

AAAI2021

# DDRel: A New Dataset for Interpersonal Relation Classification in Dyadic Dialogues

Qi Jia, Hongru Huang, Kenny Q. Zhu

JAIST 情報科学系 修士1年  
林 貴斗 (Hayashi Takato)

DDRelは対話を用いた社会的関係分類データセットである.

DDRelは複数の映画脚本から収集した694組の話者の6300の対話セッションが含まれている

対話者との社会的関係によって話し方や話題などが変化すると考えられる. また, 同じような表現でも社会的関係が異なれば, 異なる感情や態度が反映されることもある.

このような特徴を考慮して, 社会的関係性の予測が可能になれば, 社会的関係に応じて適切な応答を生成する対話システムやチャットの推薦システムなどに応用できる.

本論文では, CNN,LSTM,BERTなどのモデルを使って, 対話を用いた社会的関係分類のためのSOTAモデルの構築も行っている.

# Task Definition

本論文のタスクは対話テキストから対話者間の社会的関係を推定することである。  
各対話には、社会的関係が $R = \{R_1, R_2, \dots, R_m\}$  ( $m$ は関係の種類数)としてアノテーションされている。また、同一の対話ペアは複数セッションの対話を持っていて、それらは同一の関係がアノテーションされている。

単一の対話から社会的関係を推定する場合は

$$R_j^i = \arg \max_R f_s(D_j^i) \quad \cdot \cdot \cdot (1)$$

ただし、対話ペア $i$ の $j$ 個目の対話を $D_j^i$ とする。

また同一対話ペアのすべての対話から社会的関係を予測する場合は

$$R^i = \arg \max_R f_p(D_i) = \arg \max_R f_p(D_1, D_2, \dots, D_n) \quad \cdot \cdot \cdot (2)$$

ただし、対話ペア $i$ は $n$ 個の対話を持つものとする。

- ・フィルタリング

IMSDBから941本の映画脚本を収集し, 「IMDB1に登録されていないもの」, 「英語ではないもの」, 「非常に人気のない作品」を除外する.

「2人の話者が3人目に遮られることなく交互に話す」対話を対話セッションとして定義する.  
また, ペア間の合計発話数が20以下のものを取り除く

- ・アノテーション

アノテータは映画のタイトル, 登場人物情報, インターネットへのアクセス権が与えられ, 各対話ペアに13種類の関係または該当なし (NA)のいずれかのラベルをアノテーションする. 関係性が明確なペアにだけ, 関係性をアノテートするように指示され, 関係性のクラスがアノテートされたのは全対話セッションのうち47%であった. 2人目のアノテータに同じ設定で100ペアだけアノテーションさせ, 最初のアノテーションとの一致率を算出したら82.3%であった.

4 classes	6 classes	13 classes	# Sessions	% Sessions	# Pairs	% Pairs	# Turns	% Turns
Family	Elder-Junior	Child-Parent	414	6.57	67	9.65	3,377	6.36
		Child-Other Family Elder	91	1.44	12	1.73	632	1.19
	Family Peer	Siblings	211	3.35	27	3.89	1,585	2.98
		Spouse	568	9.02	51	7.34	4,784	9.01
Intimacy	Intimacy	Lovers	1,852	29.40	244	20.75	17,474	32.89
		Courtship	146	2.32	15	2.16	1,323	2.49
Others	Peer	Friends	1,049	16.65	124	17.87	8,900	16.75
		Neighbors	21	0.33	2	0.29	189	0.36
		Roommates	120	1.90	8	1.15	966	1.82
Official	Elder-Junior	Workplace Superior-Subordinate	536	8.51	79	11.38	3,958	7.45
		Colleague/Partners	710	11.27	76	10.95	5,455	10.27
	Official Peer	Opponents	203	3.22	33	4.76	1,532	2.88
		Professional Contact	56	8.07	56	8.07	2,952	5.56

## ・社会的関係

(Rais and Sprecher 2009)を基に13種類の社会関係を定義した。

さらに社会的分類，年功序列，話者間の親密度を考慮して，6クラス分類と4クラス分類の大きい分類も定義した。

	train	development	test
# Pair of Speakers	541	75	78
# Sessions	5,037	653	610
# Turns	42,564	5,210	5,352
Sessions per pair (mean)	9.31	8.71	7.82
Sessions per pair (std)	8.18	6.35	5.96
Turns per session (mean)	8.45	7.98	8.77
Turns per session (std)	6.96	5.60	7.93

- 統計

DDRelには300本の映画から集められた694組の対話ペアの6300セッションの対話が含まれている。対話ペア毎のセッション数は平均9,標準偏差は7でばらつきは大きい。セッション毎の発話数は8.4, 標準偏差は6.9であった。

社会的関係の分布は, 「恋人」「友人」「同僚・パートナー」の上位3つがデータセット全体の約半数を占めており, 一方で最小データ数の「隣人」は2ペアと21セッションしか含まれていない

- ・ベースラインモデル

CNN, LSTM, BERTを用いてベースラインモデルを構築する.  
同一対話ペアのすべての対話から社会的関係を予測する場合, 各関係の信頼度はセッション間の平均MRRとみなし, 信頼度が最大の関係を最終的な予測とする.

- ・人間の評価

人間が対話からどれくらい関係を推定できるか図るために人間による評価を追加する.  
2人のボランティアには対話レベルの13種類クラス分類を, 他の2人のボランティアにはペアレベルの13種類クラス分類を行ってもらい, その結果を人間の評価とした.

- ・評価指標

評価指標にはaccuracyとF1-Scoreを用いた.

- ・ベースラインモデル

CNN, LSTM, BERTを用いてベースラインモデルを構築する.  
同一対話ペアのすべての対話から社会的関係を予測する場合, 各関係の信頼度はセッション間の平均MRRとみなし, 信頼度が最大の関係を最終的な予測とする.

- ・人間の評価

人間が対話からどれくらい関係を推定できるか図るために人間による評価を追加する.  
2人のボランティアには対話レベルの13種類クラス分類を, 他の2人のボランティアにはペアレベルの13種類クラス分類を行ってもらい, その結果を人間の評価とした.

- ・評価指標

評価指標にはaccuracyとF1-Scoreを用いた.