

マルチモーダル情報を用いたソーシャルスキルの客観的推定

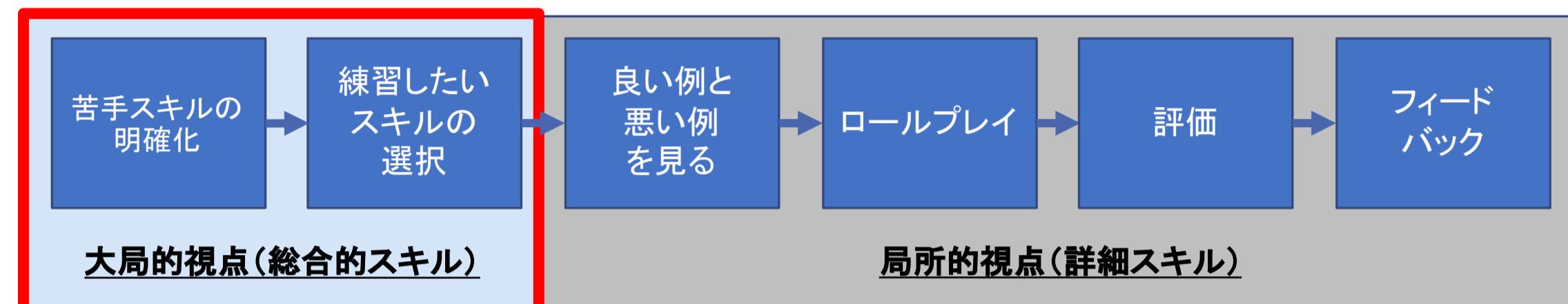
*佐賀健志、*田中宏季、**岩坂英巳、*中村哲

*奈良先端科学技術大学院大学、**奈良県立医科大学精神医学講座

背景・目的

ソーシャルスキルトレーニング (SST)

- 目的: 適切なソーシャルスキルの獲得
- 問題点: SSTセラピストが少ない → アクセシビリティが低い
- 解決策: SSTの自動化



得意・不得意スキルの評価・判断
 臨床現場: 対話を通じたコミュニケーション
 自動化システム: 標準化された評価指標 + 機械学習

想定する応用先
 ・好きなときに使える自動SSTアプリ
 ・簡易スクリーニング

データセット [1]

- 対象: 学生
- 被験者数: 27人 (男21、女6)
- タスク: 最近の楽しかった出来事について1分間話す
- 聞き手: バーチャルエージェント
- 書き起こし: 手動
- 年齢: 平均25.1、標準偏差2.13
- SRS (男): 平均65.9、標準偏差19.8
- SRS (女): 平均61.5、標準偏差21.0

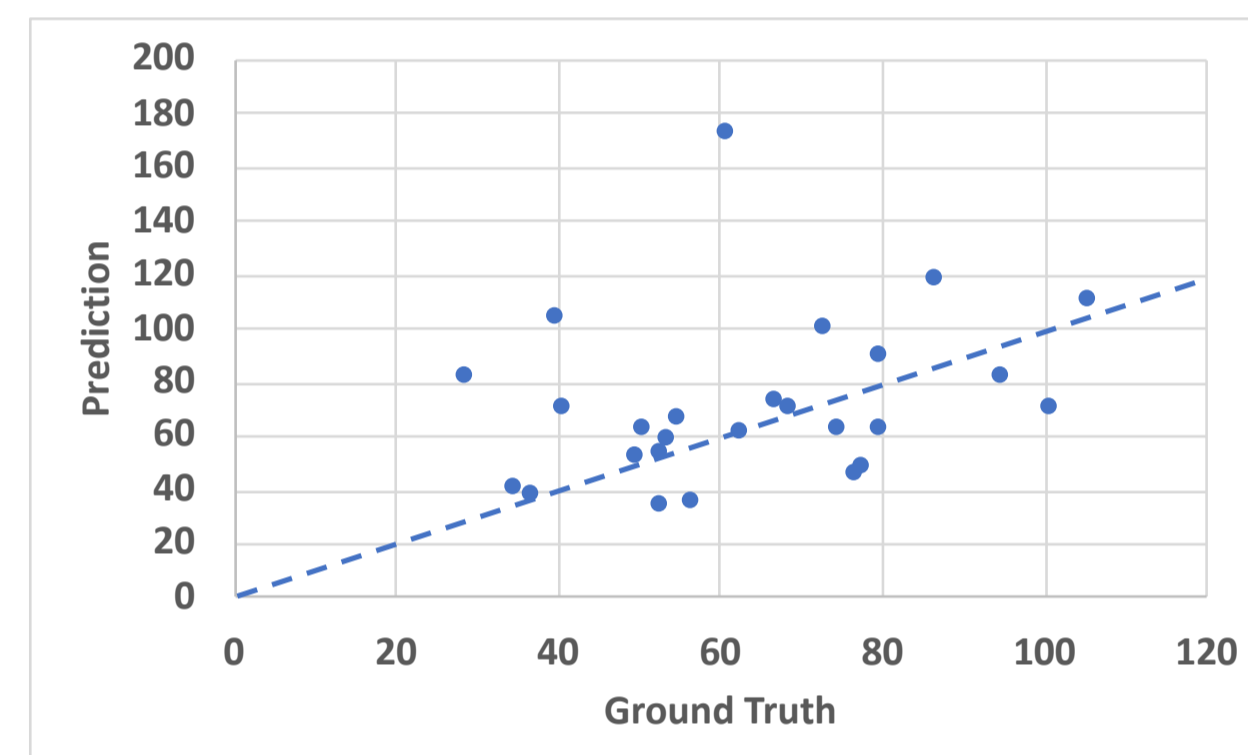
重要なソーシャルスキル「うれしい気持ちを伝える」に相当



実験結果

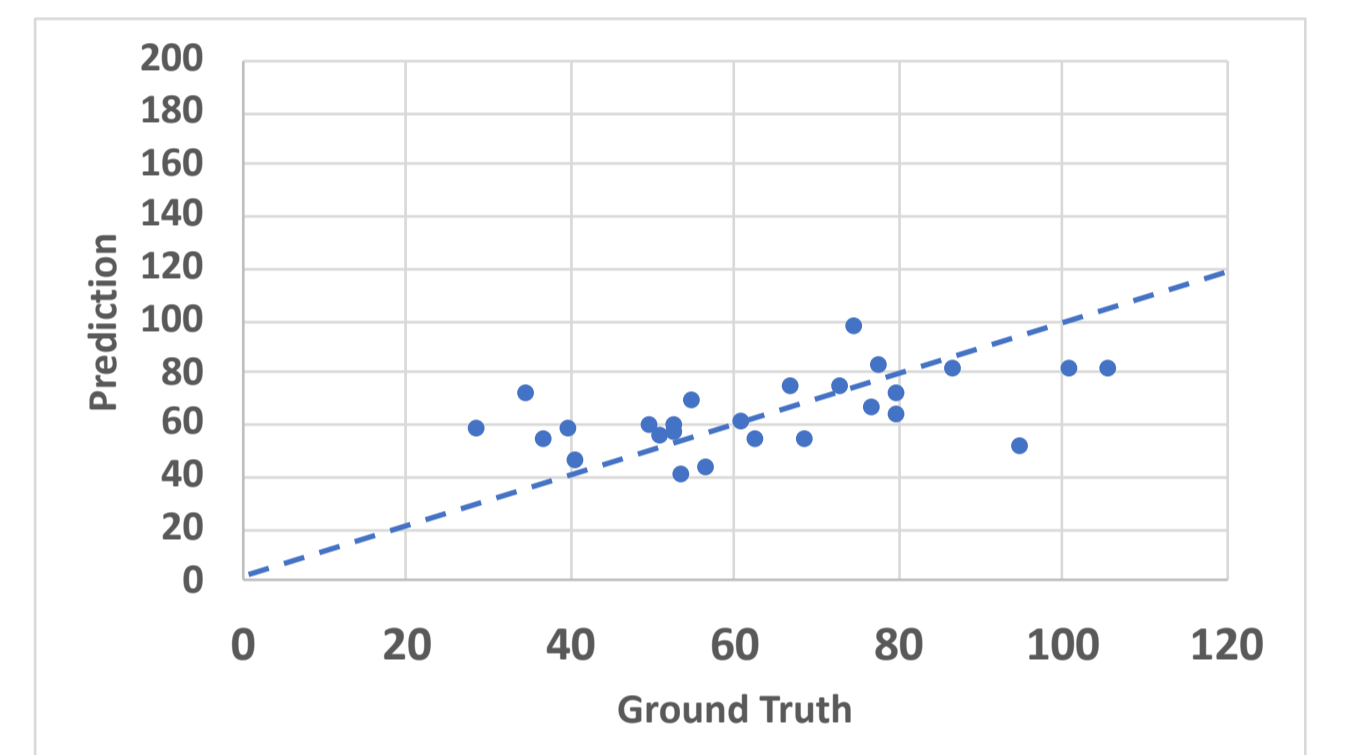
-総合スコア予測-

ベースライン [1]



RMSE標準偏差	14.38
平均RMSE	27.34
Pearson相関係数	0.35
p値	0.08

提案手法



RMSE標準偏差	10.71
平均RMSE	17.41
Pearson相関係数	0.52
p値	0.01

対人応答性尺度 (SRS-2) [3]

- 65問で構成される質問紙 (標準化済み)
 - **自閉症傾向の重症度に関する評価尺度 (健常者にも応用可)**
 - 治療下位尺度も算出可能 (総合スコアはそれらの総和)
- 1) 社会的気づき (8項目)
 - 2) 社会的認知 (12項目)
 - 3) **社会的コミュニケーション (22項目)**
 - 4) 社会的動機づけ (11項目)
 - 5) 興味の限局と反復行動 (12項目)
- (自閉症傾向が高ければSRSは高くなる。低いほどソーシャルスキルが高い)

スコアの最大幅
0-195

提案手法

-新規特徴量を導入-

機械学習モデル: 線形回帰

学習データ: 1分間の発話動画 (n=27, 健康な被験者) [4]

入力特徴量: 28 (音声) [4] + 5 (テキスト) + 22 (視覚) = 55

新規導入した特徴量

- 1) 面接評価システムで使用されていた音声特徴量の応用
- 2) BERT埋め込みに基づく発話系列類似性 (Seq-similarity)

入力特徴量は平均0分散1に正規化

考察

-重要特徴量候補を発見-

総合スコア予測 (モデル係数, 上位20項目)

Name	Coefficients	Yaw	
AU28	-409.6	AU02	-25.7
Seq-similarity	-115.6	Smile Ratio	-12.7
Pitch	45.9	B Energy	-12.4
AU25	-42.7	AU07	12.2
AU15	-39.0	AU12	-12.2
AU01	35.5	C F2/F1 Mean	8.0
AU26	-33.6	AU05	6.1
AU45	31.9	AU06+12	-6.0
AU17	-29.3	AU04	4.9
AU09	27.3		

SRS ソーシャルスキル
:高い→低い
:低い→高い

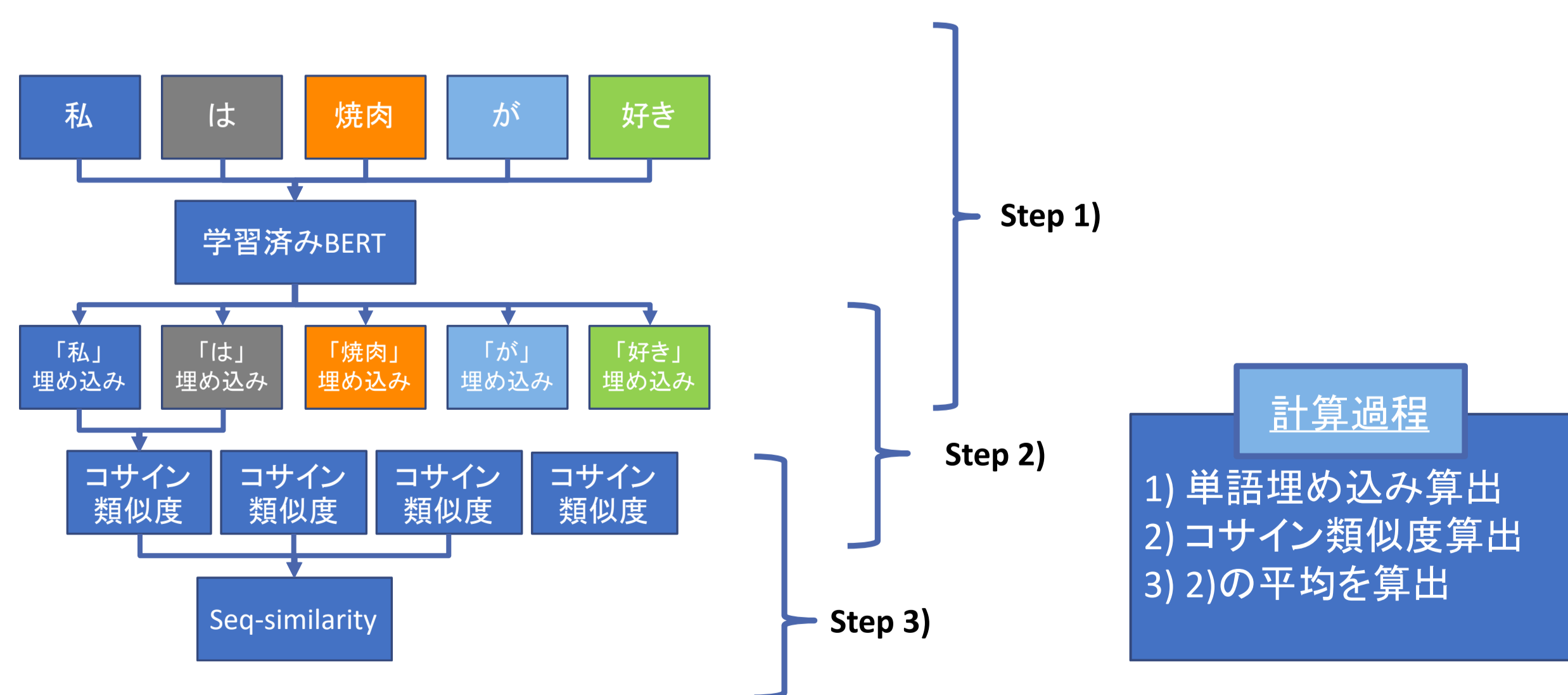
解釈

- A. 意味的一貫性を保って話す
- B. 大きい声で話す
- C. 長母音をより多く使う

予稿から修正点

	Seq-similarity
総合スコア	-0.31
社会的コミュニケーション	-0.42

BERT埋め込みに基づく発話系列類似性(Seq-similarity)



まとめと今後の展望

まとめ

- マルチモーダル特徴量でSRSスコアを相関係数0.52で予測
- BERTに基づく発話系列類似性が重要な特徴量の可能性が示唆

今後の展望

- 精神疾患患者を対象とした検証
- 結果に対する各特徴量の重要度および解釈についての検討
- データ収集時の仮想エージェントによる心理的影響

参考文献

[1] Hiroki Tanaka et al., 2017, "Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders", PLOS ONE 12, 8 (08 2017), 1-15. <https://doi.org/10.1371/journal.pone.0182151>

[2] Mohammad Ali et al., 2018, "A Virtual Conversational Agent for Teens with Autism: Experimental Results and Design Lessons", ArXiv, 1811.03046

[3] John N. Constantino et al., 2012, "Social Responsiveness Scale, Second Edition (SRS-2)", Back. Western Psychological Services.

[4] Iftekhar Naim et al., 2015, "Automated Analysis and Prediction of Job Interview Performance", IEEE Transactions on Affective Computing, 2016. 2614299. <https://doi.org/10.1109/TAFFC.2016.2614299>

[5] Hiroki Tanaka, Hideaki Iwasaka, Hideki Negoro, Satoshi Nakamura, "Analysis of Conversational Listening Skills toward Agent-based Social Skills Training", Journal on Multimodal User Interfaces, volume 14, issue 1, pp.73-82, 2020