

グループディスカッション遠隔指導支援のための SHAP 値による特徴量重要度の分析

小松眞子*1, 高田知裕*1, 根子稚絢*2, 卯木輝彦*3, 敷田幹文*1

*1 高知工科大学, *2 NRI システムテクノ, *3 フォトロン

An Analysis of Student Behavior by SHAP Values for Remote Teaching of Group Discussion

Mako Komatsu*1, Chihiro Takada*1, Chihiro Neshi*2, Teruhiko Unoki*3, Mikifumi Shikida*1

*1 Kochi University of Technology, *2 NRI System Techno, Ltd., *3 Photron Limited

近年、採用試験の一部としてグループディスカッションが実施されており、大学では就職担当の教員がグループディスカッションの指導を行うことがある。遠隔での指導が多い現在、教員が複数グループを指導する場合、対面時のように部屋全体を見回すことが難しく、指導機会を見逃してしまうといった課題がある。このような課題に対して、我々は各グループの指導が必要な場面を教員に提示する遠隔指導支援システムの実現を目指している。本論文では、実現計画の一部として、映像から得られる特徴量を用いて学生の評価を3段階に分類するモデルを作成し、分類に影響が大きく、支援システムに利用可能な特徴量を SHAP 値分析によって明らかにする。

キーワード: グループディスカッション, 遠隔指導, 映像分析, 機械学習

1. はじめに

近年、日本では企業の採用試験の一部として、グループディスカッションが実施されている⁽¹⁾。グループディスカッションは、コミュニケーション力や積極性など、書類や面接からは分からない能力を図ることができるため、重要視されている。

大学では、学生を対象に就職担当の教員がグループディスカッションの指導を行うことがある。ディスカッションを行うグループが1つの場合、指導者が十分な指導を行うことが可能である。しかし、複数グループが同時にディスカッションを行う場合、遠隔指導では対面時のように部屋全体を見回すことが難しく、1グループを指導している間に他のグループの指導機会を見逃してしまうといった課題がある。これらの背景から、我々はグループディスカッションに着目した研究を行っている。現在、遠隔での指導機会が増加していることから⁽²⁾、本研究計画では、複数グループのグ

ループディスカッションの遠隔指導を対象とする。

これらの課題を解決するために、本計画ではセンサーなどの特殊なデバイスや大掛かりな設備を使用せず、指導教員に対して各グループの指導が必要な場面を提示する支援システムの実現を目指す。この計画の一部として、本論文ではカメラ映像のみから得られる特徴量を用いて、学生の評価を合格、ボーダーライン、不合格の3つに分類するモデルを作成し、分類に影響の大きい特徴量を SHAP 値分析によって明らかにする。

2. 関連研究

Aran らは、グループの会話の中の参加者の非言語行動から特徴量を抽出し、性格を推定するモデルを構築評価している⁽³⁾。このモデルでは、音声情報など外部デバイスを使用して取得した情報を使用し、高い精度での推定に成功している。また、我々の想定する状況と似た研究として、Muralidhar らは就職面接での評価と非言語特徴量の相関関係を分析している⁽⁴⁾。こ

の研究は一対一での面接を対象としており、特徴量の分析にはセンサーなどのデバイスを使用している。このように、分類モデルや特徴量計測に関連した研究は数多くあり、それらの多くはセンサーなどのデバイスを用いている。本計画では、教育現場での実際の運用を想定し、専用のデバイスや設備が不要な支援を目指すため、参加者の上半身が映った対面グループディスカッションの映像を使用し、映像から得られる特徴量によって分類推定が可能であるかを検討する。

3. 実験

3.1 実験概要

本実験では、グループディスカッション中の映像から特徴量を計測し、採用試験の結果として学生を合格、ボーダーライン、不合格の3つに分類するモデルの作成を行う。また、SHAP 値分析を行い、分類モデルの分類に影響度の大きい特徴量は何かを明らかにする。

3.2 グループディスカッション映像

本実験では、男性2名、女性2名の4名を1グループとして、30分間決められたテーマについてディスカッションを行い、グループ内で意見をまとめて発表する様子を撮影した映像を使用する。ディスカッション参加者は、男性4人女性4人の20代の大学生計8名で、2つのグループのディスカッションをiPadで撮影

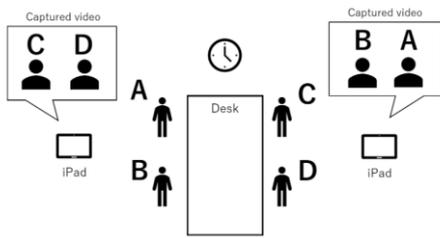


図 1 グループディスカッションの様子



図 2 使用した映像の一部

している。図1はディスカッションの様子を表した図である。実際に使用した映像の一部が図2である。

3.3 特徴量の計測と学生評価

グループディスカッションの30分の動画を10秒間のスロットに分割し、スロット毎に実験参加者の特徴量の計測と学習者としての評価を行う。このとき、スロット毎に「発言者」と「非発言者」に分類して、それぞれの項目で計測、評価を行う。これらの工程は全て人手により行う。

特徴量の計測では、映像から分かる特徴量を計測する。X+数字と特徴量の対応を表1に示す。各特徴量の計測項目は表2に示す。この特徴量は、映像と音声のみから計測可能な客観的指標とする。表1より、発言者に12項目(X1~X12)、非発言者に18項目(X1~X10、X13~X20)の指標を設定し、5段階でスロット毎に発言者1名、非発言者3名を対象に各項目を計測する。

学習者としての評価では、それぞれの実験参加者を表3の項目により評価する。これは指導者の視点から見た、協調性や他者への配慮などコミュニケーション能力についての評価である。「他の人に伝えようと努力しているか」など発言者に10項目、非発言者に5項目の評価を設定し、「非常にあてはまる」を5、「当てはまらない」を1とする5段階でスロット毎に発言者1人と非発言者3人の各項目を評価する。

表 1 特徴量

発言者	非発言者	番号	特徴量
✓	✓	X1	笑顔の度合い
✓	✓	X2	笑顔の時間
✓	✓	X3	片方の肘をついている時間
✓	✓	X4	両方の肘をついている時間
✓	✓	X5	顔や髪を手で触っている時間
✓	✓	X6	腕を組んでいる時間
✓	✓	X7	ペンを両手で持っている時間
✓	✓	X8	ペンを回している時間
✓	✓	X9	手を動かしている時間
✓	✓	X10	発言者以外とのアイコンタクトの回数
✓		X11	発言の音量
✓		X12	発言の時間
	✓	X13	顔の左右の向き (発言者を0度)
	✓	X14	顔の上下の向き (発言者を0度)
	✓	X15	うなずきの回数
	✓	X16	首を傾けた回数
	✓	X17	何らかの意思表示以外の顔の向きの度合い
	✓	X18	メモをしている時間
	✓	X19	メモ中に顔を挙げて発言者を見た回数
	✓	X20	口を閉じている時間

表 2 特徴量の計測項目

対象者	項目	数字の意味				
		1	2	3	4	5
全体	X1	真顔（笑顔でない）	口角が若干上がっている	口角が少し上がっている	口角が上がっている	歯が見える
	X2	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
	X3	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
	X4	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
	X5	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
	X6	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
	X7	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
	X8	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
	X9	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
	X10	0 回	1 回	2 回	3 回	4 回以上
発言者	X11	注意しないと聞こえない	少し聞き取りづらい	普通（日常会話程度）	少し大きめに話している	はっきりと聞こえる
	X12	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
非発言者	X13	それ以上または発言者不在	約+90 度	約+60 度	約+30 度	約 0 度
	X14	約+60 度または発言者不在		約+30 度		約 0 度
	X15	0 回	1 回	2 回	3 回	4 回以上
	X16	0 回	1 回	2 回	3 回	4 回以上
	X17	あまり動いてない	少し動いている	動いている	大きく動いている	激しく動いている
	X18	0 秒	0～2 秒	2～4 秒	4～6 秒	6 秒超過
	X19	0 回	1 回	2 回	3 回	4 回以上
	X20	0～2 秒	2～4 秒	4～6 秒	6～8 秒	8 秒超過

表 3 学習者としての評価項目

対象者	項目
発言者	他の人に伝えようと努力しているか
	論理的であるか
	話の内容が理解しやすいか
	話の流れを無視していないか
	議論を推し進めるような発言ができたか
	周囲に意見を求めるなど議論の展開に貢献できたか
	様々な意見をまとめて整理できたか
	議論の修正ができたか
	場を和ませるなどして、話しやすい雰囲気を作り出せたか
	頭ごなしに人の意見を否定していないか
非発言者	会議に積極的に参加しようとしているか
	注意力が散漫でないか
	協調性があるか
	自身の意思を主張していたか
	好感が持てるか

3.4 分類モデルと SHAP 値分析

本論文では、3.3 節で得た特徴量と評価を用いて、採用試験の結果として学生を「合格」、「ボーダーライン」、「不合格」の 3 段階に分類するモデルを作成する。まず、発言者と非発言者のそれぞれで合計値の分布図を作成する。この時の合計値とは、表 2 に示したスロット毎の特徴量計測項目の点数の合計である。この分布図を 3 段階に分け、分類推定をする機械学習モデルを作成する。このとき、合計値のデータの 75% をトレーニングデータ、25% をテスト用データとして XGboost を用いて作成する。XGboost は、機械学習アルゴリズムの 1 つである勾配ブースティングのライブラリであ

る⁵⁾。その後、SHAP 値を用いて、特徴量重要度の評価を行う。SHAP (SHapley Additive exPlanations) 値とは、ゲーム理論を応用した、機械学習の推定結果の理由を説明する方法の 1 つである⁶⁾。SHAP 値を用いることによって、分類モデルの分類結果に各特徴量がどの程度影響しているかを表す特徴量重要度を求めることができる。

4. 結果

4.1 分類モデル

XGboost で作成したモデルの Confusion Matrix を表 4、表 5 に示す。A が合格、B がボーダーライン、C が不合格を示している。横軸の Predicted の A, B, C が作成モデルの予測した 3 分類であり、縦軸の Actual の A, B, C が実際の 3 分類である。表 4 を例とすると、横軸 Predicted A、縦軸 Actual A のマスに

表 4 発言者の Confusion Matrix

		Predicted		
		A	B	C
Actual	A	3	10	0
	B	2	48	4
	C	1	6	4

表 5 非発言者の Confusion Matrix

		Predicted		
		A	B	C
Actual	A	62	40	1
	B	30	62	6
	C	1	19	60

ある 3 という数字は、分類モデルが学生を合格と分類して、実際に正解した個数である。分類モデルの正解率は発言者が 71%、非発言者が 65%と、支援システムの実現に有用な精度で分類できているといえる。

4.2 特徴量重要度

各特徴量が分類モデルの分類にどれほど影響しているのか、SHAP 値を用いて分析した結果を図 3、図 4 に示す。図の上部、横軸の値が大きいほど分類への影響度が高く、重要な特徴量であるといえる。各棒グラフの色は特徴量が 3 段階の分類にどの程度影響したかを示している。図 3 を例とすると、最も特徴量重要度が大きい X9「手を動かしている時間」は A の分類に最も大きく影響しており、B と C の分類には同程度影響しているということが分かる。発言者、非発言者どちらにおいても、A、C のいずれかに特に影響している項目はなく、どちらも同等な項目が多い。このことから、良い評価のみに影響する特徴量、悪い評価のみに影響するといった特徴量はないということがいえる。

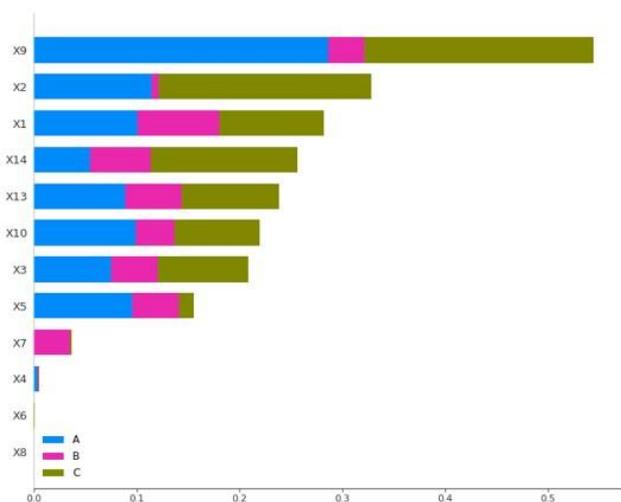


図 3 発言者の SHAP 値

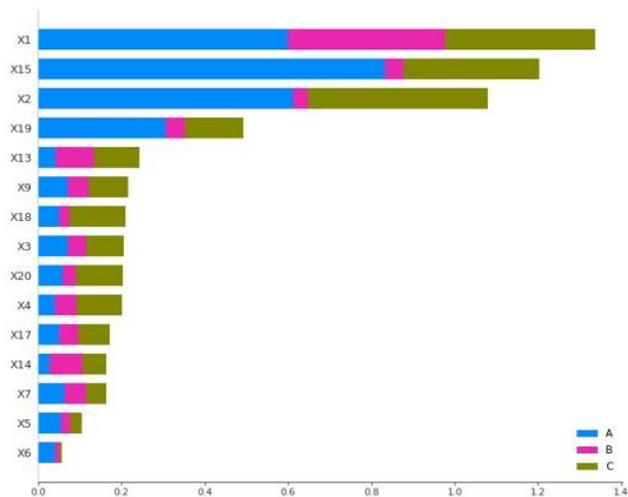


図 4 非発言者の SHAP 値

5. 考察

5.1 発言者

図 3 より、「手を動かしている時間」が発言者の分類推定に最も影響している特徴量であり、「笑顔の時間」、「笑顔の度合い」などの項目も分類に影響を与えていることが分かる。この結果から、グループディスカッションにおいて、発言者が手振りを使って意見を伝えようとしているか、他の参加者に対して笑いかけようかといった特徴量が重要であるといえる。これらの特徴量が計測できた場면을指導教員に知らせることで、評価を行う際の支援として役立てることができる。

その一方で、「ペンを両手で持っている時間」、「両方の肘をついている時間割合」、「腕を組んでいる時間」、「ペンを回している時間」の項目は、分類に対する影響が低いという結果になった。このことから、影響の低い項目は支援システムに利用する必要がないといえる。

5.2 非発言者

図 4 より、「笑顔の度合い」、「うなずきの回数」、「笑顔の時間」が非発言者の分類推定に大きく影響している特徴量であることが分かる。この結果から、非発言者が発言者のほうを向いて話を聞いているか、うなずきながら話を聞いているかといった特徴量が重要であるといえる。これらの非発言者の特徴量が計測できた場면을指導教員に知らせることで、評価の支援として

役立てることができる。

一方で、「ペンを両手で持っている時間」、「顔や髪を手で触っている時間」、「腕を組んでいる時間」は分類に対する影響が低い特徴量であることが分かった。これらの特徴量は、支援システムには利用する必要がないといえる。

5.3 支援システム実現に向けて

本論文では、映像から得られる特徴量から作成したモデルを使って、3段階評価の分類推定ができることが明らかになった。評価の分類結果や重要度の高い特徴量を計測した場面を指導教員に知らせることで、教員がグループディスカッションの映像を見て学生の評価を行う際に、指導機会を見逃してしまうといった課題への支援に役立てられると考える。

支援システムが実現した場合、指導教員はグループディスカッションを撮影した映像全体から、システムによって指導が必要な場面を知ることができるため、通知された場面の前後を含めた映像を見ることによって、すべての映像を見て指導機会を探すことや学生を評価する負担が低減できると考える。また、リアルタイムの支援が実現できると、学生と同じ教室に補助教員がいた場合、指導すべき場面があったという通知を受け取り、遠隔から指導教員が補助教員へグループの様子を見るように伝えるなど、支援システムによってより指導しやすくすることが可能ではないかと考える。

4.1 節より、本稿で作成した分類モデルは指導機会が必要な場面を通知するような支援システムの実現に利用可能であるといえる精度で分類できていることが分かる。また、計測した特徴量は、映像から手や顔の検出、音声認識などを用いることで自動的に計測することが可能であると考えられる。このことから、分類モデルと合わせることにより、自動的な支援、リアルタイムでの支援を実現できる可能性があるといえる。また、さらなる精度向上のために、その他の特徴量項目も検討すべきであると考えられる。

6. おわりに

本論文では、グループディスカッションの映像から得られる特徴量を用いて分類モデルを作成し、学生の評価の分類推定が可能であるか、分類に影響の大きい

特徴量はなにかについて明らかにした。実験結果より、映像から得られる特徴量を用いて作成した分類モデルは、システムに利用可能と言える精度の分類推定が可能であることと、分類に影響の高い特徴量が何かを明らかにすることができた。これらの結果は、指導者のグループディスカッションに対する遠隔指導を支援するシステムの実現において、映像から得られる特徴量に有用性がある事を示している。また、本実験では大学生のグループディスカッションを対象としたが、今後は他の場面への適用も視野に入れ、データ数を増加させた分析や支援システムの実現を目指す。

参考文献

- (1) マイナビ 2020 年卒学生就職モニター調査 5 月の活動状況, https://saponet.mynavi.jp/wp/wp-content/uploads/2019/06/monitor2020_4.pdf (2021 年 4 月 14 日確認)
- (2) 大学等における後期授業の実施方針の調査について, https://www.mext.go.jp/content/20200915_mxt_kouhou01-000004520_1.pdf (2021 年 4 月 14 日確認)
- (3) O.Aran and D.Gatica-Perez: “One of a kind: Inferring personality impressions in meetings”, Proceedings of the 15th ACM on International Conference on Multimodal Interaction, pp.11–18(2013).
- (4) S.Muralidhar, L.S.Nguyen, D.Frauendorfer, J.-M.Odobez, et al.: “Training on the job: Behavioral analysis of job interviews in hospitality”, Proceedings of the 18th ACM International Conference on Multimodal Interaction, pp.84–91 (2016)
- (5) Tianqi Chen, Carlos Guestrin: “XGBoost: A Scalable Tree Boosting System”, the 22nd ACM SIGKDD International Conference (2016)
- (6) Scott Lundberg and Su-In Lee: “A Unified Approach to Interpreting Model Predictions”, Advances in Neural Information Processing Systems 30 (NIPS 2017)