分類器で推定ラベルを得る。アスペクトセンチメント分析ネットでは、アスペクトカテゴリ分類ネットで得た分散表現とマルチラベル分類結果から推定ラベルと関連度が高いフレーズ（ターゲットフレーズ）や極性を分類する。

このモデルは SemEval2016^[15] の「SemEval-2016 Task 5 : Aspect Based Sentiment Analysis」で公開されたレストランのレビューから成るレストランドメインデータセット²を用いて、MLC やターゲットフレーズと呼ばれるアスペクト情報の抽出をする。注目すべきは、マルチラベルデータに対する識別性能の高さであり、従来手法と比較して F1 値は 8.5% 向上していることが報告されている。

また、三浦らは^[16] は日本語データセットを用いた SA における MLC に取り組んでいる。この研究では 2018 年に TIS インテックグループの TIS 株式会社によって機械学習で SA をするためのものとして公開された「chABSA-dataset」³を用いている。MLC モデルの識別精度は約 70% であり、区別が難しいアスペクトが含まれているため精度が高くならなかったと報告されている。

²<https://github.com/NUSTM/ACOS/tree/main/data/Restaurant-ACOS>

³<https://github.com/chakki-works/chABSA-dataset>

3 楽天トラベルレビュー

3.1 楽天トラベルレビュー: アスペクト・センチメントタグ付きコーパス

楽天グループ株式会社が事業等を通じて取得したデータから構築して、国立情報学研究所を通じて研究者に提供しているデータセット「楽天データセット（コレクション）」^[17]に含まれている「楽天トラベルレビュー: アスペクト・センチメントタグ付きコーパス」^[18]がある。表 3.1 にデータの具体例を示す。全 76623 件の楽天トラベルの日本語レビュー文章とそれぞれの文章について、「朝食, 夕食, 風呂, サービス, 施設, 立地, 部屋」の 7 項目のカテゴリが付与されており、それぞれに感情極性としてポジティブまたはネガティブのタグが付与されている。つまり、全てのデータに付与されるクラス数は 14 であり、該当するクラスには 1 が、該当しないクラスには 0 が付与されている。このデータセットを ASBA の側面から説明する。例として、「モーニングセットは割安だと感じました。」といった文章があった場合, ”モーニングセット” というエンティティは「朝食」というアスペクトに属し、エンティティに対する”割安”というアトリビュートはポジティブな感情極性を持つ。したがって、この文章は朝食に関してポジティブな内容であると解釈できる。このような観点で各テキストに対して 1 または 0 のラベルが 14 個付与されている。14 クラスのうち、ラベル 1 が 1 つしか付与されていないデータをシングルラベルデータとする。また、14 クラスのうち、ラベル 1 が 2 つ以上付与されているデータをマルチラベルデータとする。

3.2 データのクレンジング

全 76623 のデータから、ノイズを含むデータや日本語以外の言語で書かれたデータと、文章のトークン数が 101 以上のデータを除外した。この処理によってデータ数は 73461 件となった。この処理後のデータセットから更に 2 種類のデータを除外した。それら 2 種類のデータは 14 のいずれのカテゴリにも属さないニュートラルデータ 23428 件と、1 つのカテゴリにポジティブとネガティブの両方が付与されたポジネガ両ラベルデータ 1858 件である。表 3.2 にそれらのデータの具体例を示す。ニュートラルデータを除外した理由を示

表 3.1: 楽天トラベルレビュー：アスペクトセンチメントタグ付きコーパスのデータ具体例

		マルチラベルデータ	シングルラベルデータ
文章		部屋も綺麗で, 対応もよく, 朝食もおいしいのでとても満足しています.	一部の方が指摘した通り, 廊下がタバコ臭いのが気になりました.
朝食	ポジティブ	1	0
	ネガティブ	0	0
夕食	ポジティブ	0	0
	ネガティブ	0	0
風呂	ポジティブ	0	0
	ネガティブ	0	0
サービス	ポジティブ	1	0
	ネガティブ	0	1
立地	ポジティブ	0	0
	ネガティブ	0	0
施設	ポジティブ	0	0
	ネガティブ	0	0
部屋	ポジティブ	1	0
	ネガティブ	0	0

す. ABSA は文章のエンティティとアトリビュートによって構成されるアスペクト情報を利用する SA であるためである. 表 3.2 に示すニュートラルデータの文章のエンティティは「風呂, 客室 (施設, 部屋)」であるが, それらにアトリビュート (肯定的: ポジティブ, または否定的: ネガティブ) のような感情極性は含まれていない. 次に表 3.2 に示すポジネガ両ラベルデータを除外した理由を示す. 最大の理由はデータ件数の少なさである. また, 後に紹介するアトリビュートに関する 2 値分類器の作成時において, ポジネガが両方付与されているデータは扱いが難しかったためである. 最終的にモデルの学習や評価に使用するデータ数は 48353 件となった. このデータのシングルデータ数

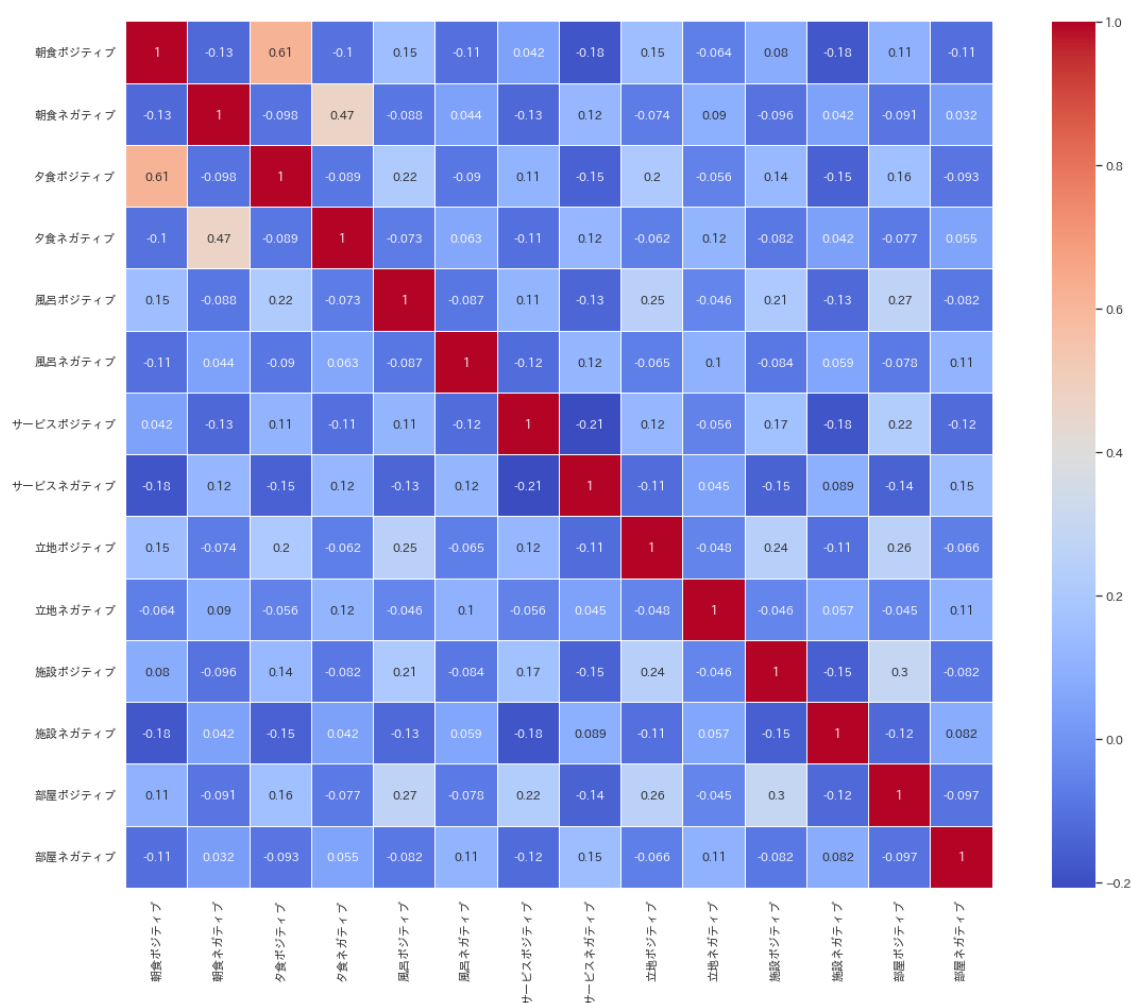


図 3.1: 各アスペクトどうしの相関関係

は 31731 件であり, マルチラベルデータ数は 16622 件である. 図 3.1 に各アスペクトどうしの相関関係を示す. ほとんどのアスペクトは互いの相関が強いことが確認できる. 一方で「朝食ポジティブと夕食ポジティブ」の相関は 0.61, 「朝食ネガティブと夕食ネガティブ」の相関は 0.47 と, 他のアスペクトに比べて強いことが確認できる. しかし全体的な相関は強くないため, アスペクト間の相関関係をモデルの学習に利用していない.

表 3.3 にデータのテキストのトークンサイズの平均と, 正例負例の数と, 各カテゴリに

表 3.2: ニュートラルデータと 1 つのカテゴリにポジネガが両方付与されたデータの例

		ニュートラルラベルデータ	ポジネガ両ラベルデータ
文章		展望風呂付客室に 宿泊しました.	朝食の品数がそこまで多くない ですが味と素材が大満足です.
朝食	ポジティブ	0	1
	ネガティブ	0	1
夕食	ポジティブ	0	0
	ネガティブ	0	0
風呂	ポジティブ	0	0
	ネガティブ	0	0
サービス	ポジティブ	0	0
	ネガティブ	0	0
立地	ポジティブ	0	0
	ネガティブ	0	0
施設	ポジティブ	0	0
	ネガティブ	0	0
部屋	ポジティブ	0	0
	ネガティブ	0	0

参考文献

- [1] W. Medhat, A. Hassan, and H. Korashy. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4):1093–1113, 2014.
- [2] 三浦義栄, 渥美雅保. 事前学習言語モデルを用いたアスペクトベースセンチメント分析ニューラルネットワーク. 人工知能学会全国大会論文集 第35回 (2021), pp. 2Yin507–2Yin507. 一般社団法人 人工知能学会, 2021.
- [3] B. Liu. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167, 2012.
- [4] 吉村皐亮, 馬場雪乃, 鹿島久嗣. 相互依存モデルによるマルチラベル分類. 人工知能学会全国大会論文集 第31回 (2017), pp. 2K21in1–2K21in1. 一般社団法人 人工知能学会, 2017.
- [5] S. L. Salzberg. C4. 5: Programs for machine learning by j. ross quinlan. morgan kaufmann publishers, inc., 1993, 1994.
- [6] M.-L. Zhang, Y.-K. Li, X.-Y. Liu, and X. Geng. Binary relevance for multi-label learning: an overview. *Frontiers of Computer Science*, 12:191–202, 2018.
- [7] G. Tsoumakas and I. Katakis. Multi-label classification: An overview. *International Journal of Data Warehousing and Mining (IJDWM)*, 3(3):1–13, 2007.
- [8] B. Lauser and A. Hotho. Automatic multi-label subject indexing in a multi-lingual environment. In *Research and Advanced Technology for Digital Libraries: 7th European Conference, ECDL 2003 Trondheim, Norway, August 17-22, 2003 Proceedings* 7, pp. 140–151. Springer, 2003.
- [9] B. Liu. Sentiment analysis: Mining opinions, sentiments, and emotions. cambridge: Cup. *Go to original source*, 2015.

- [10] L.-C. Yu, J.-L. Wu, P.-C. Chang, and H.-S. Chu. Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news. *Knowledge-Based Systems*, 41:89–97, 2013.
- [11] M. Hagenau, M. Liebmann, and D. Neumann. Automated news reading: Stock price prediction based on financial news using context-capturing features. *Decision support systems*, 55(3):685–697, 2013.
- [12] T. Xu, Q. Peng, and Y. Cheng. Identifying the semantic orientation of terms using s-hal for sentiment analysis. *Knowledge-Based Systems*, 35:279–289, 2012.
- [13] I. Maks and P. Vossen. A lexicon model for deep sentiment analysis and opinion mining applications. *Decision support systems*, 53(4):680–688, 2012.
- [14] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar. SemEval-2014 task 4: Aspect based sentiment analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pp. 27–35, Dublin, Ireland, Aug. 2014. Association for Computational Linguistics.
- [15] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. AL-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, V. Hoste, M. Apidianaki, X. Tannier, N. Loukachevitch, E. Kotelnikov, N. Bel, S. M. Jiménez-Zafra, and G. Eryiğit. SemEval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pp. 19–30, San Diego, California, June 2016. Association for Computational Linguistics.
- [16] 三浦義栄, 赤井龍一, 渥美雅保. 文中の複数アスペクトのセンチメント分析のための自己注意ニューラルネットワーク. 人工知能学会全国大会論文集 第 34 回 (2020), pp. 3Rin441–3Rin441. 一般社団法人 人工知能学会, 2020.

- [17] 楽天グループ株式会社. 楽天データセット（コレクション）, aug 2010.
- [18] 楽天グループ株式会社. 楽天トラベルレビュー: アспект・センチメントタグ付きコーパス, feb 2021.