

## 文脈から感情を分析するための感情判断器の検討 Emotional Judgment for Emotions Analysis from Context

大澤 泰我<sup>†</sup>      マッキン ケネスジェームス<sup>†</sup>      永井 保夫<sup>†</sup>  
Taiga Ohsawa      Kenneth J Mackin      Yasuo Nagai

### 1. 序論

コミュニケーションロボットの需要が高まる中,人と対話システムの自然なコミュニケーションの実現が求められている.人と対話システムの自然な対話の実現において,感情情報は必要不可欠である.本研究[7]では,深層学習を利用してテキストから感情を判断し,対話システムが感情的な会話を返すことで,人と対話システムの自然なコミュニケーションの実現を目的とする.ここでは,会話内容と感情判断器から受け取った感情値を元に,会話の中で人に近い,咄嗟に出たと思わせる発言を対話システムに起こさせることを考える.本論文では,主に対話システムのための感情判断器を検討する.

### 2. 関連研究

Ptaszynski[2]らによる感情的要素/感情表現分析システム ML-Ask は,約 2,100 語の辞書によるパターンマッチングにより求められた「喜,怒,哀,怖,恥,好,厭,昂,安,驚」の 10 クラスの感情モデルを採用している.

山下[1]らは Plutchik の感情の輪[5]における基本感情を感情モデルとして用いている.この感情モデルは 8 つの基本感情があり,各感情がペアとなって互いに反対の感情を持つという特性がある.本研究では,この Plutchik の感情の輪を元に 8 つの感情分類の特性を,システムとの親和性の高さから感情モデルとして採用する.

Ptaszynski らの感情的要素/感情表現分析システムは,感情表現辞書を使って感情を判断しており,感情推定精度は 57.6%となっている.感情表現辞書は,感情表現辞典を元に,3 人の専門家にアンケートを取り,平均をとって定義した感情語や感情をあらわす単語が収録されている.山本[3]らは,オリジナルの辞書を使って,感情を判断しており,感情判断精度は 55.0%となっている.オリジナルの辞書は,学生 20 人にアンケートをとり,「その文章はこういう感情である」と定義された文章が収録されている.

### 3. 文脈から感情を分析するための感情判断器の検討

本研究の感情判断器は,データセットを学習し,学習のモデル化を行い,学習済みモデルを用いて感情分類を行う.データセットは,顔文字付きの文章を集め,顔文字に対応する感情を正解データとし学習させる.つまり顔文字や絵文字付きの文章から感情情報を取得する.感情判断器は,Plutchik の感情の輪を元に 8 つの感情「喜び,信頼,恐れ,驚き,悲しみ,嫌悪,怒り,期待」に分類を行った.本研究の感情判断器は,対話システムから入力された会話文に対して,会話文から感情を判断し,正しい感情値を返すことに成功した.

しかしデータセットの文中には単語間をつなぐための接続語が多く存在する.接続語は学習時の誤認になる可能性があると考えられる.そのため,本論文では,接続語を含めたデータセットと,接続語を除いたデータセットの学習精度の比較を行う.

### 3.1 データセット

Twitter API を利用し Twitter から,各感情に対応する顔文字付きの文章を取得する.取得した文章は,形態素解析を行い単語ごとに区切る.データは,顔文字を元にラベリングを行う.収集したデータは 79,821 件の文章となっている.

表 1 収集データの加工例

原文	# test
@you ありがとうございます!(^^)	http://test.com
加工後	0,ありがとうございます!

本研究では,テキストから辞書を作る.ベクトル化では,Keras[6]のトークナイズを用いて,テキストをトークン列に分離して,各単語のインデックスに変換したベクトルを作る.ベクトル化したデータを入力用データ(データセット)とする.データセットの 20%をテストデータとし,残りの 80%を学習データとする.

### 3.2 感情判断器のモデル構成

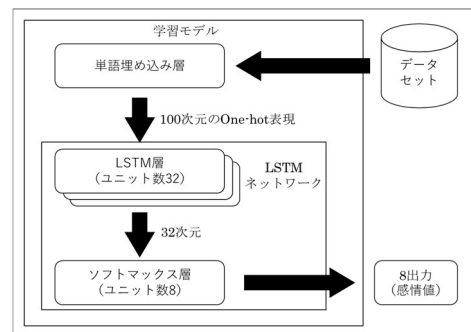


図 1 感情判断器のモデル構成

本研究では,LSTM ネットワークを用いる(図 1).単語をベクトル化したものを入力する.単語埋め込み層では One-hot 表現を用いる.

学習データにはデータセットの 80%を入力し,Embedding 層のユニット数は 100 とした.LSTM 層のユニット数は 32,全結合層のユニット数は 8 で,全結合層の活性化関数にはソフトマックスを用いた.最適化手法には Adam を用い,学習回数は 100 回とした.

### 3.3 感情判断器の動作

対話システムに対して入力された文を形態素解析し,ベクトル化したものを図 2 のように学習済みモデルへ入力する.学習済みモデルは受け取ったベクトルからラベル番号(感情)を出力する.

<sup>†</sup>東京情報大学 Tokyo University of Information Sciences

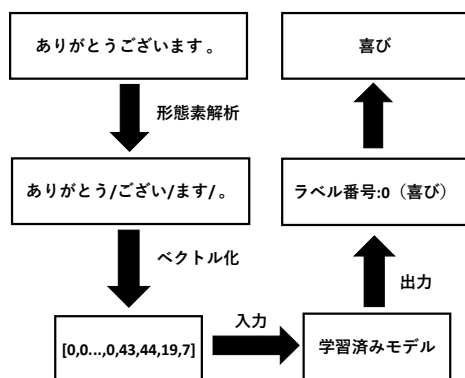


図2 感情判断器の動作

人と対話システムの、会話から感情判断までの流れを図3に示す。感情判断器は、対話システムから入力された会話文に対して、会話文から感情を判断しその結果を感情値に変換し、対話システムに返している。

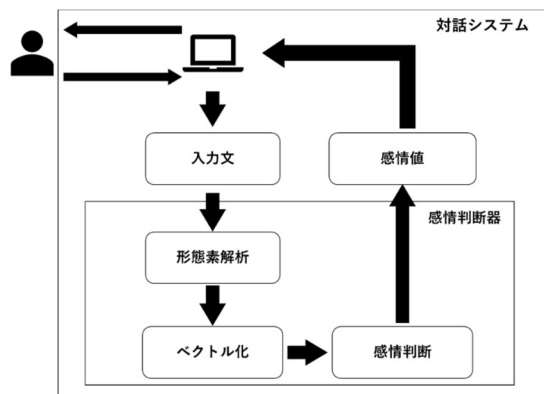


図3 対話システムと感情判断器の関係

#### 4. 評価実験

本研究で利用するデータセットの文中には単語間をつなぐための接続語が多く存在する。接続語は、少数の接続後が多く、多くの文で出現するため、感情判断においてそのままでは有益な情報にはなりえないと予想した。この予想の有効性を確認するため、接続語を含めたデータセットと、接続語を除いたデータセットの学習精度を比較した。

##### 4.1 実験データ

表2 接続語の加工例

データセット	朝 <u>は</u> きり <u>さ</u> め <u>で</u> あった
加工後	朝 きりさめ あった

表2はデータセットから接続語を除いた例であり、実際には空白は存在しない。

データセットの総単語数 1,722,475 件のうち 28.1%が、接続語の総数である。接続語はデータセットの中に 484,214 件存在し、データセットから取り除いた。

また、データセットに重複データが多く存在する場合データが偏って学習されてしまうため、重複データはデータ

セットから除いた。データセットの 79,821 件のうち、重複データは 21,202 件 (26.6%) 存在した。実験で使用するデータセットは重複データを除いた 58,619 件となる。

#### 4.2 実験結果

接続語を除いていないデータセットを用いて感情判断器に学習させた学習精度の出力と、接続語を除いたデータセットを用いて感情判断器に学習させた学習精度を出力した。

表3 実験結果

	接続語あり	接続語なし
正解率	60.5%	62.5%

接続語を含むデータセットの場合、テストデータでの正解率は 60.5%であった。それに対しデータセットから接続語を除いた場合の同じテストデータでの正解率は 62.5%であった。結果として接続語を除くことで 2%の精度向上となった。しかしこれ以上の精度向上はデータセットのチューニングだけでは難しいと考える。さらに精度を上げるためには、感情判断手法そのものを改善する必要がある。

#### 5. 結論

本研究は、人との自然なコミュニケーションを行う対話システムに必要な感情判断器の開発を目指している。本論文では、接続語をそのまま利用した場合、文の感情判断において学習に悪影響を与えると予想し、実験で確認した。

データセットから接続語を除くことにより 2%の精度向上が確認された。この結果により、接続語をそのまま分類に利用することは分類精度に悪影響を与えることを示せた。

今後は分類精度を向上するため、ベクトル化手法そのものを再検討し、現在の Keras トークナイズと word2vec を用いた 2 種類のベクトル化手法の比較を行う。

また、本論文の感情判断器に使われているデータセットに、感情表現辞書を組み合わせることを予定している。これにより単語単位の感情解析と文単位の感情解析の両方を用いた分類手法を検討することで、感情判断器としての精度向上が期待できると考える。

#### 参考文献

- [1] 山下紗苗, 上泰, 奥村紀之. 日本語文章からの感情推定難易度判定. 人工知能会第 33 回全国大会, 4M3-J-9-03 (2019)
- [2] Michal Ptaszynski, Pawel Dybala, Wenhan Shi, Rafal Rzepka, Kenji Araki. A system for affect analysis of utterances in Japanese supported with web mining. 日本知能情報ファジィ学会誌, vol.21, No.2, pp.194-213 (2009)
- [3] 山本裕司, 吉村枝里子, 土屋誠司, 渡部広一. 会話文章からの話者の意図・感情判断. 情報処理学会研究報告, vol.2010-ICS-158, No.2 (2010)
- [4] 盛舒峰, 渡辺裕. コミックのセリフの感情分析. 電子通信学会基礎・境界ソサイエティ/NOLTA ソサイエティ大会, A-10-18 (2018)
- [5] Robert Plutchik. "The nature of emotions". American Scientist, Vol.89, Iss.4, pp. 344-350 (2001)
- [6] テキストの前処理-Keras Documentation. <https://keras.io/ja/preprocessing/text/>
- [7] 大澤泰我, ケネス マッキン, 永井保夫. 対話システムのための感情判断器の検討. 情報処理学会第 82 回全国大会. 1S-06 (2020)