

卒業論文

タイトルはここ

english title

富山県立大学大学院 工学研究科 電子・情報工学専攻

2355020 水上和秀

指導教員 António Oliveira Nzinga René 講師

提出年月:

目次

図一覧	ii
表一覧	iii
記号一覧	iv
第1章 はじめに	1
§ 1.1 本研究の背景	1
§ 1.2 本研究の目的	2
§ 1.3 本論文の概要	3
第2章 関連研究	4
§ 2.1 音楽が感情・心理に与える影響	4
§ 2.2 音楽および他分野における推薦システム	8
§ 2.3 感情分析による研究事例	11
第3章 理論の説明	15
§ 3.1 テキストからの感情分析	15
§ 3.2 画像からの感情分析	18
§ 3.3 未定	18
第4章 提案手法	19
§ 4.1 楽曲特徴量の取得の流れ	19
§ 4.2 BERT による感情分析	19
§ 4.3 提案システムの概要	19
第5章 数値実験並びに考察	21
§ 5.1 数値実験の概要	21
§ 5.2 実験結果と考察	21
第6章 おわりに	22
謝辞	23
参考文献	24

図一覧

2.1	レコメンドシステムのイメージ	9
2.2	情報推薦システムの分類	9
3.1	BERT の流れ	16
3.2	BERT の入力部分	16
3.3	bert くん 1	18
3.4	bert くん 2	18

表一覽

記号一覧

以下に本論文において用いられる用語と記号の対応表を示す.

用語	記号
j 人目の使用者の名前	ϵ_j
j 人目の身長	α_j
j 人目の体重	β_j
j 人目の基礎代謝量 (下限)	B_j^L
j 人目基礎代謝量 (上限)	B_j^H
j 人目のアレルギー情報	x_j
j 人の有する生活習慣病	z_j
対象の日数	D
レシピの数	R
食材の数	Q
栄養素の数	N
データベース上の食材数	S
データベース上の食材番号	$d : 1, 2, 3, \dots, S$
日の番号	$k : 1, 2, 3, \dots, 3D$
栄養素の番号	$l : 1, 2, 3, \dots, N$
材料の番号	$m : 1, 2, 3, \dots, Q$
レシピの番号	$i : 1, 2, 3, \dots, R$
i 番目のレシピの名前	y_i
i 番目のレシピの献立フラグ	r_{ki}
i 番目のレシピの主菜フラグ	σ_i
i 番目のレシピの調理時間	T_i
i 番目のレシピの摂取カロリー	C_i
i 番目のレシピの調理コスト	G_i
i 番目のレシピの m 番目の材料の名前	q_{im}
i 番目のレシピの m 番目の材料量	e_{im}
i 番目のレシピの l 番目の栄養素の名前	n_{il}
i 番目のレシピの l 番目の栄養素の量	f_{il}
d 番目の食材名	Z_d
d 番目の食材の販売単位	W_d
d 番目の食材の値段	M_d

はじめに

§ 1.1 本研究の背景

§ 1.2 本研究の目的

く

§ 1.3 本論文の概要

本論文は次のように構成される．

第1章 本研究の背景と目的について説明する．背景ではを述べる．目的ではを述べる．

第2章 関連研究について述べる．

第3章 本研究で用いる

第4章 本研究の提案手法について述べる

第5章

第6章 本研究で述べている提案手法をまとめて説明する．また，今後の課題について述べる．

関連研究

§ 2.1 音楽が感情・心理に与える影響

音楽療法とその効果

・音楽療法は、音楽を用いて心身の健康を改善する方法で、主に「能動的音楽療法」と「受動的音楽療法」に分類される。

・音楽が感情や心理状態に与える影響

音楽は人間の感情や心理状況に大きな影響を与えることが音楽心理学の研究において広く示されている。音楽が人間心理に与える影響には、生理的な反応から心理的な効果まで、さまざまな側面がある。ここでは、音楽が心理状態に与える影響の概要について詳しく説明する。

能動的音楽療法は、音楽を自分で演奏したり歌ったりすることを通じて、感情や身体の実現を促すアプローチである。患者が音楽を作り出す過程で、自己表現や創造性を発揮でき、感情の解放や心理的な安定を得ることができる。この方法は、特に心的外傷を抱えた患者やストレスを感じている人々に効果的とされており、音楽を演奏することで、感情の整理が進み、身体的な緊張も緩和されることが多い。

一方、受動的音楽療法は、音楽を聴くことによってリラクゼーションや心理的な回復を促進する方法である。患者は音楽を聴くことによって、リラックスや集中を促し、感情のバランスを整えることができる。受動的音楽療法は、身体的なリラクゼーションを促進し、心理的な平穏を得るために広く活用されています。

・音楽が生理的な反応に与える影響

音楽は、心拍数や血圧、脳波など生理的な反応に与える影響も広く研究されている。

心拍数と血圧への影響

いくつかの先行事例では異なる特定のテンポの曲を聴くと心拍数に。Bernardi et al. (2006) の研究では、異なるテンポの音楽を聴いた場合に心拍数が変化の様子が報告されています。速いテンポの音楽（120 BPM 以上）は心拍数を上昇させ、交感神経系を活性化します。一方、ゆっくりとしたテンポ（60 BPM 以下）の音楽は心拍数を低下させ、副交感神経系を優位にすることでリラクゼーションを促進します。この効果は、音楽停止後のリバウンド反応にも現れており、音楽が生理的な調整を行う能力を有していることを示唆しています。

ストレスホルモンの減少

あ

脳波への影響

あ

自律神経の影響

あ

.. 音楽が心理的健康に与える影響

.. 音楽と感情調整の役割

音楽の特徴と感情の関係

音楽の感情的効果はいくつの特徴量に影響される。主な特徴量として、テンポ、曲調（メジャー、マイナーなど）、エネルギー、アコースティック性、そして歌詞の内容が挙げられる。

テンポにおける感情的効果

テンポは1分間の拍数(BPM)で測定される。テンポは、その違い(早いテンポ、遅いテンポ)によって感情に影響を与える。120BPM以上の速いテンポの曲は喜びや興奮、活気を喚起され、60～120BPMの中程度のテンポは安定感と穏やかな気分を提供し、リラックスしながら集中力を保てる状態をもたらす。また、60BPM以下の遅いテンポの曲は副交感神経を活性化させ、深いリラクゼーションを促進させる。このような効果は脳の活動にも影響され、早いテンポでは側頭葉の活性化が強まり、遅いテンポでは心拍数や血圧の低下がみられる。(Bernardi et al., 2006) また、中程度のテンポの音楽は、特に聴覚皮質や感情記憶に関連する脳領域を活性化し、強い感情的覚醒をもたらすとされる。

曲調における感情的効果

音楽の曲調(メジャー・マイナー)によって感情に影響を与えることがいくつかの研究によって示されている。軽度のネガティブ感情を抱える人々に対しては、長調(メジャーキー)な曲が感情を改善し、ポジティブな感情を喚起することが示されている。例えば、メジャー調や高テンポの音楽が、リラックス効果やエネルギーの向上をもたらし、ネガティブな感情を軽減することが確認されている【Soleymani et al., 2013】。長調の音楽は、聴衆に活力を与え、心理的なストレスを緩和する効果が期待される。一方で、深刻な悲しみやうつ病状態のような極度にネガティブな感情を抱えている場合、長調よりも短調の方がより効果的に穏やかでリラックスできる音楽を提供することができる。一般的に、長調(メジャーキー)の音楽は明るく快活な印象を与え、ポジティブな感情を喚起しやすい一方、短調(マイナーキー)の音楽は、より内省的で落ち着いた雰囲気を持ち、感情を穏やかに整える効果があるとされている【Juslin & Sloboda, 2010】。特に、深刻な悲しみやうつ病のような状況では、過度にポジティブな音楽が逆に感情の圧迫感や不安感を引き起こす可能性があるため、感情の急激な変化を避けるためにも、リラックスできる短調の音楽が適していると考えられる。

また、Thoma et al. (2013) の研究によれば、短調の音楽は聴衆の心理的な安定感を高め、過度の感情的な変動を防ぐ効果があるとされている。具体的には、短調の音楽が聴く人の心を落ち着かせ、心拍数や呼吸数を安定させることが示されており、これにより、リラクゼーションやストレス軽減が促進されるとされている。また、悲しみや絶望感といった強いネガティブ感情に直面した人々に対しては、短調の音楽が感情の調整をスムーズに行い、無理なく心を落ち着ける助けとなる。

エネルギーにおける感情的効果

エネルギーは音楽が持つ「活力」や「力強さ」を示す指標であり、音楽のリズム、音量、テンポなどに関連している。高エネルギーの音楽は、活気に満ちたリズムとメロディーによってリスナーの気分を明るくし、エネルギッシュな感情を引き起こす。これにより、喜び、興奮、活力といった感情を感じやすくなり、ポジティブな心情を促進する。一方、低エネルギーの音楽は、リズムが遅く、メロディーやハーモニーが穏やかで、音量も控えめである。低エネルギーの音楽は、心拍数を下げ、リラクセス効果を高めるため、ストレスや緊張を緩和する効果がある。

アコースティック性における感情的効果

音楽のアコースティック性は感情や心理状態に影響を与えることが示されている。アコースティック性とは、曲がどれだけ自然な楽器の音、または生音に近いかを示す指標のことである。アコースティック性が高い音楽は、ギター、ピアノ、弦楽器、ドラムなど、生楽器を主体とした音を使用されており、自然で心地よい音色であることが特徴である。また、多くの研究において、アコースティック性が高い音楽が感情のバランスをとる助けになることが示されている。例えば、Thoma et al. (2013) の研究では、アコースティック性の高い音楽がストレスホルモンであるコルチゾールの分泌を減少させ、心理的な安定感を高める効果があることが示されている。また、Saarikallio et al. (2013) による研究でも、アコースティック性が高い音楽が感情の調整に有効であることが確認されている。これらの研究からも、アコースティック音楽が心理的な安定や感情の調整に有効であることが示唆されている。特に、アコースティック性の高い音楽はリラクゼーション効果が高く、心拍数や血圧の低下を促すため、ネガティブな感情が強い状態にある場合に聴くことでリラクセスを促し、気分が改善されることが期待される。一方で、アコースティック性が低い音楽は、エレクトロニック音楽やシンセサイザーを多く使用した音楽を指し、リズムやビートが強調され、感情に強い影響を与える。このような音楽は、活力や興奮を促進し、身体的にも精神的にもエネルギーを増加させる効果がある。特に、ネガティブな感情が少ない人は、アコースティック性が低い音楽のエネルギーを受け入れやすく、ポジティブな感情を高めることができる。逆に、ネガティブ度合いが高い人には、アコースティック性が低い音楽を聴くことはあまり効果的ではないとされている。特に、心の状態が不安定である場合、このような音楽は感情をさらに乱す可能性があり、リラクセスや心理的な安定を促進するには適していない。

歌詞における感情的効果

曲の歌詞は、感情に影響を与えることが示されており、その影響の度合いは感情値により異なる。ポジティブな歌詞の音楽を聴くことにより、ネガティブな感情を軽減す

る効果について、多くの研究が示している。例えば、Juslin & Sloboda (2010) の研究では、ポジティブな感情を表現する歌詞（「愛」「幸せ」など）が、リスナーに幸福感や満足感を引き起こし、心の状態を改善することを示している。歌詞の感情的な内容が感情に与える影響は大きく、ポジティブな感情を持つ歌詞は、気分を向上させ、ストレスや不安を軽減する効果があることが示されている。一方、ネガティブな歌詞は共感を与えることもあるが、特に感情的に不安定な状態にある人々には、より感情を不安定にしてしまう可能性がある。Soleymani et al. (2013) の研究によると、ネガティブな歌詞は共感を誘発する一方で、過度にネガティブな感情を抱える人々にとっては、感情をさらに深める結果となり得ることを示唆している。

ここまで、音楽の特徴量における、感情への影響について述べてきたが、音楽の効果は、その時に聞く人の現在の感情状態に大きく依存する。音楽による気分の変化について、同質の原理 (Iso-principle) が提唱されている。

同質の原理

人は自分の感情状態に似た特徴を持つ音楽を選び、その音楽を聴くことで感情が安定するという効果がある原理である。例えば、ネガティブな感情を抱えている時には、その感情に同調するような音楽を聴くことで、感情的な一致感が生まれ、リラクゼーションや感情の調整が進むとされるというものである。

同質の原理は、ネガティブな感情状態にある場合には、ポジティブな音楽よりもネガティブな感情に同調する音楽が効果的であることを示唆している。例えば、深い悲しみや落ち込みを感じている人が、明るく活発な音楽を聴いても、その音楽が感情的に異質に感じられることがある。この場合、感情的に明るい音楽は逆に不快に感じ、気分転換の効果が薄れる可能性があります。逆に、悲しい気分を持っている場合、暗い音楽を聴くことで、感情的な一致感が生まれ、その後にポジティブな感情に移行する助けになることが考えられます。

同質の原理は、音楽推薦システムにおいても重要な概念です。ネガティブな感情を持つユーザーに対して、無理にポジティブな音楽を薦めるのではなく、感情的に一致する音楽を選ぶことで、感情の調整が促され、最終的にポジティブな感情への移行をサポートすることができます。このように、感情の強さや状態に合わせた音楽選択が、音楽による感情調整の効果を最大化するのです。

・音楽がポジティブ感情に与える影響の研究事例

本研究では同質の原理に基づき、ユーザのネガティブ度合に応じた音楽の推薦を行う。具体的には、ネガティブ強度が強い場合には、感情的に一致する特徴量を持つ音楽を推薦することで、ユーザが自分の感情と調和し、リラックスや心の安定を促進できるようにする。

また、ネガティブ度合いが中程度または低い場合には、少しポジティブな特徴量を持つ音楽を選択し、少しずつ気分を高めるような音楽を推薦することで、最終的にポジティブな感情への移行をサポートする。

§ 2.2 音楽および他分野における推薦システム

近年、ソーシャルメディアの発展やスマートフォンなどの ICT 機器の普及により、多くの人々がインターネットに触れる機会が増えている。その中でも、YouTube¹をはじめとする動画配信サイトや、Amazon²をはじめとする EC（Electronic Commerce：電子商取引）サイトなどを利用する機会が増加している。これらのサービスでは、動画や商品を閲覧している際に、関連性の高い別の動画や商品を提示されることが一般的である。このような技術は「情報推薦システム」と呼ばれ、現在その重要性が高まっている。また、情報推薦システムによりユーザーは膨大な選択肢の中から適切な情報や商品を効率的に発見できるようになり、現代の情報過多の時代において重要な役割を果たしている。レコメンドシステムのイメージを図 2.1 に示す。

推薦システムの目的

推薦システムには永続個人化推薦と一時的個人化、非個人化推薦の 3 種類に分類することができる [?]。それぞれのイメージを図 2.2 に示す。永続個人化推薦とはユーザーの長期的な行動履歴や嗜好を元に、継続的に進化する推薦を行う方法であり、一時的個人化推薦とはユーザーの現在の行動や最近の嗜好を基に、短期間で個別の推薦を行う方法であり、非個人化推薦とはユーザーの個別の嗜好や行動を考慮せず、全ユーザーに共通の推薦を行う方法である。

推薦システムにはいくつかの手法が存在し、主に情報推薦における嗜好の予測方法としてルールベース、コンテンツ（内容）ベースフィルタリング、協調フィルタリング、ハイブリッド法の 4 つに分けることができる。いかにそれぞれの手法について述べ、その後、他分野及び音楽分野における情報推薦の事例、その課題について紹介する。

ルールベース推薦

ルールベース推薦とは、システムの運営者があらかじめ「特定の行動をとった人や特定の属性を持つ人に対してどのような商品や情報を提供するか」というルールを設定し、そのルールにしたがってレコメンドを行う手法である。例えば、パスタを購入しようとしている人にチーズやソースを推薦するように、利用者の行動や嗜好を予測して適切だと思われるルールを設定する。この手法は比較的低コストで容易に実装可能であり、透明性が高いという利点がある。運営者が設定したルールに基づいて商品を推薦するだけなので、推薦の結果を簡単に説明できる。しかし、あらかじめ設定したルールに依存するため、利用者の多様なニーズや嗜好を細かく反映することが難しいという欠点がある。また、利用者の行動パターンが変化した場合に対応するには、ルールを手動で更新する必要があるため、運用コストが高くなる可能性がある。

コンテンツベースフィルタリング

コンテンツベースフィルタリングは、アイテム（利用者に推薦する対象やコンテンツ）の特徴量を使用し、利用者の嗜好に近いアイテムを推薦する手法である。例えば、映画を推薦するシステムを考えてみる。この場合、利用者が「アクション映画」や「サス

¹<https://www.youtube.com>

²<https://www.amazon.co.jp>

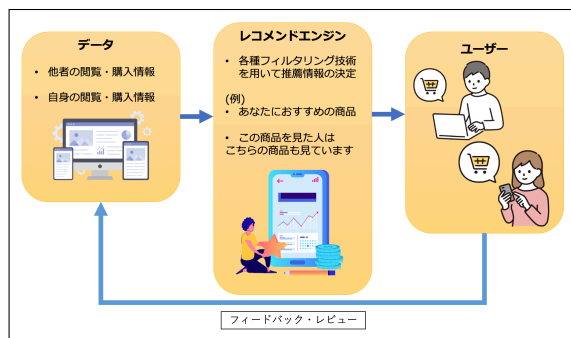


図 2.1: レコメンドシステムのイメージ



図 2.2: 情報推薦システムの分類

ペンス」といったジャンルを検索すると、システムはそれらの特徴量を分析し、ジャンルやテーマが類似する映画を選び出して推薦する。また、過去に視聴した映画のデータを基に、類似した内容や雰囲気映画を提案することも可能である。この手法は、利用者の過去の嗜好に基づき、個人に特化した精度の高い推薦ができるという利点がある。利用者が以前に好んだアイテムと特徴の類似度を計算しアイテムを優先的に推薦するため、個々の利用者に特化した推薦が可能である。一方、この手法は過去の嗜好に基づいて推薦を行うため、利用者に同じようなアイテムが繰り返し推薦され、新しいアイテムの発見が難しいという欠点がある。コンテンツベースフィルタリングには、利用者が自分の好むものを直接指定する「直接指定コンテンツベースフィルタリング」と、利用者の嗜好データからプロファイルを作成し、アイテムデータベースと比較することで嗜好を測る「間接指定コンテンツベースフィルタリング」が存在する。

協調フィルタリング

協調フィルタリングは、システム使用前に利用者の嗜好データを収集し、それを基に利用者の好みや傾向（例えば、どのアイテムが好まれ、どれが嫌われるか）を分析する手法である。この手法では、嗜好が類似している利用者同士が、共通して好むアイテムや嫌うアイテムがあると仮定し、似たような嗜好を持つユーザーを見つけ出します。その後、そのユーザーが好むと思われるアイテムを推薦する。

協調フィルタリングは大きく分けてユーザーベース協調フィルタリングとアイテムベース協調フィルタリングの二つに分類される。ユーザーベース協調フィルタリングとは似た嗜好を持つ他のユーザーを見つけ、そのユーザーが高く評価したアイテムを推薦する手法のことであり、アイテムベース協調フィルタリングとはアイテム同士の類似性に基づき、ユーザーが過去に好んだアイテムに似たアイテムを推薦手法のことである。これら手法は他の利用者の行動データを活用するため、利用者の過去の評価や行動に基づいて適切な推薦ができるという特徴がある。しかし、協調フィルタリングはシステムに新規ユーザーや新規アイテムが追加された場合に過去のデータがないため推薦することが難しくなるという「コールドスタート問題」やユーザーとアイテムの評価データが非常に少ないためにユーザー間の類似性を正確に算出できず推薦精度が低下してしまう「スパース性問題」などの課題がある。これらの問題を解決するために、コンテンツベースフィルタリングやハイブリッドフィルタリングと組み合わせて使われることが多い。

ハイブリッドフィルタリング

ハイブリッドフィルタリングとは、コンテンツベースフィルタリングと協調フィルタリングを組み合わせた手法である。コンテンツベースフィルタリングは推薦がパターン化されやすく、協調フィルタリングは、新規ユーザーや新規アイテムに対して機能しにくいといった欠点があるが、ハイブリッドフィルタリングはそれらの欠点が補われ、幅広いユーザーに適切な推薦を行うことができる。

推薦システムは、現在、様々な分野で活用されており、その適用事例も増えている。例えば、EC サイトの分野では商品の推薦、教育分野では学習コンテンツの提示、利用者のニーズや好みに応じた提案を行う仕組みが使われている。こうした他分野での具体的な推薦システムの事例について紹介し、その後、音楽分野における推薦システムの事例を紹介する。

他分野での推薦システムの事例

1. Amazon の推薦システムの事例

1

2. Coursera における推薦推薦システムの事例

2

音楽分野分野での推薦システムの事例

他分野での事例を書く

1. YoutubeMusic での推薦システムの事例

い

2. AppleMusic における推薦システムの事例

あ

本研究への関連性

他分野での事例を書く

§ 2.3 感情分析による研究事例

感情分析とは自然言語処理の分野の一つであり、データから人々の感情的な状態や意図を抽出し、それを定量的・定性的に分析する技術のことである。感情分析の対象は、テキストや音声、映像、生体情報などがあげられる。テキストによる分析では文章に含まれている単語や表現を分析することで、書き手の感情分析する。音声による感情分析では声の抑揚や声の大きさを分析することで話し手の感情を分析する。人間の感情表現は非言語による部分が大きいため、音声分析では、テキストでは読み取れない情報を読み取れることができる。また、映像による分析では顔の表情や顔の筋肉の動きを分析することで対象の感情を分析する。映像による分析では音声に頼らずに感情をとらえるため、視線やジェスチャーなどの言葉を使わない感情表現を分析することができる。生体による分析では脳波や脈拍などの生体データを分析する。生体反応は無意識的に起こるため、生体による分析は感情を客観的に分析することができる。

テキストによる感情分析の手法は大きく分けて「感情辞書による手法」、「機械学習による手法」、「深層学習による手法」3つのアプローチがある。以下にそれぞれの手法の概要及び特徴について説明する。

1. 感情辞書による手法

感情辞書による手法は感情極性辞書を用いて、単語やフレーズに割り与えられたポジティブまたはネガティブスコアを集計することでテキストの感情を推定する手法である。感情極性辞書を用いた感情推定の流れを説明する。まず、対象となる文章を形態素解析と係り受け解析により単語に分割する。次に、分割した単語を感情極性辞書に基づいてポジティブまたはネガティブに振り分ける。最後に、各単語の情報を集計し、文章全体の感情を推定する。

形態素解析とは、文章を言語において意味を持つ最小単位(形態素)に細分化することである。形態素解析で文章を形態素に分割し、係り受け解析により各形態素の係り先を決定する。日本語の感情分析で使われる感情極性辞書は以下の二つがあげられる

・ 単語感情極性対応表

単語感情極性対応表は、東京工業大学の高村教授が公開した、スピンモデル (SPIN Model) を用いて作成された感情辞書である。スピンモデルとは物理学のスピン理論を応用し、単語の感情極性を文脈に基づいて推定する手法である。単語が同じ文脈で共起する場合、それらの単語は同じ感情極性を持つと仮定し、語彙ネットワークを構築する。このネットワークを用いて、単語同士の感情的な関係性を計算し、感情極性を決定している。

・ 日本語感情極性辞書

日本語感情極性辞書は、東北大学の乾研究室が公開した極性辞書であり、約 8,500 の名詞および複合名詞に感情極性情報を人手で付与している。名詞の評価極性は、評価・感情、状態、行為などを基準に決定されている。

感情辞書による手法では、感情辞書があれば比較的容易に導入可能であり、また、解釈がシンプルであり、直感的に理解しやすいという利点がある。一方、文脈を考慮できないた

め、単語単体では誤った感情判定を行ってしまう場合がある。また、未知語に対応できず、比喩や曖昧表現に対して精度が落ちてしまうという問題がある。

2. 機械学習による手法

機械学習による手法は、データからパターンや規則を学習し、未知のデータを予測する教師あり学習を用いた手法である。教師あり学習とは、正解のデータが用意されており、正しい出力ができるように入力データの特徴やルールを学習していく手法のことである。感情分析の場合、感情ラベル(ネガティブか、ポジティブかなど)が付与されたデータを学習させ、対象の文章が持つ感情を分類する。

機械学習手法を用いた感情分析では、

step1. データ収集及び前処理

感情分析を行うためには分析対象となるテキストを収集することが必要である。分析対象によって異なるが、一般的にはSNSの投稿やレビューサイトのデータ、アンケートデータなどが用いられる。また、収集したテキストデータは、このままでは機械学習のモデルに適用できないため、「テキストクリーニング」、「トークン化」、「ストップワードの除去」、「正規化」などの前処理を行う。テキストクリーニングでは、不要な記号や特殊文字、URLなどを削除し、データのノイズを取り除く。トークン化ではテキストを単語単位で分割し、モデルが処理可能な形に変換する。日本語の場合、文章に対して形態素解析を行い、文章を単語ごとに分解することが一般的である。ストップワードの除去では、意味を持たない単語を除去する。日本語の場合、助詞である「は」「の」「に」などを除去する場合が多い。

また、この段階で、モデルの学習に利用する正解データも収集する。感情分析においての正解データは、「ポジティブ」、「ネガティブ」などの感情ラベルを付けたデータセットを収集する。データセットは、既存のデータセットを用いるか、独自にデータを収集し、活用する

step2. 特徴量の抽出

特徴量抽出では、テキストデータを数値に変換する処理を行う。トークン化によって得られた単語や文節を数値的な形式に変換し、機械学習モデルで利用できる特徴量に変換する。これにより、機械学習アルゴリズムがテキストデータの情報を計算で扱えるようになる。テキストを数値化する手法として、「Bag of Words (BoW)」、「TF-IDF」、「Word2Vec」がある。BoWは文書内の単語の出現頻度を数え、それをベクトル形式に変換する手法である。BoWはシンプルで実装が容易である。しかし、単語の出現頻度のみを考慮するため、文書内の単語の順序や重要度、文法構造は考慮されない。TF-IDFは、単語が特定の文書内でどれだけ重要であるかを考慮した手法である。TF-IDFでは、単語の文書内での出現頻度を表す「TF (Term Frequency)」と、単語がどれだけ多くの文書で使われているかを逆数で示す「IDF (Inverse Document Frequency)」を掛け合わせた値を計算する。これにより、「です」や「ます」などの頻出するが重要ではない単語の影響を減らすことができる。しかし、全文書に対してIDFを計算する必要があるため、大量のデータを扱う際には計算コストが高くなるという課題がある。BoWとTF-IDFは、単語の意味的な関係を考慮できない一方、Word2Vecは単語間の意味的な関係性を反映した数値ベクトルを作成する。Word2Vec

ではCBOW (Continuous Bag of Words) を用いてCBOWは、文脈情報を利用して中心の単語を予測であり、中心単語の周囲にある複数の単語を入力として、それらを平均化することで文脈の情報を集約し、その情報を元に中心単語を推定する。CBOWは計算量が比較的少なく、効率的に学習を進めることができたため、データ量が比較的少ない場合や、迅速なモデル構築が必要な場合に有効である。ただし、文脈全体を一度に平均化してしまうため、単語間の詳細な関係性を捉えるのが難しい。

step3. モデルの学習

この段階では、特徴量抽出の過程で数値化された文章及び正解データを用いて、モデルの学習を行う。学習は、モデルが予測を行い、その予測と実際の正解データとの誤差を最小化するように繰り返し行われる。以下に具体的な学習の流れを説明する。

まず、学習プロセスが始まる前に、モデルのパラメータはランダムに初期化される。次に、ベクトル化されたテキストデータを入力として与える。モデルは、テキストデータのベクトル値を特徴量として予測を行う。この時、予測値を $\hat{y}$ 、入力データを X 、モデルのパラメータを W (重み) および b (バイアス) としたとき、予測値は次のように計算される

$$\hat{y} = \sigma(WX + b) \quad (2.1)$$

ここで、 $W \in \mathbb{R}^d$ は重みベクトル、 b はバイアス、 σ はシグモイド関数であり、以下の式で定義される

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (2.2)$$

次に、モデルが予測した値 $\hat{y}$ と実際の正解データ y との間の誤差を測るために、損失関数 $L(y, \hat{y})$ を使用する。感情分析における2クラス分類 (ネガティブ、ポジティブ) の場合、損失関数はクロスエントロピーとして以下のように表される。

$$L(y, \hat{y}) = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})] \quad (2.3)$$

次に、損失関数を最小化するために、パラメータ W および b を更新する必要がある。そのためには、損失関数 $L(y, \hat{y})$ をそれぞれのパラメータで微分し、勾配 ∇_W および ∇_b を計算する。勾配降下法に基づくパラメータ更新式は次のように表される

$$W := W - \eta \cdot \nabla_W L(y, \hat{y}) \quad (2.4)$$

$$b := b - \eta \cdot \nabla_b L(y, \hat{y}) \quad (2.5)$$

ここで、 η は学習率であり、パラメータの更新幅を調整する。

この更新を訓練データに対して繰り返し行い、損失関数の値を最小化していく。

step4. モデルの評価

この段階では、学習に使用しなかったデータを用い、モデルが新しいデータに対してどの程度正確に予測できるかを検証する。性能の評価には、主な評価指標として「精度」、「適合率」、「再現率」、「F1スコア」が用いられることが多い。精度とはモデルの全予測に対す

る正しい予測の割合であり、適合率はモデルが正と予測した中で実際に正だった割合のことであり、再現率は実際に正だった中でモデルが正と予測できた割合のことであり、F1スコアは適合率と再現率の調和平均をとった値である。以上の指標を用いることでモデルの強みや弱みを分析することができる。たとえば、適合率が高く再現率が低い場合、モデルは慎重すぎる傾向があるといえる。

step5. 感情の予測

この段階では、ベクトル化された文章データを、作成したモデルに適応し、結果を予測する。感情分析の場合、感情スコア（0から1の間の連続値とし、1に近いほどポジティブ、0に近いほどネガティブ）として感情の強さや確信度を出力する。

例えば、ポジティブな文章に対しては高いスコア（1に近い値）を予測し、ネガティブな文章に対しては低いスコア（0に近い値）を予測する。このようにして得られた予測結果は、感情分析の目的に応じて様々な方法で解釈され、評価される。

3. 深層学習による手法

深層学習 (Deep Learning) は、人工ニューラルネットワークを用いた学習手法で、ニューラルネットワークを多構造に用いることで、大量のデータを使って特徴量の抽出や感情分類を行うことが可能となる。従来の機械学習手法に比べて、文脈を考慮した高度な理解が可能となっている。

深層学習にはさまざまなモデルがあり、画像認識に適した CNN(Convolutional Neural Network) や自然言語処理に適した Transformer、生成モデルに適した GAN(Generative Adversarial Network) など、それぞれ特定のタスクに適応するモデルが存在する。自然言語処理や時系列データの分析においては、文脈や時間的な依存関係を考慮することが重要であり、これを実現するためのモデルとして RNN やその改良版が広く用いられている。

• RNN

RNN(Recurrent Neural Network) は、時系列データや文章など、順序が重要なデータを扱うために設計されたニューラルネットワークである。従来のニューラルネットワークでは、入力データの長さが固定である必要があったが、RNN は各時点の出力を次の時点の入力にフィードバックする構造を持つため、データ間の時系列的な依存関係を捉えることが可能である。ただし、長い時系列データにおいては、情報が時間の経過とともに失われる「勾配消失問題」が生じるという欠点がある。

• LSTM

LSTM(Long Short Term Memory) は、RNN を改良したものであり、RNN が長期的に情報を保持できない問題を解消している。LSTM は不要な過去の情報を削除する忘却ゲート、与えられた情報の重要性を記憶する入力ゲート、記憶された情報のうち必要な部分だけを外部に伝える出力ゲートという3つのゲートを備えることにより、長期的な依存関係を学習している。

顔画像を用いた感情分析の手法

顔画像を用いた感情分析の手法は、テキストによる感情分析とは異なる手法があり、以下の3つの手法が一般的に使われている。

1. 特徴ベースによる手法

特徴ベースによる手法は、顔画像から特徴を抽出して感情を分類する手法である。

3. 機械学習による手法

3. 深層学習による手法

感情分析は

感情分析の応用分野

理論の説明

§ 3.1 テキストからの感情分析

BERT の技術的説明を書く。

BERT

BERT は、Google が提案した自然言語処理モデルの一つであり、Transformer の Encoder 部分を基盤とし、Attention メカニズムを用いて単語間の関係性をモデル化している。

BERT による処理の流れを図??に示す。持っていることが最大の特徴であり、文脈を考慮して単語の意味を理解することができる。また、BERT は、テキスト内の単語の位置情報を学習することができるため、単語の順序を考慮した文脈理解が可能である。これにより、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。

結果として、多くの NLP タスクで従来手法を上回る精度を達成しており、BERT は分散表現と転移学習において大きな進展をもたらした。

Transformer

近年、翻訳などの入力文章を別の文章で出力するというモデルは、Attention を用いたエンコーダー、デコーダ形式の RNN が主流であった。しかし RNN は逐次的に単語を処理しているため、訓練時に並列処理ができないという欠点があった。それに対し Transformer は、RNN や CNN を用いず Attention のみを用いている。Transformer は、再帰も畳み込みも一切行わないので並列化が容易であり、他のタスクにも汎用性が高いという特徴がある。Transformer はベクトル化された文章を入力とし、Attention を多数並列に配置した Multi-Head Attention が用いて分析を行っている。

BERT における入力部分

言語を用いたタスクを解く際には、モデルが言語を扱えるように数値化する必要がある。BERT では、まず MeCab を使用して文を単語に分割し、その後 WordPiece を用いて単語をさらにトークンに分割する。BERT の日本語モデルでは、32,000 個のトークンが定義されており、各トークンには固有の ID が割り振られている。BERT への入力時には、このトー

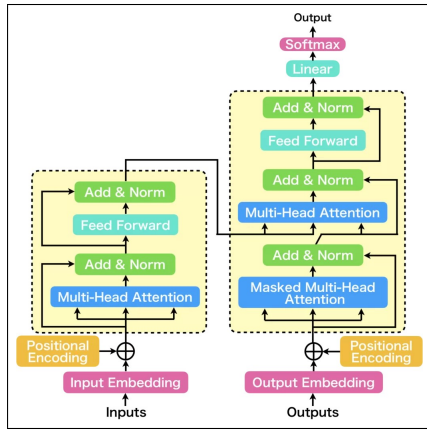


図 3.1: BERT の流れ

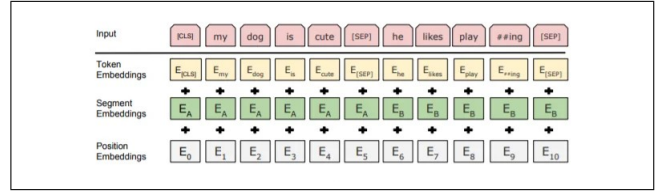


図 3.2: BERT の入力部分

クン ID が使用される。トークン ID に変換されたデータは、BERT モデルに入力される前に、以下の 3 種類の埋め込みを加えることで、モデルに適した数値ベクトルとして表現される。

1. トークンの埋め込み

トークンごとに、事前学習された埋め込み行列 $\mathbf{E} \in \mathbb{R}^{|V| \times d}$ を用いてベクトル表現に変換する。ここで、 $|V|$ は語彙サイズ、 d は埋め込みベクトルの次元である。各トークン ID t_i に対応する埋め込みベクトル $\mathbf{e}_i$ は次のように定義される。

$$\mathbf{e}_i = \mathbf{E}[t_i] \quad (3.1)$$

このベクトル $\mathbf{e}_i$ は、語彙内のトークンの意味的な特徴を学習した固定長ベクトルである。

2. 位置埋め込み

トークン埋め込みを行っただけでは入力の順序に関する情報を持たないため、文章を正しく扱えなくなる可能性がある。そのため、トークン列内での順序情報をモデルに追加するため、位置埋め込みを加える。位置埋め込み行列は $\mathbf{P} \in \mathbb{R}^{n \times d}$ として定義され、各位置 i に対応するベクトル $\mathbf{p}_i$ は次のように計算される

$$\mathbf{p}_{(pos, 2i)} = \sin \left(\frac{pos}{10000^{\frac{2i}{d_{model}}}} \right) \quad (3.2)$$

$$\mathbf{p}_{(pos, 2i+1)} = \cos \left(\frac{pos}{10000^{\frac{2i}{d_{model}}}} \right) \quad (3.3)$$

ここで、 pos はトークンの位置 (0, 1, 2, ...)、 i は埋め込み次元のインデックスで

ある。これにより、固定長の埋め込みベクトルを用いながら、トークンの相対的な順序や距離をモデルが学習可能となる。

3. セグメント埋め込み

BERT では、1つの入力が単一文か複数文かを区別するために、セグメント埋め込みを使用する。各トークンには、対応するセグメント ID s_i (文1なら0、文2なら1) が割り振られ、埋め込みベクトルは以下のように計算される。

$$\mathbf{s}_i = \mathbf{S}[s_i] \quad (3.4)$$

ここで、 $\mathbf{S} \in \mathbb{R}^{2 \times d}$ はセグメント埋め込み行列であり、 s_i が0または1に応じて適切なベクトルが選択される。

4. 埋め込みベクトルの統合

最終的に、トークン埋め込み、位置埋め込み、セグメント埋め込みを加算して、モデルに入力するベクトル $\mathbf{x}_i$ を生成する。

$$\mathbf{x}_i = \mathbf{e}_i + \mathbf{p}_i + \mathbf{s}_i \quad (3.5)$$

これにより、各トークンには、その意味（トークン埋め込み）、位置（位置埋め込み）、文区別（セグメント埋め込み）の情報が含まれるベクトルが追加される。

Attention

Attention は、入力の各トークンが他のすべてのトークンにどれだけ関連しているかを学習するメカニズムであり、この機構は、一般的に自己注意（Self-Attention）として知られている。Self-Attention では、各トークンの埋め込みベクトルを Query、Key、および Value という3つのベクトルに変換し、その相関関係を計算して、重み付けされた値を集約する。Query、Key、および Value は入力単語 x_i にそれぞれの重み W_Q 、 W_K 、 W_V を用いて以下の式で定式化される。

$$Q_i = W_Q \cdot x_i, \quad K_i = W_K \cdot x_i, \quad V_i = W_V \cdot x_i \quad (3.6)$$

また、3つの相関関係は各トークンのクエリとキーとの内積を計算することで求めることができる。

$$\text{score}(Q_i, K_j) = \frac{Q_i K_j^T}{\sqrt{d_k}} \quad (3.7)$$

ここで、 d_k は Query と key の次元数である。得られた式にソフトマックス関数を適応し、 V_i を付加重することによって Attention を求めることができる。

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3.8)$$

Multi-Head Attention

Attention は一つの計算を逐次的に行っているため、1つの視点からしか文脈を読み取ることができない。これに対し、Multi-Head Attention は複数の Attention を並列に連結した構造をしているため、各ヘッドが独自の視点から情報を捉え、最終的にそれらを結合して出力を得ます。このアプローチにより、情報の多様な側面を同時に捉えることができる。Multi-Head Attention は以下の式で定式化される。

$$\text{Multi-HeadAttention}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W_o \quad (3.9)$$

$$\text{where } \text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (3.10)$$

モデルの学習方法

- 事前学習

事前学習では、BERT は Masked Language Modeling (MLM) と Next Sentence Prediction (NSP) という2つのタスクを使ってモデルを訓練する。MLM では、入力文の中でランダムに選ばれた単語を「[MASK]」トークンで置き換え、モデルにその単語を予測させる。MLM において、Multi-Head Attention は文脈を双方向的に考慮し、マスクされた単語を予測する。Multi-Head Attention は、文全体の意味的および構文的な関係を理解し、正しい予測を行うための情報を集約する。

- ファインチューニング

ファインチューニングをする

§ 3.2 画像からの感情分析

§ 3.3 未定

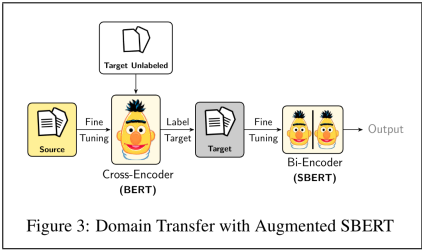


Figure 3.3: bert $\prec \mathcal{L}$ 1

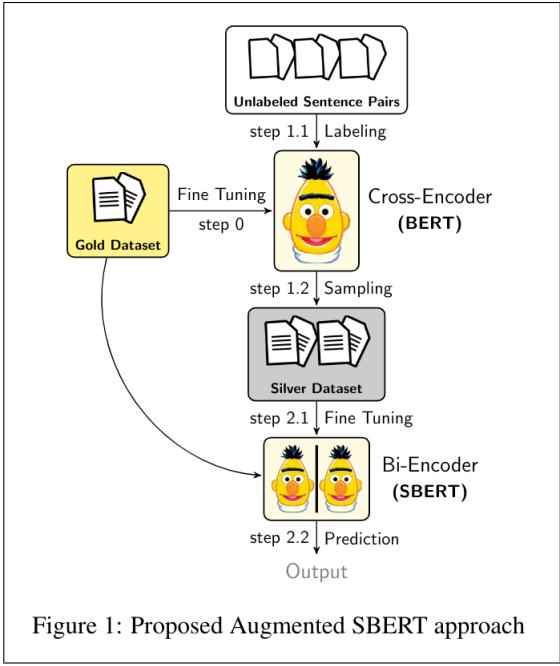


Figure 3.4: bert $\prec \mathcal{L}$ 2

提案手法

- § 4.1 楽曲特徴量の取得の流れ
- § 4.2 BERT による感情分析
- § 4.3 提案システムの概要

数値実験並びに考察

§ 5.1 数値実験の概要

§ 5.2 実験結果と考察

おわりに

謝辞

本研究を遂行するにあたり，多大なご指導と終始懇切丁寧なご鞭撻を賜った富山県立大学工学部電子・情報工学科情報基盤工学講座の António Oliveira Nzinga René 講師，奥原浩之教授に深甚な謝意を表します．最後になりましたが，多大な協力をしていただいた研究室の同輩諸氏に感謝致します．

2025 年 2 月

水上和秀

参考文献

- [1] 公益社団法人 千葉県栄養士会, “生活習慣病の予防、食生活 生活習慣病の予防と食事”, <https://www.eiyou-chiba.or.jp/commons/shokuji-kou/preventive/seikatusyukan/>, 閲覧日 2023.1.7.
- [2] 国立研究開発法人 国立循環器病研究センター, “食事療法について”, <https://www.ncvc.go.jp/hospital/pub/knowledge/diet/diet02/>, 閲覧日 2023.1.7
- [3] ソフトム株式会社, “ソフトム通信 第 79 号「給食業界における A I 活用」”, https://data.nifcloud.com/blog/food-service-provider_ai-use-case_01/, 閲覧日 2022.12.28.
- [4] 貝沼やす子, 江間章子, “日常の献立作りの実態に関する調査研究 (第 1 報)”, 日本調理学会誌, Vol.30, No. 4, pp. 364-371, 1997.
- [5] 株式会社おいしい健康, “おいしい健康”, <https://oishi-kenko.com/>, 閲覧日 2022.10.16.
- [6] 総務省統計局, “小売り物価統計調査による価格調査”, <https://jpmarket-conditions.com/>, 閲覧日 2022.10.11.
- [7] J. W. Ratcliff and D. Metzener, “Pattern Matching: The Gestalt Approach”, *Dr. Dobb's Journal*, p.46, 1988.
- [8] C. A. Coello Coello and M. S. Lechuga, “MOPSO: a proposal for multiple objective particle swarm optimization, ” *Proceedings of the 2002 Congress on Evolutionary Computation (CEC'02)*, Vol. 2, pp. 1051-1056, 2002.
- [9] Q. Zhang and H. Li, “MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition”, *IEEE Trans. Evolutionary Computation*, Vol. 11, No. 6, pp. 712-731, 2007.
- [10] LeftLetter, “多目的進化型アルゴリズム MOEA/D とその改良手法”, <https://qiita.com/LeftLetter/items/a10d5c7e133cc0a679fa>, 閲覧日 2023.1.6.
- [11] J. H. Holland, “Adaptation in Natural and Artificial Systems”, 1975.
- [12] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, “A Fast and Elitist Multi-objective Genetic Algorithm: NSGA-II”, *IEEE Tran. on Evolutionary Computation*, Vol. 6, No. 2, pp. 182-197, 2002.
- [13] D. E. Goldberg, “Genetic algorithms in search, optimization and machine learning ”, *Addison-Wesly*, 1989.
- [14] メディカル・ケア・サービス株式会社, “制限食にはどんな種類があるの?”, 健達ネット, <https://www.mcsg.co.jp/kentatsu/health-care/12106>, 閲覧日 2023.1.6.

- [15] ときわ会栄養指導課, “減塩について”, 栄養指導,
<http://www.tokiwa.or.jp/nutrition/diet/low-salt.html>, 閲覧日 2023.01.15
- [16] 全国健康保険協会, “ちょっとした工夫で脂質をコントロール”,
<https://www.kyoukaikenpo.or.jp/g4/cat450/sb4501/p004/>, 閲覧日 2023.01.15
- [17] 厚生労働省, “日本人の食事摂取基準 (2020 年度版)”,
<https://www.mhlw.go.jp/content/10904750/000586559.pdf>, 閲覧日 2023.01.15
- [18] 東京医科大学病院, “カリウムは調理のくふうで減らせます”, 内臓内科,
<https://articles.oishi-kenko.com/syokujinokihon/dialysis/05/>, 閲覧日 2023.01.15
- [19] 厚生労働省, “糖尿病”, <https://www.mhlw.go.jp/content/10904750/000586592.pdf>,
閲覧日 2023.01.17
- [20] 厚生労働省, “慢性腎臓病”, <https://www.mhlw.go.jp/content/10904750/000586595.pdf>,
閲覧日 2023.01.17
- [21] 腎臓内科, “慢性腎臓病の食事療法”, 東京女子医科大学,
<https://www.twmu.ac.jp/NEP/shokujiryohou.html>, 閲覧日 2023.01.17
- [22] 厚生労働省, “脂質異常症”, <https://www.mhlw.go.jp/content/10904750/000586590.pdf>,
閲覧日 2023.01.17
- [23] 厚生労働省, “高血圧”, <https://www.mhlw.go.jp/content/10904750/000586583.pdf>,
閲覧日 2023.01.17
- [24] 厚生労働省, “食べ物アレルギー”, アレルギーポータル,
<https://allergyportal.jp/knowledge/food/>, 閲覧日 2023.01.17
- [25] J. Blank, “pymoo: Multi-objective Optimization in Python ”,
<https://www.egr.msu.edu/kdeb/papers/c2020001.pdf>, 閲覧日 2023.1.22.
- [26] 和正敏, “多目的線形計画問題に対する対話型ファジィ意思決定手法とその応用”, 電子情報通信学会論文誌 Vol. J 65-A, No. 11, pp. 1182-1189, 1982.
- [27] 厚生労働省, “日本人の食事摂取基準 (2020 年版) ”,
<https://www.mhlw.go.jp/content/10904750/000586553.pdf>, 閲覧日 2022.12.26.
- [28] 農林水産省, “一日に必要なエネルギー量と摂取の目安”,
https://www.maff.go.jp/j/syokuiku/zissen_navi/balance/required.html, 閲覧日 2023.1.22.