

放送コンテンツに対するツイートの＜一様率＞分析

太田 暢[†] 若宮 翔子[†] 荒牧 英治[†]

[†] 奈良先端科学技術大学院大学 〒630-0192 奈良県生駒市高山町 8916-5

E-mail: †{ota.itaru.og4,wakamiya,aramaki}@is.naist.jp

あらまし テレビ番組やラジオ番組などの放送コンテンツの評価指標は、調査会社によって集計される視聴率や聴取率が従来より一般的である。しかし、近年ではインターネット上における感想や実況を分析して評価に役立てる研究が行われたり、インターネット配信プラットフォームにおける視聴数が番組制作者によって重視されるなど、多角化が見られる。その延長線上にある本研究では、新たな評価手法として Twitter におけるツイートの言語的特徴が番組内でどの様に推移していくか可視化した。放送コンテンツに対するツイートの類似度の代表値として平均値を＜一様率＞として定義し、ラジオ番組の面白さとツイートの一様率との相関関係を検証した。

キーワード ソーシャルメディア, Twitter, テレビ, ラジオ, 自然言語処理, 可視化

1 はじめに

放送コンテンツの面白さの評価指標のうち、最も有名なものとして、テレビ番組では視聴率、ラジオ番組では聴取率が挙げられる。視聴率は、地上波などのテレビ放送をリアルタイムで視聴している世帯または個人の割合であり、ランダムに抽出された世帯に設置された測定機によって1分ごとに計測される。¹ 聴取率は、インターネットリサーチパネルより無作為にメールを送信された調査対象者による1週間の日記式調査によって計測される。² 日本においては、調査会社であるビデオリサーチ社が集計したこれらのデータが多くのテレビ局やラジオ局で用いられており、視聴率1%あたり関東地方では約40万人が視聴する計算となる。テレビやラジオ全体の視聴率・聴取率の低下が叫ばれている中でも、10万人単位で情報を届けられるマスメディアの影響力は大きく、そのマスメディアが重視する指標の1つである視聴率や聴取率の数値は放送局やスポンサー企業から重視されている。

一方で、インターネットやスマートフォンの普及により視聴環境や形態が変化した近年では、視聴率や聴取率以外にも番組評価指標が多角化している。例えば、テレビ番組の配信プラットフォームであるTVerやラジオ番組のradiko、また各放送局が独自に展開している配信サービスなどにおける再生数や視聴者属性の分析が行われている。またソーシャルメディアが普及した2010年代からは、番組制作者がソーシャルメディアを通じて番組の宣伝をしたり、番組と連動してキャンペーンを行うことも定着している。テレビとその他のメディアを取り巻く環境について特集したNHKの番組³では、ソーシャルメディ

アで一定の評価を受けていても、視聴率には全く反映されないといったねじれ現象が見られることもしばしばあり、営利性も考慮しなければならない番組の制作にあたって、ソーシャルメディアにおける評価を偏重することはリスクにもなることも指摘されている。こうした視聴率や配信プラットフォームにおける指標は、主に番組の作り手である放送局や制作会社、広告会社やスポンサー企業間で共有される。そのため、一般の視聴者がリアルタイムにこれらの指標情報を取得して視聴行動に反映することは難しく、視聴者は今後もソーシャルメディアの投稿を評価基準の一つとして重視することが想定される。

このように放送コンテンツとソーシャルメディアの関係性が高まっているが、番組の作り手が評価に用いる指標以外に視聴者が入手できる指標の数が少なく、入手も難しいという情報の非対称性が存在している。さらに、放送コンテンツとソーシャルメディアの関係性に注目した研究は数多く行われているにも関わらず、時系列における番組の面白さを図る指標として言語的な共通性や一様性を利用する研究はない。そこで本稿では放送コンテンツのソーシャルメディア関連投稿に対して言語的特徴を時系列的に分析し、別視点での指標として定義する。具体的には、番組放送中のTwitterのツイートを1分間ごとに分析して視聴者の意見の一致度を＜一様率＞として定義し、番組放送中における変化の可視化を試みた(図1)。

こうした指標の応用先として、広告主のための評価指標としてよりも、ユーザの視聴体験の向上を目的として使用する、ツイートデータを倍速視聴の速度の最適化に用いる研究[1]などが考えられる。そこで、ラジオ番組の面白さとツイートの一様率との相関関係を検証した。

2 関連研究

2.1 ソーシャルメディアと放送コンテンツの関係性

放送コンテンツにはネット配信コンテンツのコメント欄のような感想の共有方法はないが、ソーシャルメディアの関連投稿

1: 『視聴率』ビデオリサーチ

<https://www.videor.co.jp/service/media-data/tvrating.html>

2: 『ラジオ個人聴取率調査』ビデオリサーチ

<https://www.videor.co.jp/service/media-data/rdrating.html>

3: 『あたらしいテレビ』2023年1月1日放送 NHK

<https://www.nhk.jp/p/ts/R32JGQLRW6/>

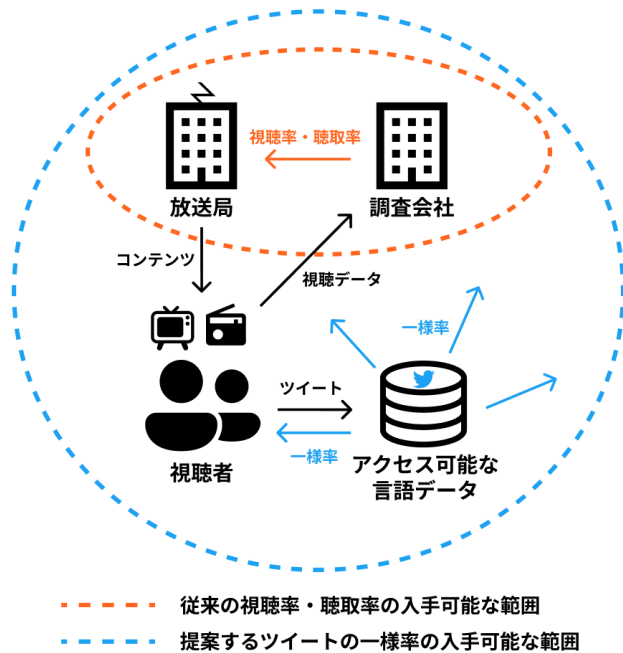


図 1 従来手法と、提案手法である一様率の入手可能な範囲。

において同様の傾向が見られることを利用して、感想や高評価数、拡散数などのデータを用いた分析や評価が試みられている。具体例として、Twitter における投稿を用いて視聴率を評価する研究 [2] や、スポーツ番組のネタバレ検知についての研究 [3] がある。一方でソーシャルメディアとラジオ番組の関係については、Bonini ら [4] の、ラジオ局による Twitter を用いたリスナーエンゲージメント向上施策についての調査に限られる。

2.2 文章中の共感の検出

視聴行為において、重要な要素の一つとして共感があると言われる [5]。共感とは他人の考えや感情を感じ、理解する能力である [6]。視聴者がソーシャルメディアを通じて番組の情報収集を行う際、トレンドランキングにおける関連単語の順位や投稿数、それに付随して表示される複数の代表的な関連用語を通じて事象や他人の考えの理解に繋げていることが想定される。共感を理解するためには自然言語から共感を検出するモデルや、モデルの効果的な学習のための質の高いラベル付きデータが必要となるが、アノテーションコストが大規模になる。Hosseini ら [7] は、感情やセンチメントに関する利用可能な関連リソースから知識を抽出することで、テキストベースの共感度分類のタスクに関する既存のデータセットを用いた共感分類を効果的に利用できることを示した。

本稿では基準を共感まで拡大することによるコストを考慮し、学習に必要なモデルの提示ではなく、汎用性が高い教師なし学習により共通点を提示する。

2.3 文章類似度の算出

本研究では、ソーシャルメディアテキストの文章類似度を測定する。コサイン類似度は 2 つのベクトル間の角度を比較す

る尺度で、単語埋め込みで表現された文の類似度の比較に用いることができる。ユークリッド距離は 2 つのベクトル間の距離を測定する尺度で、単語埋め込みの文の類似度の比較に利用できる。Word Mover's Distance (WMD) [8] は教師なし手法で、ある文の単語を他の文に移動するのに必要な最小限の「作業量」を求めることにより、2 文の非類似度を測定する方法である。Recurrent Neural Networks (RNN) や Bidirectional Encoder Representations from Transformers (BERT) [9] などの Transformer モデルは、ラベル付きデータを用いて文の埋め込みを生成するために学習することができ、その後、余弦類似度を用いて文の類似性を比較することができる。

2.4 文章分散表現の獲得

本研究では、ソーシャルメディアテキストのベクトル化 (文の埋め込み) を行う。文の埋め込みには、Word2Vec, Doc2Vec, SentenceBERT など多くの手法があり、これらを用いて文の埋め込み表現を獲得し、コサイン類似度などの尺度を用いることにより文の類似度が得られる。Word2Vec [10] は、ニューラルネットワークベースで単語の分散表現を学習するモデルであり、2 層のニューラルネットワークとスキップグラムモデルを用いて表現を学習する。学習された単語埋め込みは言語モデリングや構文解析などの自然言語処理タスクの特徴として利用することができる。Doc2Vec [11] は、Word2Vec を拡張したモデルであり、段落や文書の分散表現を学習する。学習時に追加パラメータを設定することで、同じ綴りで異なる意味を持つ単語を異なる文書間で曖昧性をなくすことができる。SentenceBERT [12] は、BERT モデルを改良したものであり、大規模なテキストデータのコーパスで事前に学習され、特定のタスクのために微調整して学習させ文レベルの埋め込みを生成する。

2.5 ツイートの言語指標

ツイートは自然言語の文章や単語であり、これらの内容を言語処理解析することにより、番組の面白さの推定が可能であると考えられる。ツイートの言語指標として、センチメント、トピック、意味的類似度などが挙げられる。

センチメントやトピックは個人のツイート単位で測る指標として用いることが可能であり、ツイートのセンチメント分析により、番組の内容がポジティブまたはネガティブであることを場面の特定に使用することが出来る。Latent Dirichlet Allocation (LDA) [13] や Latent Dirichlet Allocation (LDA) [14] などのトピックモデリングも有効である。これにより番組で取り上げられているトピックを特定し、その時間的推移を通じて番組に対する認識の時間的経過を観測することや、異なる番組間の類似性を知ることが可能である。

意味的類似度は複数のツイート単位で測る指標として用いることが可能であり、ツイート間の意味的類似度が高い場合、番組の視聴者が同様の感情を抱いている可能性が高い。一方、ツイート間の意味的類似度が低い場合には話題が散漫になっていたり、議論を巻き起こしている可能性が高い。ただし、少量のテキストデータやノイズの多いデータを扱う際にはバイアスの

影響を受け得る。

本稿では、番組に対する実況や感想のツイートの意味的類似度を言語指標として用いた。類似度の測定にはコサイン類似度を、ツイートの分散表現獲得にはコストを最低限に留めるため、教師なしで使うことが出来る Word2Vec を用いた。

3 ツイートの類似度の算出方法

テレビまたはラジオ番組の放送中に Twitter に投稿されたツイートを 1 分間ごとに分析し、埋め込み表現の類似度を算出した。ツイートの取得には Twitter API for Academic Research を用いた。まず、テレビまたはラジオ番組の放送時間中のツイートを番組公式ハッシュタグ (#) をクエリとして取得し、リツイートを除いた。取得したツイートデータを、以下の 4 ステップで処理した。2022 年 12 月 18 日のある時間帯の #annkw を含むツイートをを用いた類似度算出処理を図 2 に示す。

STEP 1. 文字の置換・削除 前処理として、各ツイートテキストから、ハッシュタグ (#以降)、リプライ先 (@以降)、空白文字、絵文字、記号、URL を削除する。数字は小数点とカンマを消して 0 に置換し、英語は小文字に統一する。

STEP 2. 分散表現獲得 Word2Vec を用いて各ツイートの分散表現を獲得する。本稿では、ツイートの分散表現獲得の精度向上のための Word2Vec モデルのファインチューニングは行わず、東北大学乾研究室が公開している日本語 Wikipedia エンティティベクトル⁴を用いた。ツイートの分かち書きには、オープンソース形態素解析エンジン MeCab⁵ および mecab-ipadic-NEologd 辞書 [15] を用いた。トークンには、名詞、動詞、形容詞、副詞を使用した。また、ストップワードは設定しなかった。

STEP 3. リスト化 ツイートの投稿時間をもとに、ツイートの分散表現を 1 分間ごとのリストにまとめる。

STEP 4. コサイン類似度算出 各分のリスト中の分散表現ペアのコサイン類似度を総当たりで算出する。ある t 分から 1 分間の類似度の総数 N_{S_t} は、ツイートのリストの長さを n_t としたとき式 (1) のようになる。

$$N_{S_t} = \frac{n_t(n_t - 1)}{2} \quad (1)$$

4 ツイートの類似度の推移傾向

ある番組放送中の類似度の代表値の推移と、類似度の標準偏差とツイート数の推移例を図 3 に示す。類似度の代表値である中央値と平均値の推移には、大きな差は見られなかった。複数の番組の推移から、以下のような傾向が得られた。

類似度の標準偏差が低い (図 3 の上方向) ケースではツイ



図 2 2022 年 12 月 18 日 2 時 33 分に投稿された #annkw を含むツイートの類似度算出処理。

トに含まれている単語が特有のイベントやコーナー名、登場人物を指す名詞を含む場合が多かった。ツイートが少ない場合では特に類似度の標準偏差が低くなる傾向にあった。類似度の標準偏差が高い (図 3 の下方向) ケースでは、ツイート数が増える場面で内容に散らばりが大きくなっており、唐突なイベントが発生して感想が多様になる場合が多かった。例として、放送開始後 53 分のツイートのワードクラウドは図 3 の右側の様になっており、パーソナリティの個人的な報告に対する反応となっていた。注目すべき点として、番組開始 0~2 分後の時間帯にツイートが最も多くなる傾向にあり、データを確認すると多くがハッシュタグのみのツイートであった。これはフォロワーへの番組視聴の意思表示として慣習化している行為であり、ハッシュタグを削除する本稿の処理では類似度計算においては、残りの文章の分散表現のみが類似度計算に用いられている。そのため、番組開始直後には標準偏差は他の箇所比べて高くなる傾向にあった。

これらの傾向を参考に、ツイートの類似度の代表値として平均値を <一様率> として定義する。

5 番組の面白さスコアとツイートの一様率の相関検証

ツイートの類似度の代表値として平均値を <一様率> として定義し、ラジオ番組の面白さと一様率にどのような相関があるかを調査するために実験を行った。

5.1 設定

20 代の男女 4 名の被験者がラジオ番組の音源を聴取し、主観的な面白さを 1 分ごとに評価した。音源には放送時間帯および

4: 日本語 Wikipedia エンティティベクトル
http://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki_vector/
5: MeCab: Yet Another Part-of-Speech and Morphological Analyzer
<http://taku910.github.io/mecab/>

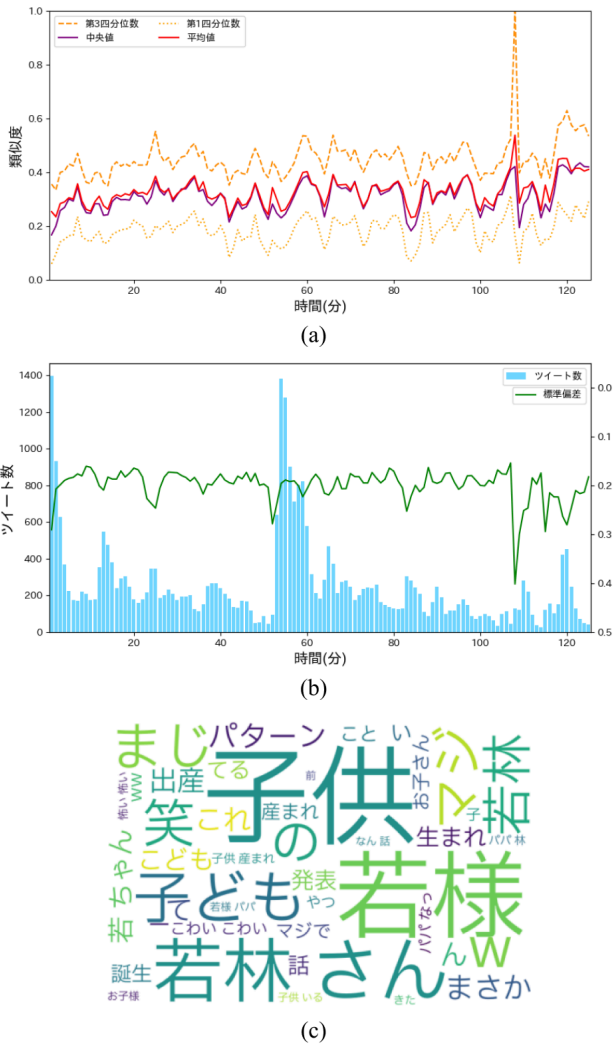


図3 2022年2月6日の#annkwを含むツイートの類似度の代表値推移 (a), 類似度の標準偏差とツイート数の推移 (b), 放送開始後53分のツイートのワードクラウド (c).

radikoにおける番組ジャンルを参考にし、首都圏で放送されている3番組のそれぞれ一部分を用いた(表1). 評価には「1分あたりの音源が面白いと思ったか」という質問に対する「全くそう思わない」「そう思わない」「そう思う」「非常にそう思う」の4段階のリッカート尺度(1~4)を用いて面白さスコアを集計した. 番組が公式に周知しているハッシュタグをクエリとして、検証音源として使用した時間とその後4分間のリツイートを除く全ツイートを取得した. 全ツイート数はann220116が1981, tama220120が507, trad220124が455であった.

5.2 結果

各音源における面白さスコア平均値と、ツイートの一様率の推移を図4に示す. 著者ら[1]が行ったラジオ番組とツイート行為に関する研究結果である、「ラジオ番組の面白さがツイート数の盛り上がり」に1~2分かけて反映される」という原則を考慮して、面白さスコア平均値と放送から0~4分後の一様率との相関係数を求めた(表2). ann220116では+4分との相関係数で0.5を上回り, tama220120では+3分との相関係数で

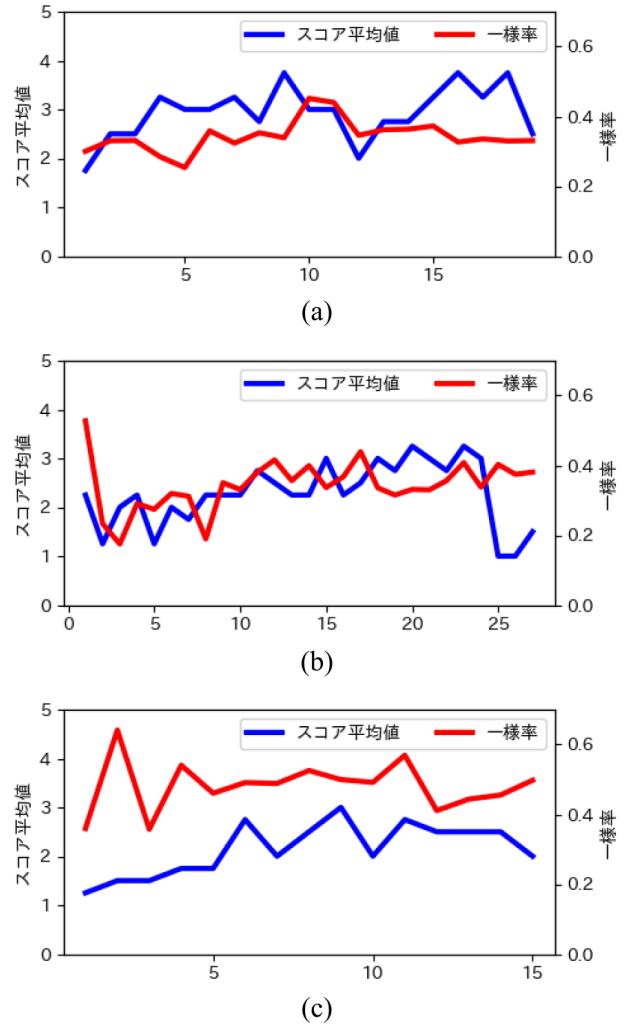


図4 ann220116 (a), tama220120 (b), trad220124 (c)の面白さスコア平均値と一様率の推移.

0.4を上回り, trad220124では+2分との相関係数で0.4を上回るなど, かなり正の相関関係が見られた. このことから, ラジオ番組の面白さ平均値は, 約3分かけて一様率に反映されることがわかった.

6 おわりに

本稿ではテレビ番組やラジオ番組などの放送コンテンツの新たな評価手法としてTwitterにおけるツイートの言語的特徴の定量分析の結果を放送コンテンツに対するツイートの「一様率」として定義し, 番組内でどの様に推移していくか可視化を試みた. さらに, ラジオ番組の面白さとツイートの一様率との相関関係を検証し, ラジオ番組の面白さ平均値は約3分かけて一様率に反映されることがわかった.

今後は一様率としてより妥当性の高い手法の開発・検証を行う予定である. 応用先として動画サイトのコメント欄も想定している. 本稿ではベクトル化した文章のコサイン類似度のみを用いて共感の変化の可視化を試みたが, 一様性という観点から共感に拡大し, ポジネガ分析や感情分析を本実験の結果とそれらを複合的に組み合わせることで共感の実感に近づけることも

表 1 関連検証に用いた番組

ID	番組名	放送日時 (使用した時間)	ジャンル
ann220116	オードリーのオールナイトニッポン	2022/1/16 1:00~3:00 (19 分間)	トーク, 芸人
tama220120	赤江珠緒たまむすび	2022/1/20 13:00~15:30 (27 分間)	バラエティー, アナウンサー
trad220124	THE TRAD	2022/1/24 15:00~16:50 (15 分間)	音楽, ミュージシャン

表 2 面白さスコア平均値と一様率の相関係数.

ID	+0 分	+1 分	+2 分	+3 分	+4 分
ann220116	0.041383	0.157486	0.152175	0.238314	0.519623
tama220120	0.147146	0.253411	0.216174	0.410463	0.329285
trad220124	0.195269	-0.314477	0.423021	-0.340215	-0.227978

今後の課題とする.

謝 辞

本研究は, JST, 未来社会創造事業, JPMJMI21J2 の支援を受けたものである.

文 献

- [1] 太田暢, 村山太一, 矢田竣太郎, 荒牧英治, and 若宮翔子. Twitter を用いたラジオ番組圧縮手法の提案. 第 14 回データ工学と情報マネジメントに関するフォーラム (C21-2), 2022. URL <https://proceedings-of-deim.github.io/DEIM2022/papers/C21-2.pdf>.
- [2] Shoko Wakamiya, Ryong Lee, and Kazutoshi Sumiya. Towards Better TV Viewing Rates: Exploiting Crowd's Media Life Logs over Twitter for TV Rating. In *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*, page 1–10, 2011.
- [3] Yuji Shiratori, Yoshiki Maki, Satoshi Nakamura, and Takanori Komatsu. Detection of football spoilers on twitter. In *Proceedings of International Conference on Collaboration Technologies*, pages 129–141. Springer, 2018.
- [4] Tiziano Bonini and Toni Sellas. Twitter as a public service medium? A content analysis of the Twitter use made by Radio RAI and RNE. *Communication & Society*, 27(2):125–146, 2014.
- [5] Jennifer Edson Escalas and Barbara B Stern. Sympathy and empathy: Emotional responses to advertising dramas. *Journal of Consumer Research*, 29(4):566–578, 2003.
- [6] Jean Decety and Philip L Jackson. The functional architecture of human empathy. *Behavioral and cognitive neuroscience reviews*, 3(2):71–100, 2004.
- [7] Mahshid Hosseini and Cornelia Caragea. Distilling knowledge for empathy detection. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 3713–3724, Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.findings-emnlp.314. URL <https://aclanthology.org/2021.findings-emnlp.314>.
- [8] Matt Kusner, Yu Sun, Nicholas Kolkin, and Kilian Weinberger. From word embeddings to document distances. In *International conference on machine learning*, pages 957–966. PMLR, 2015.
- [9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [10] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space, 2013. URL <https://arxiv.org/abs/1301.3781>.
- [11] Quoc Le and Tomas Mikolov. Distributed representations of sentences and documents. In *International conference on machine learning*, pages 1188–1196. PMLR, 2014.
- [12] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*, 2019.
- [13] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.
- [14] Scott Deerwester, Susan T Dumais, George W Furnas, Thomas K Landauer, and Richard Harshman. Indexing by latent semantic analysis. *Journal of the American society for information science*, 41(6):391–407, 1990.
- [15] Sato Toshinori. Neologism dictionary based on the language resources on the web for mecab, 2015. URL <https://github.com/neologd/mecab-ipadic-neologd>.