

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

方向性と進捗

水上 和秀 (Kazuhide Mizukani)
u355020@st.pu-toyama.ac.jp

富山県立大学 工学部 電子情報工学専攻

November 26, 2024

1. はじめに

2/34

研究の背景 (仮)

- SNS や Web サイトの利用者は増加し、個人のインターネット利用率は 80 % を越えており、大勢の人がインターネット上の情報に触れることができるようになった。
- 現代の消費行動の特徴としてインターネット上の口コミや評価を重視し、意思決定や行動に大きな影響を与えている。
- しかしインターネット上の口コミは対面のクチコミよりも利用可能な非言語的手がかりが乏しく、相手に伝わりにくい。

研究の目的 (仮)

- BERT でレビューの感情分析を行い、分析結果に大きく影響を及ぼしている単語を SHAP で可視化を行う。
- トピックごとにクラスタリングを行うことでユーザにとってより解釈可能なレビューの分析を行うシステムを提案する。

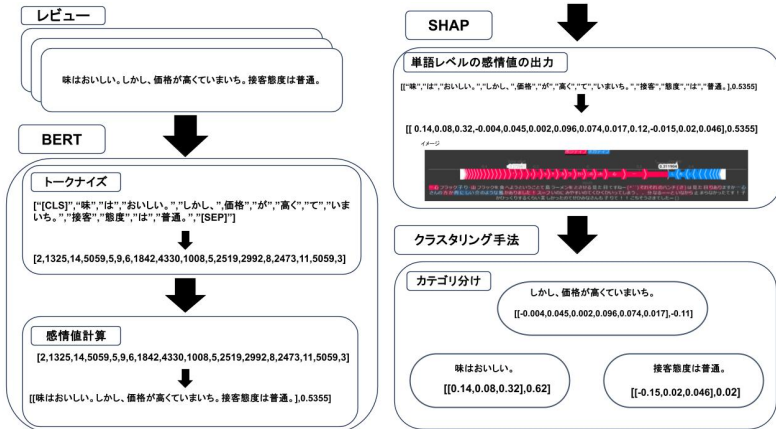
実装の流れ

- 1** データセットの取得
→解きたいタスクの学習用データに使用。
ネガポジ分析の場合、ラベル付けしたネガティブな文章とポジティブな文章を大量に用意する
- 2** 事前学習モデルの構築 (BERT)
→「事前学習」「ファインチューニング」の二つの学習を行うモデルを構築する
- 3** トークナイザーの構築 (BERT)
→文章を語彙（トークン）に分割したうえで、BERT モデルに入力できる形に変換する処理
- 4** 分析レビューの取得
→感情分析を行うためのレビューをスクレイピングする
- 5** BERT による感情分析
- 6** SHAP の実装
→ BERT で分析した結果を可視化 (数値化) する
- 7** レビュー文をクラスタリング
- 8** 出力
→可視化した結果をわかりやすく表示する

実装のイメージ (システム)

4/34

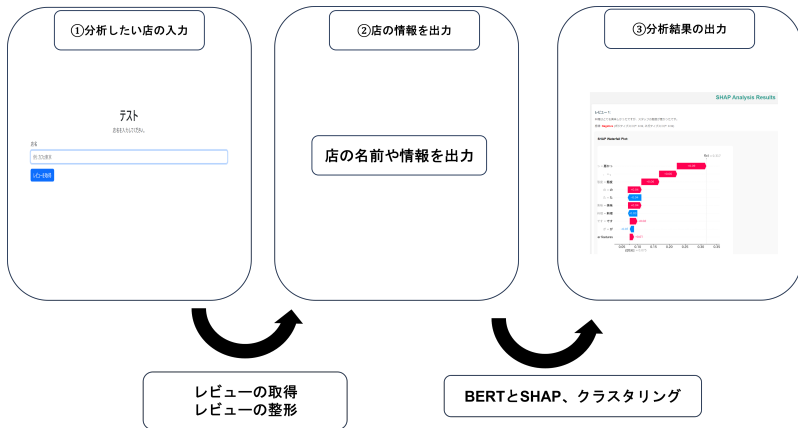
- はじめに
- 提案手法
- はじめに
- 関連研究?
- BERT と SHAP、クラスタリングについて
- 提案手法
- 数値実験ならびに考察
- おわりに



システムの流れ (ユーザーの操作)(仮)

5/34

1. はじめに
2. 提案手法
 1. はじめに
 2. 関連研究?
 3. BERT と SHAP、クラスタリングについて
 4. 提案手法
 5. 数値実験ならびに考察
 6. おわりに



できてること

- レビューサイトのレビューをスクレイピング
- データセットの収集
- 自分で BERT モデルの構築
- SHAP の実装

問題点

- 感情分析の精度がまだよくない
- レビュー文のクラスタリングについて

検討中

- システム化する (サーバー上でできるようにする)
- ほかの機械学習の手法と比較
- わかりやすく出力する

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

1. はじめに
2. 提案手法
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

進捗

- レビューの取得部分について
- レビューの整形
- bert のパラメータの調整
- レビュー文クラスタリングの検討
- システムの出力部分の作成

- 1. はじめに
- 2. 提案手法
 - 1. はじめに
 - 2. 関連研究?
 - 3. BERT と SHAP、クラスタリングについて
- 4. 提案手法
- 5. 数値実験ならびに考察
- 6. おわりに

レビューの取得について

分析する店の名前を入力すると、入力されて店のレビューを食ベログから取得できるようにした

テスト

店名を入力してください。

店名

例: カフェ東京

レビューを取得

また、最近の急激な金利上昇は日本銀行の政策の方向性からいって一時的な現象だと考えられる。つまり、日本銀行は「金融政策の正常化」を目標として、金融市場の流動性を確保し、金融機関の資金調達コストを抑制し、企業投資を促進することを目的としている。金利上昇は、市場の資金需要が過剰なため、日本銀行が金融政策を正常化するまでの一時的な現象と見られる。金利上昇は、市場の資金需要が過剰なため、日本銀行が金融政策を正常化するまでの一時的な現象と見られる。金利上昇は、市場の資金需要が過剰なため、日本銀行が金融政策を正常化するまでの一時的な現象と見られる。

BERT のパラメータについて

以下は BERT に用いるパラメータ。どのパラメータが分析精度に影響しているか検討中。

- エポック数: トレーニングデータを学習させる回数
- バッチサイズ: 一度にモデルに与えるデータの数
- ウォームアップステップ: 学習率を徐々に増加させるためのステップ数
- 重みの減衰率: パラメータ更新の際に重みの値を少しずつ減らす手法

```

1
# モデルの初期化
model = BertForSequenceClassification.from_pretrained('bert-base-uncased', num_labels=2)

# トレーニング引数の設定
training_args = TrainingArguments(
    output_dir='./results',
    num_train_epochs=3, # エポック数
    per_device_train_batch_size=16, # トレーニングのバッチサイズ
    per_device_eval_batch_size=16, # 評価のバッチサイズ
    warmup_steps=500, # ウォーミングアップステップ
    weight_decay=0.01, # 重みの減衰率
    logging_dir='./logs',
    logging_steps=10,
)
    
```

1. はじめに

2. 提案手法

1. はじめに

2. 関連研究?

3. BERT と
SHAP、クラスタ
リングについて

4. 提案手法

5. 数値実験ならび
に考察

6. おわりに

クラスタリングについて

文章をカテゴリごとに分類するために、クラスタリング手法を用いる。ラベル付き教師用データがないので、教師なし学習によるクラスタリング手法を用いる。

kmeans 法と LDA 法

kmeans 法はデータを数値化し、互いに近いデータを同じクラスに分類するクラスタリング手法。

LDA は、文書内の「トピック」を潜在的に見つけ出し、そのトピックに基づいて文書を分類する手法。LDA では、各文書は複数のトピックの混合であると仮定し、トピックごとの単語の出現確率分布を計算する。

実装してみた

レビューを文節ごとに区切り、kmeans 法で実際にクラスタ分けを行った。
(味、価格、接客態度、その他) に分類されることを目的に、クラスタ数は 4 つに設定を行った。

[illegible]

1. はじめに

2. 提案手法

1. はじめに

2. 関連研究?

3. BERT と
SHAP、クラスタ
リングについて

4. 提案手法

5. 数値実験ならび
に考察

6. おわりに

考察

- kmeans 法は似たような文章を自動でクラスタリングする手法だが、思ったようなクラスタ分けができていなかった→レビュー文を適正な方法で数値化する必要がある
- また、クラスタ数を適切に決める必要がある

分析結果の出力の作成

BERT と SHAP で分析した結果を html で出力する部分を作成した

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

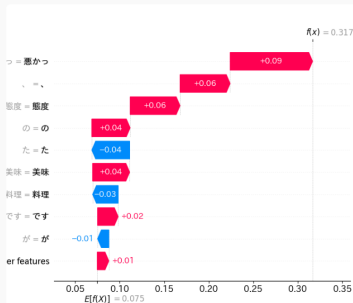
SHAP Analysis Results

レビュー 1:

料理はとても美味しかったです。スタッフの態度が悪かったです。

感情: **Negative** (ポジティブスコア: 0.32, ネガティブスコア: 0.32)

SHAP Waterfall Plot:



1. はじめに

2. 提案手法

1. はじめに

2. 関連研究?

3. BERT と
SHAP、クラスタ
リングについて

4. 提案手法

5. 数値実験ならび
に考察

6. おわりに

付け加えれそうな機能

分析するレビューにさくらレビューが紛れている可能性があるので、さくらレビューの判断 (清水さん)

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

やったこと

- レビューの整型を行った
- BERT のパラメータを調整した
- レビュー文クラスタリングを行った

これからやること

- BERT のモデルの調整
- それぞれの精度を上げる
- システム化

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

以下、前回の資料

1. はじめに

18/34

研究の背景 (仮)

- SNS や Web サイトの利用者は増加し、個人のインターネット利用率は 80 % を越えており、大勢の人がインターネット上の情報に触れることができるようになった。
- 現代の消費行動の特徴としてインターネット上の口コミや評価を重視し、意思決定や行動に大きな影響を与えている。
- しかしインターネット上の口コミは対面のクチコミよりも利用可能な非言語的手がかりが乏しく、相手に伝わりにくい。

研究の目的 (仮)

- BERT でレビューの感情分析を行い、分析結果に大きく影響を及ぼしている単語を SHAP で可視化を行う。
- トピックごとにクラスタリングを行うことでユーザにとってより解釈可能なレビューの分析を行うシステムを提案する。

2. 関連研究?

19/34

1. はじめに

2. 提案手法

1. はじめに

2. 関連研究?

3. BERT と
SHAP、クラスタ
リングについて

4. 提案手法

5. 数値実験ならび
に考察

6. おわりに

2.1

2.2

2.3

3.1 BERT について?

20/34

BERT とは

Bidirectional Encoder Representations from Transformers (BERT) は, Devlin らが提案した自然言語処理モデルであり, 双方向 Transformer というニューラルネットワークを利用している. BERT は大量の教師なしデータによる事前学習を行ったモデルを, 比較的少量の教師ありデータでファインチューニングして個々のタスクに合わせて学習する. これにより事前学習済みモデルをそのまま適用するよりも性能の向上が期待できる.

また, BERT の特徴として, 文章を文頭と文末の双方向から学習することで, 文脈を考慮できることが挙げられる. BERT ではこの双方向からの学習を, 「事前学習」の中の「MLM」「NSP」の2つのタスクを行うことで可能としている

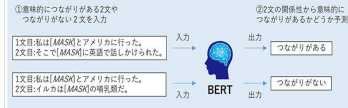
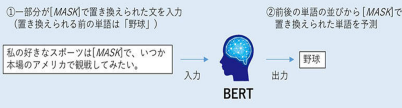
- 1. はじめに
- 2. 提案手法
- 1. はじめに
- 2. 関連研究?
- 3. BERT と SHAP, クラスタリングについて
- 4. 提案手法
- 5. 数値実験ならびに考察
- 6. おわりに

3.1 BERT について?

21/34

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスティングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

- ある文章において一部のトークンを特殊トークンである [MASK] に置き換えて、その [MASK] に入るトークンを予測する言語モデルのこと。
→単語に対応する文脈情報を獲得できる
- 「意味的につながりがある 2 文」、または「意味的につながりのない 2 文」を入力し、2 文の関係性を考慮することで「入力された 2 文に意味的につながりがあるかどうか」を予測する
→単語の関係性だけでなく文章の関係性の情報も獲得できる



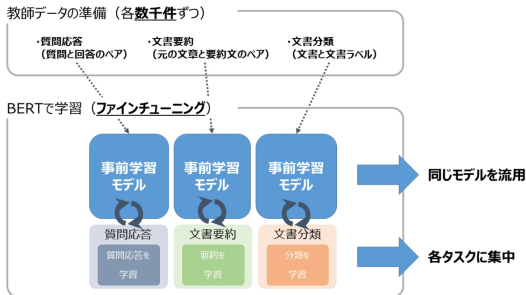
3.1 BERT について?

22/34

ファインチューニング

BERT の学習済みモデルは、そのまま使うことは珍しく、一般に、解きたいタスクに応じて特化するようにする。

ファインチューニングを行うときにはモデルの初期値として、事前学習で得られたパラメータを用い、新たに加えられた分類器のパラメータにはランダムな値を与える。そして、ラベル付きデータを用いて BERT と分類器の両方のパラメータを学習する → ファインチューニングの際事前学習で得られたパラメータを初期値として用いることで比較的少数の学習データでも高い性能のモデルを得ることができる



1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

3.2 SHAP について?

SAHP とは

機械学習で導出した予測値に対して各特徴量がどのくらい寄与しているかを算出する手法で、シャープレイ値の考え方に基づいている

シャープレイ値とは

協力ゲーム理論において複数のプレイヤーの協力によって得られた利得を各プレイヤーの貢献度に応じて構成に分配するための手段の一つ

- 3 人のプレイヤー (1.2.3) が協力してゲームに挑戦し、利得として、以下の賞金
が得られるとする
- このときの 1.2.3 にそれぞれどのようにお金を分配するか。

表1 協力ゲームの例

参加プレイヤー	賞金/万円
1	4
2	6
3	10
1, 2	16
1, 3	22
2, 3	30
1, 2, 3	60

- はじめに
- 提案手法
- はじめに
- 関連研究?
- BERT と SHAP、クラスターリングについて
- 提案手法
- 数値実験ならびに考察
- おわりに

- このとき各プレイヤーの限界貢献度を導入する。限界貢献度とは、プレイヤー i が参加したときの利得の増加分である。
- 例えば、プレイヤーの参加順「 $1 \rightarrow 2 \rightarrow 3$ 」のときのプレイヤー 3 の限界貢献度は、 $v(1, 2, 3) - v(1, 2) = 60 - 16 = 44$ のように計算できる。
- 各プレイヤーのシャープレイ値は以下ようになる

プレイヤー 1: $(4 + 4 + 10 + 30 + 12 + 30)/6 = 15$

プレイヤー 2: $(12 + 38 + 6 + 6 + 38 + 20)/6 = 20$

プレイヤー 3: $(44 + 18 + 44 + 24 + 10 + 10)/6 = 25$

表2 限界貢献度

プレイヤーの参加順	各プレイヤーの限界貢献度		
	1	2	3
$1 \rightarrow 2 \rightarrow 3$	4	12	44
$1 \rightarrow 3 \rightarrow 2$	4	38	18
$2 \rightarrow 1 \rightarrow 3$	10	6	44
$2 \rightarrow 3 \rightarrow 1$	30	6	24
$3 \rightarrow 1 \rightarrow 2$	12	38	10
$3 \rightarrow 2 \rightarrow 1$	30	20	10

1. はじめに
2. 提案手法
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

シャープレイ値の定式化

- 一般的には、プレイヤー i のシャープレイ値は次式によって定式化される。ただし、 s は連携 S に含まれるプレイヤー数である。

$$\phi_i = \sum_{S: i \in S \subset N} \frac{(s-1)!(n-s)!}{n!} \{v(S) - v(S - \{i\})\}$$

シャープレイ値と SHAP について

協力ゲーム理論のシャープレイ値の概念を応用して、特徴量の貢献度を計算

協力ゲーム (Shapley値)

貢献が違う複数人で仕事をしたときに
各人の仕事に対する貢献度を考慮して
公平に報酬を分配する方法



SHAP

貢献が違う複数の特徴量で予測を
したときに、各特徴量の予測に対
する貢献度を考慮して、公平に
予測値への貢献を分配する方法



SHAP の定式化

解釈したい予測モデルを f 、バイナリ変数 (0 か 1 の変数) を z 、各特徴量に対する貢献度を ϕ_i とすると以下のようにあらわす

$$g(z) = \sum_{i=1}^p \phi_i z_i$$

$$\phi_i(f, x) = \sum_{z \subseteq x} \frac{|z|!(p - |z| - 1)!}{p!} [f(z) - f(z \setminus i)]$$

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

文章のカテゴリ分類について

- レビューの文章を文節ごとに区切り、カテゴリ (味に関する文、価格に対する文、接客態度に関する文、など) ごとにクラスタリング分けを行う。

クラスタリング手法

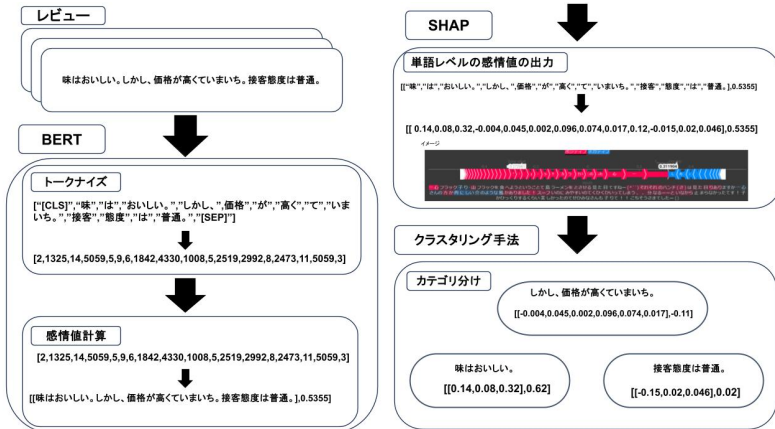
クラスタリング手法で定番の k-means 法か LDA を検討中

実装の流れ

- 1** データセットの取得
→解きたいタスクの学習用データに使用。
ネガポジ分析の場合、ラベル付けしたネガティブな文章とポジティブな文章を大量に用意する
- 2** 分析レビューの取得
→感情分析を行うためのレビューをスクレイピングする
- 3** 事前学習モデルの構築 (BERT)
→「事前学習」「ファインチューニング」の二つの学習を行うモデルを構築する
- 4** トークナイザーの構築 (BERT)
→文章を語彙 (トークン) に分割したうえで、BERT モデルに入力できる形に変換する処理
- 5** BERT による感情分析
- 6** SHAP の実装
→ BERT で分析した結果を可視化する
- 7** レビュー文をクラスタリング
- 8** 出力
→可視化した結果をわかりやすく表示する

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

- はじめに
- 提案手法
- はじめに
- 関連研究?
- BERT と SHAP、クラスタリングについて
- 提案手法
- 数値実験ならびに考察
- おわりに



進捗 1(データセットの取得)

- 1 TIS が公開している感情分析を行うためのデータセット「chABSA-dataset」を取得できるようにした。
- 2 「chABSA-dataset」は上場企業の有価証券報告書 (2016 年度) をベースに作成されたデータセットで、各文に対してネガティブ・ポジティブのラベル付けがされている
- 3 このデータセットを BERT に読み込ませ、モデル構築を行う

text	labels
売上収益は、抗がん剤「レンビマ」(欧州における腎細胞がんに係る製品名「Kisplyx」)および抗てん	1
セグメント別には、日本医薬品事業およびアジア医薬品事業が増収となりました	1
また、すべての海外セグメントにおいて現地通貨ベースで成長を果たしました	1
グローバルブランド4品目合計では、為替の影響を受ける中で、前期から14.5%増の728億22百万円と	1
営業利益は、オペレーションの高質化・効率化に加え、EAファーマ株式取得に伴う一時収益(割安購	1
親会社の所有者に帰属する当期利益は、393億58百万円(前期比28.4%減)となりました	0
基本的1株当たり当期利益は、137円63銭(前期より54円61銭減)となりました	0
当期利益にその他の包括利益を加減した当期包括利益は、前期に円高の進行によって為替換算差額が	1
売上収益の内訳は、医療用医薬品が2,439億99百万円(同4.3%増)、ジェネリック医薬品が280億27百万	1
品目別売上収益については、ニューロロジー領域で、不眠症治療剤「ルネスタ」が80億12百万円(前期	1
ファイザー社と共同販促を展開している疼痛治療剤「リリカ」の共同販促収入は242億68百万円(同1.8	0
オンコロジー領域では、「ハラヴェン」が77億63百万円(同14.2%増)、「レンビマ」が27億15百万円	1
さらに、ヒト型抗ヒトTNFαモノクローナル抗体「ヒュミラ」は376億62百万円(同15.4%増)と順調に	1
売上収益は、1,172億17百万円(前期比4.1%減、現地通貨ベースでは6.3%増)となりました	0
セグメント利益は、効率的なマーケティング活動による販売管理費の減少および前期に発生した米国	1
品目別売上収益については、ニューロロジー領域で、抗てんかん剤「Banzel」が138億48百万円(前期	1
「Belviqu」は37億14百万円(同16.0%減)でした	0
オンコロジー領域では、制吐剤「Aloxi」が480億82百万円(同12.1%減)、「ハラヴェン」が166億19百	0
売上収益は492億74百万円(前期と同等水準、現地通貨ベースでは17.0%増)、セグメント利益は136億96	1

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

進捗 2(分析レビューの取得)

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

- 1 食べログのラーメンのレビューをスクレイピングし、利用できるようにした。
- 2 1つの店に対して上位 20 件分を取得し、感情分析をして、その店の分析結果を表示する

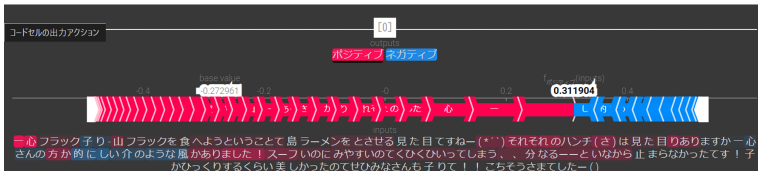
動画	水見庵煮干ラーメン 1000円今日は、富山県を代表する富山1の人気を誇るラーメン屋である貴賓館さんにお邪魔しました！ 初見！読むことが不可能なお店の名前はとんじろ
貴賓館	旅行中、通りかかった水見に不味いお店を探していたところ、断り1点数のうちに遭遇、すぐに向きました。平日13:20に15人の列に接続、すぐに店員さんがやってきました。朝の朝に相いました。6時半到着で7時とあつた。朝は遅いので待ち時間も長くなりませんでした。13時迄で入店。7時に20人以上の列が伸びてます。スープが大好き美味し、自宅みたいで
貴賓館	日曜、富山市の初日は水見！ 湯呑みで野菜とあつた、折角のため10名ほど店員を待ちかゝり！という事で水見駅前のタクシーまで。11時48分に着いたところと54人(赤ちゃんとあつた)
貴賓館	今日のお昼は富山県水見市の貴賓館(とんじろ)さんにお伺いします。営業時間：11:00-15:00 / 17:00-23:00 定休日：月曜日 夜店：カウンター 10席 / テーブル8席 駐車
貴賓館	〇氣に入った点・旨みだけを抽出した透明感ある煮干しスープ〇追加情報・朝方が始まるとが専用メニューのよう〇駐車場有〇丸い・列は待つ所を挟んだ側から割
貴賓館	〇氣に入った点・旨みだけを抽出した透明感ある煮干しスープ・様々な具材の組み合わせが美味とよくまとめた白濁スープ〇追加情報・1時後は待つ所を挟ましよう
貴賓館	〇氣に入った点・旨みだけを抽出した透明感ある煮干しスープ〇注意・1時間待つ所を挟ましよう
貴賓館	貴賓館に訪問しました。今回お昼に富山でしていること〇朝に訪問風ラーメンを注文したスープは美味しいものであり、列は美味しくて〇列は美味しくて〇列は美味しくて
貴賓館	貴賓館に訪問白濁油以外ありけ状態です……………湯出汁は金分が効いてあり美味いです。また、白濁油にコクがあり良いです。麺は中細麺使用喉越しがよいです。チャ
貴賓館	水見煮干ラーメン1000貴賓館〇水見水見に移動して富山1に念願の初訪問。店の前に二人、向かい一人並ぶんですが直ぐ案内され美味し無し。営業システムス
貴賓館	※3.4タッパ付3.8タッパ4.0タッパあつたのもり更新されていなかったようであり直し湯呑み営業はあつた人並ぶの噂を聞いていたのが最良でした朝方から拭きと
貴賓館	水見庵煮干ラーメン1000★★★★★☆☆☆☆☆をみ取って訪れる富山県。富山県No.1の評価で、永らくBMにて立ち。平日12時半の時間帯に着くと25人の行列が。1時間
貴賓館	最高に美味かつたです！うまかつたBuono!はのお店を評価できるような者ではないです。でもジャッジし点数を付けての評価は投えてさせていただきます。
貴賓館	最高に美味かつたです！うまかつたBuono!はのお店を評価できるような者ではないです。でもジャッジし点数を付けての評価は投えてさせていただきます。
貴賓館	土日は朝方ラーメンをやてとる事で来ました。こちらラーメン百名店「貴賓館(とんじろ)」さんです。開店1時間前の6時30分頃に書いてお目です。ちょっと物を置いて
貴賓館	せっかく富山に来たのだからと思い、貴賓館さんへ訪問しました。富山駅から富山駅駅り換えて約60分、水見駅に着。駅から約10分の距離です。お店は開店2時間前の
貴賓館	以前から行きたいと思っていた、貴賓館(とんじろ)さん。富山のラーメン名店の本店です。水見湯通しに並びてはてな415号車一本入ったところから、人気のあつた
貴賓館	富山県「ラーメン人気No.1」のトピックが今日1日朝方ラーメンを始めたので行ってみたい。朝ののそんじろに並びてはてな415号車一本入ったところから、人気のあつた

進捗 3(BERT モデルの構築、SHAP の実装)

- 1 自分でモデルを作成し、レビュー文に適応させ、感情値を出力できるようにした。
- 2 BERT で出力した結果対して、SHAP で可視化できるようにした。

※ テストデータ

```
test_texts = ["** 一山ブラック煮玉子入り ¥980- 富山ブラックを食べようということで訪問徳島ラーメンを彷彿とさせる見た目ですね(・_・)時それぞれ醤油のパンチ(濃さ)は見た目通りありますが一山さんの方が全体的に優しい魚介のような風味がありました！スープ濃いのに飲みやすいのでぐびぐびいってしまう、塩分過多なるーと思いながら止まらなかったです！煮玉子がびっくりするくらい美味しかったのでぜひみなさんも煮玉子入りで！！ごちそうさまでしたー(・_・)**"]
test_encodings = tokenizer(test_texts, truncation=True, padding=True, max_length=128, return_tensors='pt')
```



- 1 出力結果の精度がまだよくないので、モデルの設定を変えて精度を上げる
- 2 文章をカテゴリごとに分類できていないため、レビュー文をカテゴリごとに分類する手法を検討

- 1. はじめに
- 2. 提案手法
- 1. はじめに
- 2. 関連研究?
- 3. BERT と SHAP、クラスタリングについて
- 4. 提案手法
- 5. 数値実験ならびに考察
- 6. おわりに

有効性の検証

- 1 ほかの感情分析の手法も実装し、どれだけ優れているか比較する
- 2 実際にシステムを利用してもらい、どれだけつかいやすかったかアンケートを取る
- 3 ほかに有効性を示す手法があれば検討

6. おわりに

34/34

1. はじめに
2. 提案手法
1. はじめに
2. 関連研究?
3. BERT と SHAP、クラスタリングについて
4. 提案手法
5. 数値実験ならびに考察
6. おわりに

やったこと

- あ

これからやること

- い