

聴解テストから得られたラベルセットによる
アニメーション会話の属性の
ゼロショット分類の検証

Ni Yangdi, Shan Junjie, 西原陽子
立命館大学情報理工学部

研究背景：テキスト分類

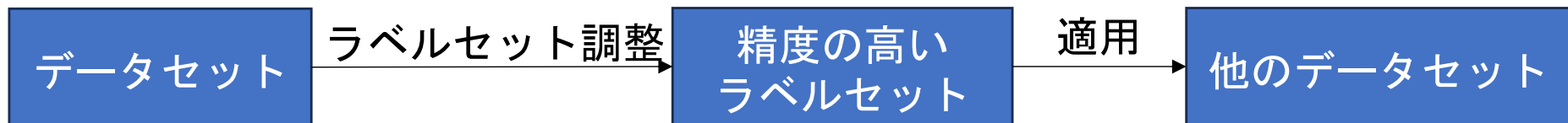
- ・ 与えられたテキストデータセット内のテキストを属性ごとに分類することをテキスト分類と呼ぶ
 - ・ メールのスパム分類, センチメント分析, トピック分類など



- ・ 従来の機械学習の方法では, 以下の制限があった:
 - ・ 学習のために訓練データを作成する必要がある
 - ・ 分類タスクごとにモデルを作成する必要がある

研究背景：ゼロショット分類

- ゼロショット分類はラベルセットとテキストを入力すると、テキストを自動分類できる。訓練データの作成が不要
- ゼロショット分類は入力するテキストが同じでも、入力するラベルセットが変わると、出力結果が変化する。
- ラベルセットの調整により、ゼロショット分類の精度が向上する [Ni, 2024]。
- **しかし、調整されたラベルセットの他のデータセットに対する精度は検証されていない。**

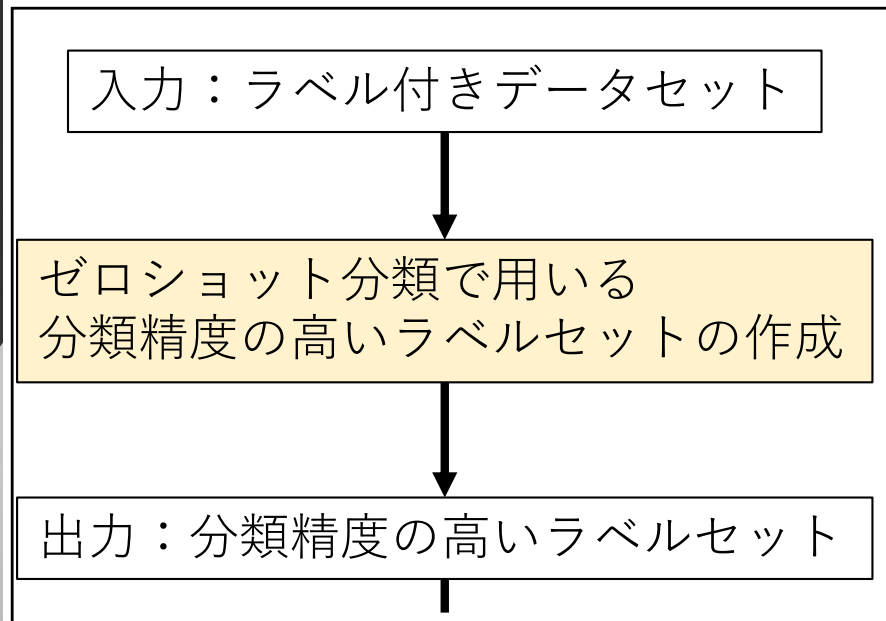


研究の目的

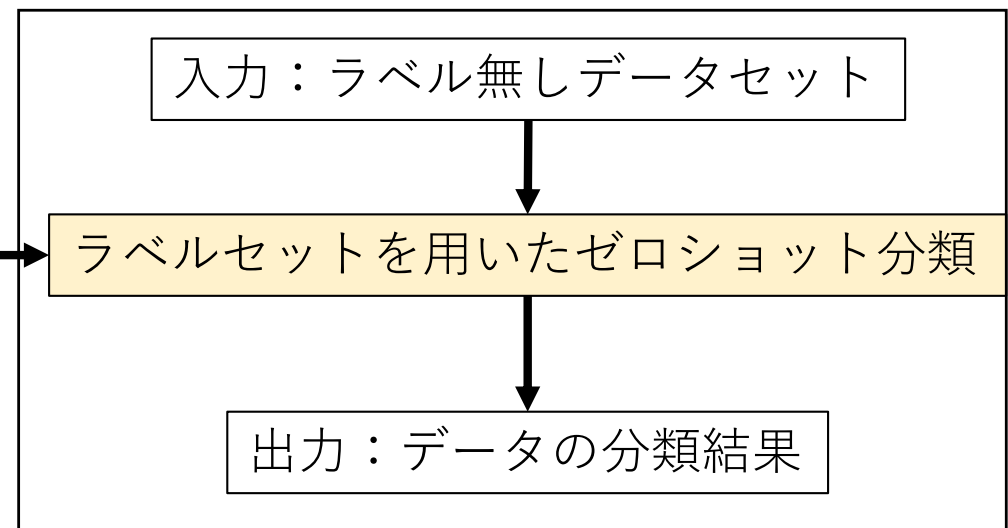
- 本研究では、1つのデータセットから得られたゼロショット分類のラベルセットが他のデータセットにも適用可能であるかを評価する
 - 得られたラベルセットを他のデータセットに適用した際の分類精度を評価する

提案手法

第1段階：ラベルセットの作成



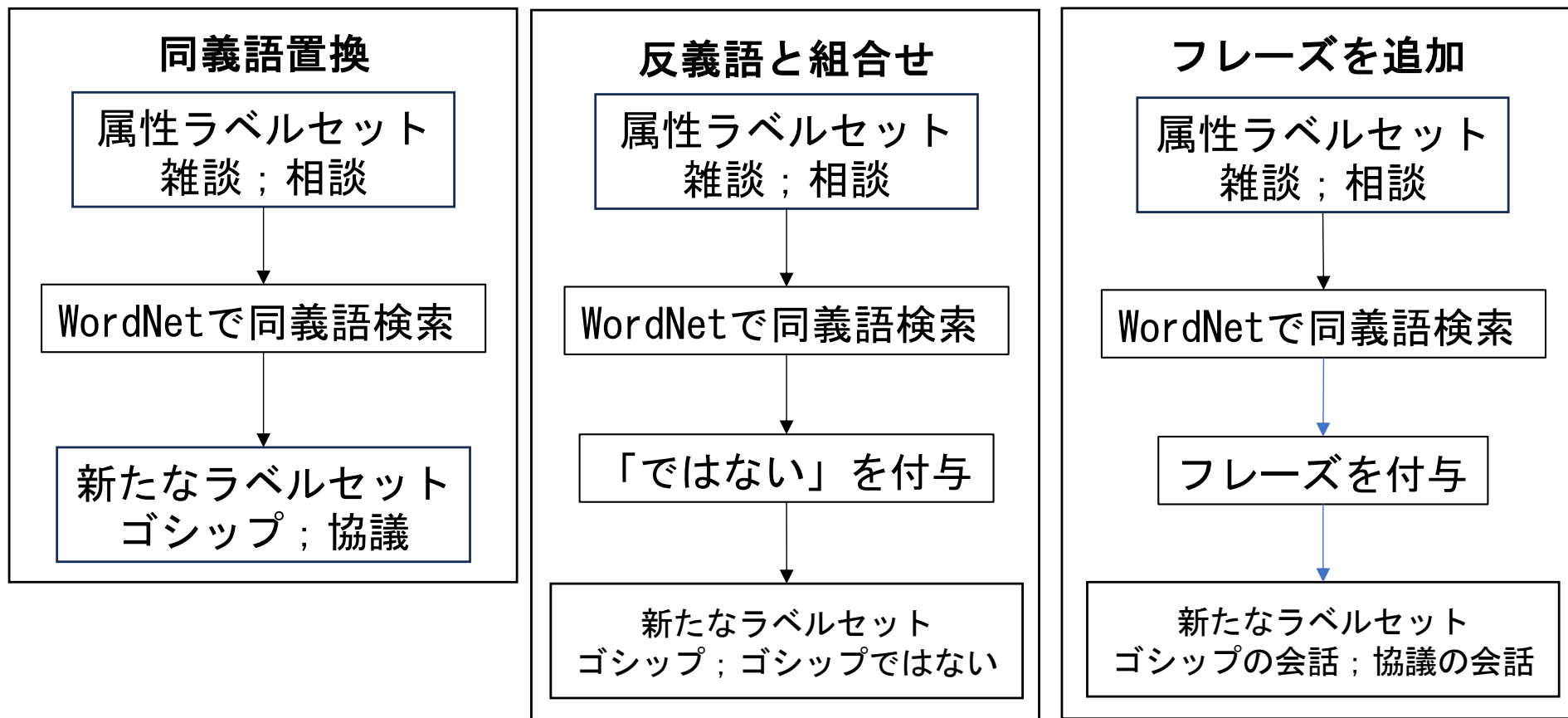
第2段階：ラベルセットを用いた ゼロショット分類



ゼロショット分類で用いるラベルセットの作成

入力するテキストが同じでもラベルセットが異なると結果が変わるため、分類精度の高いラベルセットを作成する

本研究では、分類対象のテキストとして会話テキストを用いた。



付与したフレーズ：「の会話」「間の会話」「的な会話」「で話す」「での会話」

ラベルセットを用いたゼロショット分類

2つのゼロショット分類手法を用いる

(1) テキストembeddingのcos類似度

入力:

- ・ テキスト
- ・ ラベルセット [知り合い, 初対面]

テキストとラベルをembedding

テキストとラベルのcos類似度を算出

知り合いとのcos類似度 : 0.964
 初対面とのcos類似度 : 0.036

Open-Aiのtextembedding-ada-002を利用

(2) 事前訓練済みモデルを用いたゼロショット分類

入力:

- ・ テキスト
- ・ ラベルセット [知り合い, 初対面]

事前訓練済みモデルを用いてゼロショット分類

知り合いの尤度 : 0.953
 初対面の尤度 : 0.047

MoritzLaurerモデルを利用

Open-Aiのtextembedding-ada-002: <https://platform.openai.com/docs/guides/embeddings>

MoritzLaurer: <https://huggingface.co/MoritzLaurer/mDeBERTa-v3-base-xnli-multilingual-nli-2mil7>

予備実験（分類精度の高いラベルセットの特定）

実験目的：

各ゼロショット分類方法による分類精度の高いラベルセットの特定

手順：

1. 候補のラベルセットを作成する。
 - 「同義語置換」「反義語と組合」「フレーズを追加」で合計162セット
2. 作成されたラベルセットにより、属性ごとにゼロショットで入力テキストを分類する。
3. 分類結果を正解ラベルに比べて、各候補ラベルセットの分類精度を算出。
4. 分類精度の高いラベルセットをゼロショット分類方法ごとに特定する。

評価指標：

二乗平均平方根誤差 (RMSE) : RMSEが小さいほど分類精度が高い

予備実験の使用データ

入力テキスト：（ラベル付き検証テキスト）

250件の日本語能力試験（JLPT）の聴解テスト会話文。

項目	値
会話数	250
最小文字数	37
最大文字数	743
平均文字数	211.5
文字数の標準偏差	124.3

分類属性：

- 会話場所：[自宅, 職場, 学校, 施設, 屋外, 電話]
- 会話形式：[雑談, 用談相談]
- 話者間の関係：[初対面, 知り合い, 友達, 家族]

（「日本語日常会話コーパス」（CEJC）に参考）

属性	元ラベル	数
関係 (4分類)	初対面	49
	知り合い	82
	友達	98
	家族	21
形式 (2分類)	雑談	75
	用談相談	175
場所 (6分類)	自宅	39
	職場	79
	学校	53
	施設	42
	屋外	14
	電話	23

予備実験の結果 (ゼロショット分類手法ごとのラベルセット)

ゼロショット分類方法	属性 (分類数)	分類精度の高いラベルセット	RMSE	RMSE (元ラベルセット)
Open-Aiのembeddingによるコサイン類似度	関係 (4)	[初対面, 知辺, 馴じみ, 親類]	2.74	5.04
	形式 (2)	[雑談, 相談]	3.17	3.21
	場所 (6)	[自宅で話す, 職場で話す, 学校で話す, 施設で話す, 街並みで話す, 電話で話す]	2.54	3.26
事前訓練ゼロショット分類モデル MoritzLaurer	関係 (4)	[初対面間の会話, 御近付き間の会話, 身方間の会話, 親類間の会話]	2.39	4.22
	形式 (2)	[世間話, 相談]	3.43	3.63
	場所 (6)	[自宅で話す, 職場で話す, 学校で話す, 施設で話す, 街並みで話す, 電話で話す]	1.75	2.32

- 元ラベルセットのRMSEと比べると、提案手法で作成した分類精度の高いラベルセットのRMSEは低であった。ラベルセットの調整により、ゼロショット分類の精度が向上した。
- MoritzLaurerモデルで抽出したラベルセットのRMSEの平均値は2.54であった。Open-Aiのembeddingコサイン類似度での場合のRMSEの平均値は2.81であった。平均値より、MoritzLaurerモデルの方が分類精度が高いラベルセットが抽出されたことがわかった。
- 「形式」の分類においては、Open-Aiのembeddingコサイン類似度を用いた方がRMSEの値が低いラベルセットが抽出された。

ゼロショット分類ラベルの特定

ゼロショット分類方法	モデル：属性(分類数)	分類精度の高いラベルセット
最優モデル： MoritzLaurer	MoritzLaurer：関係(4)	[初対面間の会話, 御近付き間の会話, 身方間の会話, 親類間の会話]
	MoritzLaurer：形式(2)	[世間話, 相談]
	MoritzLaurer：場所(6)	[自宅で話す, 職場で話す, 学校で話す, 施設で話す, 街並みで話す, 電話で話す]
最優ラベル： Open-Ai (形式) + MoritzLaurer (関係、場所)	MoritzLaurer：関係(4)	[初対面間の会話, 御近付き間の会話, 身方間の会話, 親類間の会話]
	Open-Ai：形式(2)	[雑談, 相談]
	MoritzLaurer：場所(6)	[自宅で話す, 職場で話す, 学校で話す, 施設で話す, 街並みで話す, 電話で話す]

属性ごとに特定されたゼロショット分類手法とラベルセット：

1. 最優モデル：
MoritzLaurerで得られたラベルセットで3つの属性を分類
2. 最優ラベル：
MoritzLaurerで得られたラベルセットで「場所」「関係」属性、
Open-Aiのembeddingで得られたラベルセットで「形式」属性を分類

評価実験

実験目的：

ラベル付き検証テキスト（聴解テスト）で特定した「ゼロショット分類手法」と「ラベルセット」の組み合わせの分類効果の確認

手順：

1. 特定した「ゼロショット分類手法」と「ラベルセット」の組合せにより、ラベル無し実験テキストを属性ごとに分類
2. 検証テキストの属性ラベルにより、3種類の属性の組合せに頻出するパターンを抽出
 - 上位10件の頻出パターンまで
3. 抽出した頻出パターンと同様の属性組合せに分類されたラベル無し実験テキストからランダムにサンプリングする
 - 各属性組合せごとに10件ずつサンプリング
4. サンプリングした実験テキストの属性分類精度を確認する
 - 各属性の分類精度、該当する属性組合せと一致した数

実験データ

ラベル無し実験テキスト：
ラベル付き検証テキスト（聴解テスト）
の統計情報により抽出したアニメーション
の会話文。

－ 検証テキストの平均文字数 ± 標準偏差

項目	実験 (アニメーション)	検証 (聴解テスト)
会話数	314, 930	250
最小文字数	87 = 211-124	37
最大文字数	335 = 211+124	743
平均文字数	189. 0	211. 5
文字数の 標準偏差	73. 4	124. 3

検証テキスト（聴解テスト）
に頻出した上位10個の属性組
合せ：（形式, 場所, 関係の順）

順位	属性の組合せ
1	用談相談, 職場, 知り合い
2	用談相談, 施設, 初対面
3	用談相談, 学校, 知り合い
4	雑談, 職場, 友達
5	用談相談, 学校, 友達
6	雑談, 学校, 友達
7	用談相談, 職場, 友達
8	雑談, 自宅, 友達
9	用談相談, 自宅, 家族
10	用談相談, 電話, 初対面

属性の一致数の評価方法

1つのアニメ会話シーンのテキストに対し、3種類の属性で分類を行った。

よし後はタイミングだタイミング
なに一人でテンション上がってんだ
お前高校の学園祭だぞ
とくに代わり映えしない彼も何故か格
好良く見え
恋のカオスと言われるあの宴だぞ
そしてラストを締めくくる後夜祭
...

ゼロショットで分類された結果：
場所：学校 ○
人物間の関係：友達 ○
形式：用談相談 ×

一致する属性数は2種類である

実験結果（各属性分類での正解率）

ゼロショット分類方法	「形式」の 分類正解率	「場所」の 分類正解率	「関係」の 分類正解率	平均正解率
最優モデル： MoritzLaurer	81.0%	53.0%	58.0%	<u>64.0%</u>
最優ラベル： Open-Ai+MoritzLaurer	75.0%	54.0%	51.0%	60.0%
属性平均正解率	<u>78.0%</u>	53.5%	54.5%	62.0%
チャンス確率	50.0%(2分類)	16.7%(6分類)	25.0%(4分類)	—

- 属性「形式」：2分類
チャンス確率(50.0%)より28.0% (=78.0%-50.0%) 上昇
- 属性「関係」：4分類
チャンス確率(25.0%)より29.5% (=54.5%-25.0%) 上昇
- 属性「場所」：6分類
チャンス確率(16.7%)より36.8% (=53.5%-16.7%) 上昇

実験結果（該当パターンと一致した属性数）

ゼロショット分類方法	一致した属性の種類数での正解率(%)			合計
	1種類	2種類	3種類	
最優モデル： MoritzLaurer	31.0%	36.0%	27.0%	<u>94.0%</u>
最優ラベル： Open-Ai + MoritzLaurer	33.0%	29.0%	21.0%	83.0%

- 2種類以上の属性が一致した割合：
 - 最優モデル：63.0% (=36.0%+27.0%)
 - 最優ラベル：50.0% (=29.0%+21.0%)

まとめ

本研究では、分類ラベルセットの記述を変えることにより、テキスト属性のゼロショット分類の精度を改善する手法を提案した。

提案手法では、「同義語置換」「反義語と組合」と「フレーズを追加」の3つの方法で記述の異なるゼロショット分類のラベルセットを作成し、分類効果を評価した。

2つのゼロショット分類方法と、3つのテキスト属性で分類精度の高いラベルセットをラベル付き検証テキストにより特定し、ラベル無し実験テキストにより分類精度を確認した。

今後、同じ属性を持つテキスト間のマッチング・推薦の実験を行う予定。