労働者の日常的な精神衛生モニタリングに向けた心的状態推定モデルの性能評価

諏訪部 開[1,a)]　前川　哲志[1]　水本　旭洋[2]　諏訪 博彦[1]　安本 慶一[1]

概要：1970 年以降の先進７カ国 (G7) の中で，日本の労働生産性は一貫して最下位である．これを改善するため，労働者が健康にパフォーマンスを発揮できるような環境を整える働き方改革が進められているが，依然として半数以上の労働者は職業生活に強いストレスを感じているのが現状である．現在，労働者の精神衛生を保つ目的では，労働安全衛生調査が義務付けられているが，これはアンケートにより１年に１回程度実施するもので，日常的なモニタリングは難しい．そこで我々は，マルチモーダル情報を用いて，労働者の仕事に対する心的状態（ワーク・エンゲイジメントやリカバリ経験等を含めた「Work Attitude」）の推定を持続的に行うことができる「Work Attitude PLR (Personal Life Record) 収集基盤」を提案し，構築している．本発表では，大規模な実証実験に先立つ予備実験で収集したデータでの推定性能を示し，本手法の実現可能性が肯定的に示唆されたことを述べる．

1. はじめに

1970 年以降の先進７カ国 (G7) の中で，日本の労働生産性は一貫して最下位である．これを改善するため，労働者が健康にパフォーマンスを発揮できるような環境を整える働き方改革が進められているが，依然として半数以上の労働者は職業生活に強いストレスを感じているのが現状である [1].

現在，労働者の精神衛生を保つ目的では，労働安全衛生調査 [2] が義務付けられているが，これはアンケートにより１年に１回程度実施するもので，日常的なモニタリングはできていない．スマートフォンの普及に伴い，モバイルデバイスを使用してユーザの精神状態を監視し，セルフケアを促進することへの関心が高まっており [3]，スマートデバイスを用いた心理状態の推定手法も提案されている．しかし，多くの既存手法は，スマートウォッチなどの接触デバイスの情報を用いており，ユーザの負荷を高めている．

そこで我々は，多くのユーザが日常的に所持しているスマートフォンのみから得られるマルチモーダル情報を用いて，労働者の仕事に対する心的状態（ワーク・エンゲイジメントやリカバリ経験等を含めた「Work Attitude」）の推定を持続的に行うことができる「Work Attitude PLR (Personal Life Record) 収集基盤」を提案し，構築している．

本稿では，我々の提案手法について述べ，その実現可能性を調査するために行った評価実験とその結果を示す．なお，今回の評価実験は，大規模な実証実験に先立つ予備的なものである．

これにより，充分なサンプル数を得られた推定においては，チャンス・レベルを上回る性能を得られた．逆に，サンプル数が不足した推定においては，有効な結果を示すことができなかった．また，特徴量選択の結果から，推定対象により，使用される特徴量に異なる傾向が見られることがわかった．以上の結果から，本手法に対する大規模な実証実験の必要性と実現可能性が肯定的に示唆された．

2. 関連研究

本章では，本研究の目的である Work Attitude の持続的推定の関連研究として，アンケートやスマートデバイスを用いた心理状態のモニタリング手法について述べる．

2.1 アンケートを用いた心理状態のモニタリング研究

心理状態を定量的に把握することは，うつ病や不安障害などといった精神疾患を予防し，人々の健康を守る上で重要である．目には見えない心理状態を定量的に扱うために，アンケートベースでの測定がさかんに行われている．例えば，ALS（Affective Lability Scale）[4] は，心理状態を表す文章を読み，それがどれだけ自分に当てはまるかを答えさ

1　奈良先端科学技術大学院大学
Nara Instituite of Science and Technology
〒 630-0192, 奈良県生駒市高山町 8916-5
2　大阪大学
Osaka University
〒 565-0871, 大阪府吹田市山田丘 1-1
a)　suwabe.kai.sz1@is.naist.jp

© 2023 Information Processing Society of Japan

せ気分を定量的にセンシングするものである．これは一般的に使用されている臨床調査手段の１つであるが，過去の経験の偏りによる想起バイアスや個人の記憶能力の問題から，即時性に欠けるという問題がある．EMA（Ecological Momentary Assessment ）[5] は，この問題を解消する手法として注目されている．EMA の質問項目は，人の経験に関するリアルタイムな情報を取得するように設計された短い質問であり，既存のアンケートベースでの測定では困難であった場所や時間に囚われない測定を可能としている．しかしながら，回答者の負担がかかるという課題は解決できておらず，アンケートベースでの測定はデータ収集の粒度と参加者の負担とのトレードオフを考慮しなければならない．

こうした問題の解決法として，スマートデバイスにより収集したデータを使用し，回答者の負担を減らしながら測定の質を担保しようとする研究が行われている．

2.2 スマートデバイスを用いた心理状態のモニタリング研究

スマートデバイスを用いて心理尺度を推定しようとする研究は，最近注目されている．雨森らは，スマートデバイスを用いた学生の QOL 簡易評価に関する研究 [6] を行っている．生体データや活動データを特徴量とし、QOL の推定モデルを構築している．Natasha らは，スマートデバイスを用いた学生の幸福さなどを予測する研究 [7] を行っている．生体データや活動データを特徴量とし，健康・エネルギー・ストレスなどの幸福さに関わるような５つの尺度を組み合わせた複合的な心理尺度を推定するモデルを構築している．Sano らは，ウェアラブルセンサとスマートフォンによるデータ収集とアンケートによるストレス，精神的健康状態に関する調査を実施し，客観的データを使用してアカデミックパフォーマンス（GPA）やピッツバーグ睡眠指標（PSQI）の値を 67〜92%の分類精度の範囲で推定している [8]．Boukhechba らは，スマートフォンによる自動客観センサデータの収集と，アンケートによる社会的不安およびうつ病の病状に関する調査を実施し，データと心理的指標との間の多くの重要な相関関係を示している [9]．Fukazawa らは，スマートフォンのセンサーログとアプリケーション履歴データと不安を表す STAI スコアから，日常生活における無意識の不安の変化を予測しようと試み，74.2%の F 値を示す性能の推定モデルを構築している [10]．

これらの先行研究では，主観的評価である心理指標を，ウェアラブルデバイスやスマートフォンによって収集される客観的なデータから一定の水準で推定することに成功している．しかしながら，スマートウォッチや特殊なウェアラブルデバイスなど，ユーザが特定のデバイスを所持している必要があるため，ユーザの参加コストが高い．また，殆どの研究がストレスやうつ病などのネガティブな指標を対象としており，労働者の好調さをあらわすポジティブな指標は対象とされていない．ストレスがあったとしても，仕事に対する意欲やリカバリーが高い場合は問題ないことが指摘されており，「Work Attitude」を計測するためには，ポジティブ・ネガティブ両面の指標を対象とする必要がある．

我々は，日常的な精神衛生モニタリングを可能にするために，労働者の所持率が高いスマートフォンのみを用いてポジティブ・ネガティブ両面の心理尺度を含んだ「Work Attitude」を日常的に推定することを目的とする．

3. 提案手法

本章では，労働者の日常的な精神衛生モニタリングを可能にするために我々が提案している [11]，Work Attitude PLR 収集基盤について述べる．まず，心的状態の評価尺度として用いる Work Attitude について説明し，その後に，Work Attitude 推定に用いる，スマートフォンでの収集データについて述べる．最後に，先に述べた評価尺度やデータを用いた Work Attitude PLR 収集基盤の構成を示す．

3.1 Work Attitude

日常的な精神衛生モニタリングのためには，すでにいくつかの心理尺度が開発されている．それぞれの尺度は，１年に１回程度の調査を想定している職業性ストレス指標と異なり，少ない項目数で心的状態を評価できるため，負担が少ない．しかし，労働者の精神衛生モニタリングを行うには，先の章で述べたとおり，ポジティヴ・ネガティヴの両面を評価する必要があるため，単一の心理尺度だけでは不充分である．

そこで，本研究では，既存尺度であるワーク・エンゲイジメント，リカバリ経験，リカバリ状態の３つをパイロットスケールとし，複合して Work Attitude 評価尺度とする．以下に，それぞれのパイロットスケールについて述べる．

ワーク・エンゲイジメントについては，活力，熱意，没頭という３つの下位尺度を持つユトレヒト・ ワーク・エンゲイジメント (UWES) [12] [13] [14] [15] [11] について，３項目で評価可能な UWES-3 [16] の日本語版を採用する．本手法では，この３項目を WE1, WE2, WE3 と呼ぶ．この尺度は，労働者の労働生産性にも関わる重要な尺度である．

リカバリ経験 (Recovery Experience Measures) [17] [18] は，ストレス負荷により高まったストレスレベルを回復するための活動が余暇に行えたかを評価する尺度である．この尺度は，Psychological detachment (仕事からの離脱性)，Relaxation (リラクゼーション)，Mastery (仕事以外での挑戦)，Control (余暇の制御) という４つの下位尺度を評価する尺度である．各下位尺度について４項目の計 16 項目に回答する必要があるが，本手法では，各尺度に対し，因子

負荷量が最も高い項目を１項目ずつ抽出し，下位尺度として採用する．本手法では，この４項目を RE1, RE2, RE3, RE4 と呼ぶ．高ストレスレベルの状態が慢性的に続くことで，精神障害を発症するリスクが高まるため，リカバリ経験の測定は重要である．また，開発者の Sonnentag らの実験 [17] [18] においても，労働生産性への影響が確認されていることから，リカバリ経験は重要な尺度と考えられる．

リカバリ状態 [19] は，仕事前における回復状態を評価する尺度である．前回の仕事後から次回の仕事前の間，つまり，余暇における回復の結果を示す．仕事前に充分に回復できている状態は，身体的・心理的リソースを仕事に多く利用可能であることを表している．逆に，回復が不充分な状態は，リソースが不足していることを表している．リカバリ状態は，全体的な回復，身体的な回復，精神的な回復，活力について，４項目で評価する尺度である．本研究では，全体的な回復およびワーク・エンゲイジメントと重複する活力の２項目を省き，個別の回復状態を評価可能である，身体的な回復・精神的な回復の２項目を採用する．また，リカバリ状態は，睡眠後の回復状態を評価するものであり，これには，睡眠の質も関係している [20]．そのため，ピッツバーグ睡眠質問票 [21] のうち，睡眠の質に関する質問を，回復状態の１項目として取り入れる．本手法では，以上の３項目を RS1, RS2, RS3 と呼ぶ．

以上の尺度を用いて労働者の精神衛生をモニタリングする．ワーク・エンゲイジメント (WE1〜3) は，仕事が規定要因となると考えられるため，仕事後の余暇（夜）に測定を行う．リカバリ経験 (RE1〜4)，リカバリ状態 (RS1〜3) は余暇の過ごし方と睡眠が規定要因と考えられるため，仕事前（朝）に測定を行う．

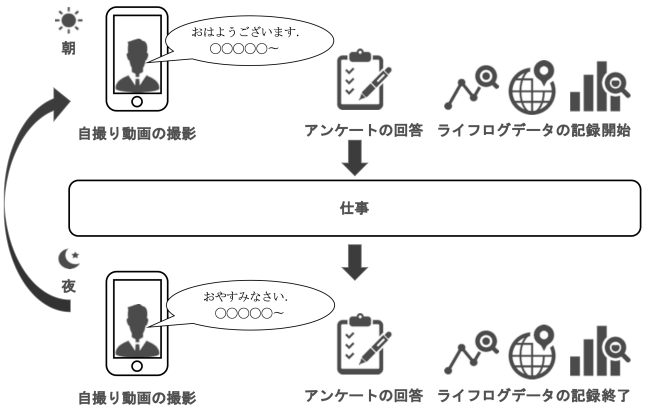

図 1　データ収集の流れ [11]

3.2　スマートフォンによる収集データ

労働者の Work Attitude 推定を行うために，本手法では，スマートフォンによるデータ収集を行う．収集するデータは，活動量データと，労働者による 10 秒程度の自撮り動画データ，そして，アンケートデータである．

活動量データは，活動範囲と歩数からなる．これらは，ストレスや抑鬱状態について調査する多くの研究 [22] において，心的状態と関係があるとして，広く採用されてきたものである．活動量データは，スマートフォンのアプリケーションによって自動的に収集されたものを用いる．

自撮り動画データは，労働者がその時点での気持ちを一言のメッセージとして発言したものである．したがって，顔画像，音声，テキスト（発話内容）からなる．顔画像データは，表情を認識することで感情推定が可能であるとされ [23]，音声データは，鬱病に関する心理尺度を推定する研究 [24] が存在している．これらに加え，本研究では新たに発話内容も含めてマルチモーダルな感情推定を行うことで，既存研究では明らかにならなかった心理尺度との関係が明らかになる可能性がある．

アンケートデータは，自撮り動画撮影後に収集し，労働者の睡眠効率を算出する．なお，このアンケートは評価実験における Ground Truth 取得のためにも用いたので，詳細は４章で述べる．

図 1 にデータ収集の流れを示す．労働者は，朝に起きてから仕事へ向かうまでに自撮り動画の撮影を行い，ライフログデータの記録を開始する．仕事から帰宅した後，夜の就寝前に自撮り動画の撮影を行い，ライフログデータの記録を終了する．各撮影での第一声は，「おはようございます」（朝）と「おやすみなさい」（夜）とし，ベースタームとする．これにより，同じタームに対する声質の違いを比較できる．そして，ベースタームの後に，一言のメッセージ（たとえば，「今日はよく寝た．元気いっぱいだ」，「まだ眠い．会社へ行きたくない」，「今日はお客さんに褒められた．我ながらよく頑張った」，「課長に怒られた．寝て忘れよう」など）を発言する．

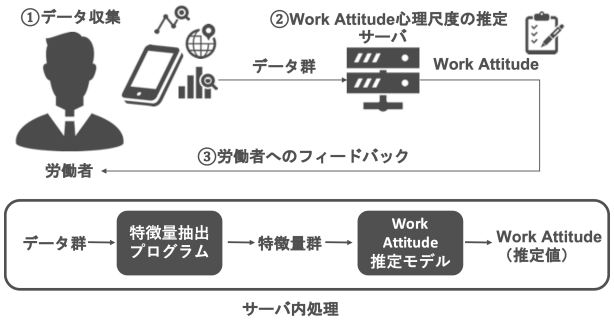

図 2　Work Attitude PLR 収集基盤の構成 [11]

3.3　Work Attitude PLR 収集基盤

図 2 に Work Attitude PLR 収集基盤の構成を示す．システムは，労働者のデータを収集するスマートフォンとそれを処理・蓄積するサーバで構成される．

© 2023 Information Processing Society of Japan

スマートフォンは，労働者の活動量データと自撮り動画データを収集し，サーバに送信する．サーバは，特徴量抽出プログラムにより，受信したデータ群から特徴量群を抽出し，Work Attitude 推定モデルに入力することで，Work Attitude の推定値を出力する．推定モデルが推定した Work Attitude は，PLR (Personal Life Record) としてデータベースに記録される．また，この Work Attitude は労働者にフィードバックし，これにより，労働者は働き方を客観的に振り返ることができる．

特徴量抽出プログラムは，活動量データから，１日の歩数と総移動距離を抽出し，特徴量とする．また，自撮り動画データにおいては，顔画像，音声，テキスト（発話内容）の３つのチャネルからそれぞれ得られる感情値を特徴量として抽出する．感情値の取得には，それぞれ Face API (Microsoft 社)，Empath API (Empath 社)，Cotoha API (NTT コミュニケーションズ社) を使用する．なお，顔画像チャネルの特徴量に関しては，動画から５枚の画像を抽出して，それぞれの感情値の平均値，中央値，分散，最大値を特徴量とした．このようにして抽出した特徴量は，**表 1** のとおりである．

表 1 特徴量

データ種	特徴量	備考
顔画像	smile anger contempt disgust fear happiness neutral sadness surprise	０から１までの数値． １に近いほど，その感情に当てはまる．
音声	calm anger joy sorrow energy	０から 50 までの数値． 50 に近いほど，その感情に当てはまる．
テキスト	positive/neutral/negative P N PN 悲しい 不安 好ましい 興奮 喜ぶ 安心 嫌	positive/neutral/negative は，テキスト全体の気分を判定する． その他の項目は，テキスト内に該当する感情の言葉がいくつ含まれているかをカウントする．
活動量	Distance Step	総移動距離 [m] 歩数 [歩]
アンケート	Sleep Efficiency	睡眠効率 [%]

Work Attitude 推定モデルは，事前にデータを収集し，機械学習を用いて学習することで構築する．推定モデルには，多数の説明変数を使用する学習に適する Random Forest を用いる．

4. 評価実験

本章では，Work Attitude PLR 収集基盤の実現可能性を実証するため，スマートフォンにより収集したデータを用いて Work Attitude 推定モデルを構築し，この性能を評価した．本発表では，大規模な実証実験に先立つ予備実験で収集したデータでの推定性能を示す．

4.1 収集データ

本実験には，９人の男性被験者が参加した．期間は２週間で，データは，そのうち休日を除く 10 日分を収集した．被験者は，朝夜の動画撮影のみを課され，それ以外は普段どおりに行動した．その結果，朝のデータを 78 サンプル，夜のデータを 77 サンプル，合計 155 サンプルを得ることができた．

なお，この実験では，Work Attitude の推定性能を評価するため，Ground Truth として用いる Work Attitude の値を，動画撮影後のアンケートで調査した．

4.2 Work Attitude 推定モデル

Work Attitude の推定は，すべての特徴量を用いた場合と，特徴量選択を行った場合の２通りでモデルを構築した．評価はどちらも，一つのサンプルをテストデータとし，それを除く，すべての被験者のサンプルをトレーニングデータとして推定を行う Leave-One-Out の交差検証で行った．

表 1 で示す 57 個の特徴量のうち，Sleep Efficiency（睡眠効率）は，その後の仕事中のワーク・エンゲイジメントにのみ作用すると考えられるので，WE1〜4 の推定ではこれを除く 56 個の特徴量を，RE1〜4 および RS1〜3 の推定では 57 個すべての特徴量を用いて推定を行った．

特徴量選択は，ジニ重要度を基準に行った．重要度の高い特徴量から順に一つずつ特徴量を追加して逐次的に推定を行い，最も性能が高くなったときの特徴量のサブセットを採用した．

なお，推定モデルの構築は，MATLAB で行った．

5. 結果

本章では，４章で述べた評価実験の結果について述べる．まず，被験者がアンケートによって答えた Work Attitude の分布について述べ，その後に Work Attitude 推定モデルの性能評価の結果を示す．性能評価は，正答率 (Accuracy) と F1 スコアで行った．

© 2023 Information Processing Society of Japan

5.1 アンケート結果

被験者がアンケートによって答えた Work Attitude の Ground Truth は，0と答えたサンプル数が非常に少ないなど，評価検証が難しいものだった．そこで，各クラスのサンプル数を確保するため，5，7段階の評価項目（RS3以外のすべての項目）は3段階評価に，4段階の評価項目（RS3）は2段階評価にカテゴリを統合している．その結果を**図 3** に示す．

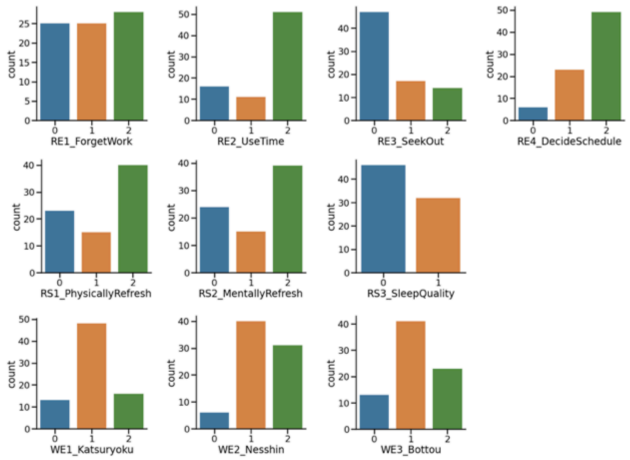

図 3 加工後のアンケート回答結果

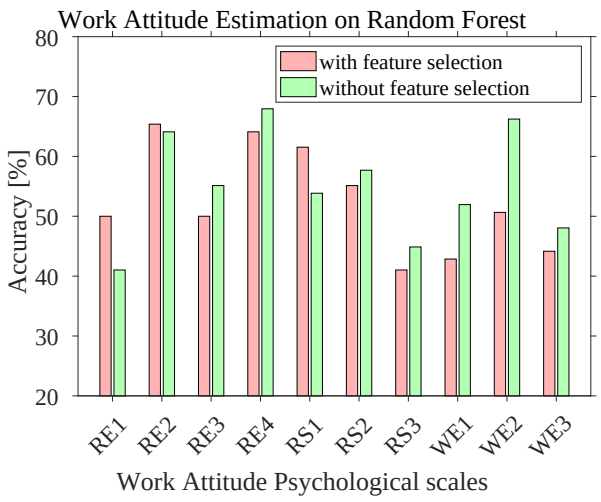

図 4 Work Attitude 推定の正答率

5.2 推定結果

Work Attitude 推定モデルの各尺度における正答率 (accuracy) を**図 4** に示す．リカバリ経験 (RE)，リカバリ状態 (RS)，ワーク・エンゲイジメント (WE) の平均値はそれぞれ，特徴量選択を行ったとき，57.37 (± 8.53)%, 52.96 (± 10.49)%, 45.89 (± 4.17)%，特徴量選択を行わなかったとき，57.05 (± 11.96)%, 52.14 (± 6.58)%, 55.41 (± 9.57)% であった．また，すべての項目についての平均値は，特徴量選択を行ったときが 52.48 (± 8.82)%，特徴量選択を行わなかったときが 55.08 (± 9.07)%であった．

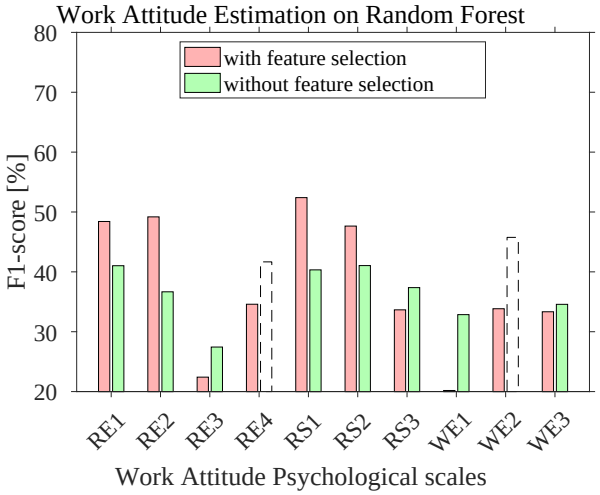

図 5 Work Attitude 推定の F1 スコア

また，Attitude 推定モデルの各尺度における F1 スコアを**図 5** に示す．ここで，点線 RS, RS, WE の平均値はそれぞれ，特徴量選択を行ったとき，38.65 (± 12.73)%, 44.57 (± 9.74)%, 29.12 (± 7.74)%，特徴量選択を行わなかったとき，36.70 (± 6.56)%, 39.58 (± 1.95)%, 37.73 (± 7.01)% であった．また，すべての項目についての平均値は，特徴量選択を行ったときが 37.57 (± 11.36)%，特徴量選択を行わなかったときが 37.87 (± 5.26)%であった．なお，特徴量選択を行わなかったとき，RE4 と WE2 の推定において，特定のクラスに分類されたケースがなかったため，適合率 (Precision) を用いて計算する F1 スコアが計算できなかった．破線のバーが示すのは，該当する適合率を0として補間し，F1 スコアを算出した値である．

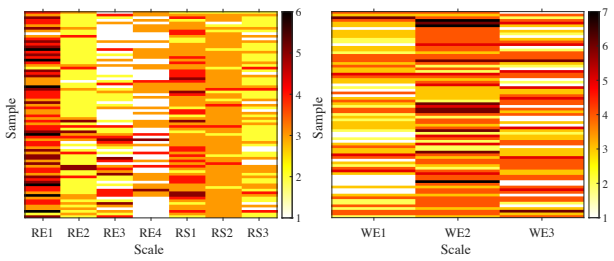

図 6 各項目の推定で選択された特徴量数

次に，特徴量選択の結果を示す．

各項目の推定で選択された特徴量数は**図 6**，選択された特徴量の度数分布は**図 7** のとおりであった．度数分布は，上段が RE1〜4，中段が RS1〜3，下段が WE1〜3 の推定におけるものであり，各横軸は表 1 で示した特徴量，各縦軸は交差検証を行ったときに選択された度数である．

© 2023 Information Processing Society of Japan

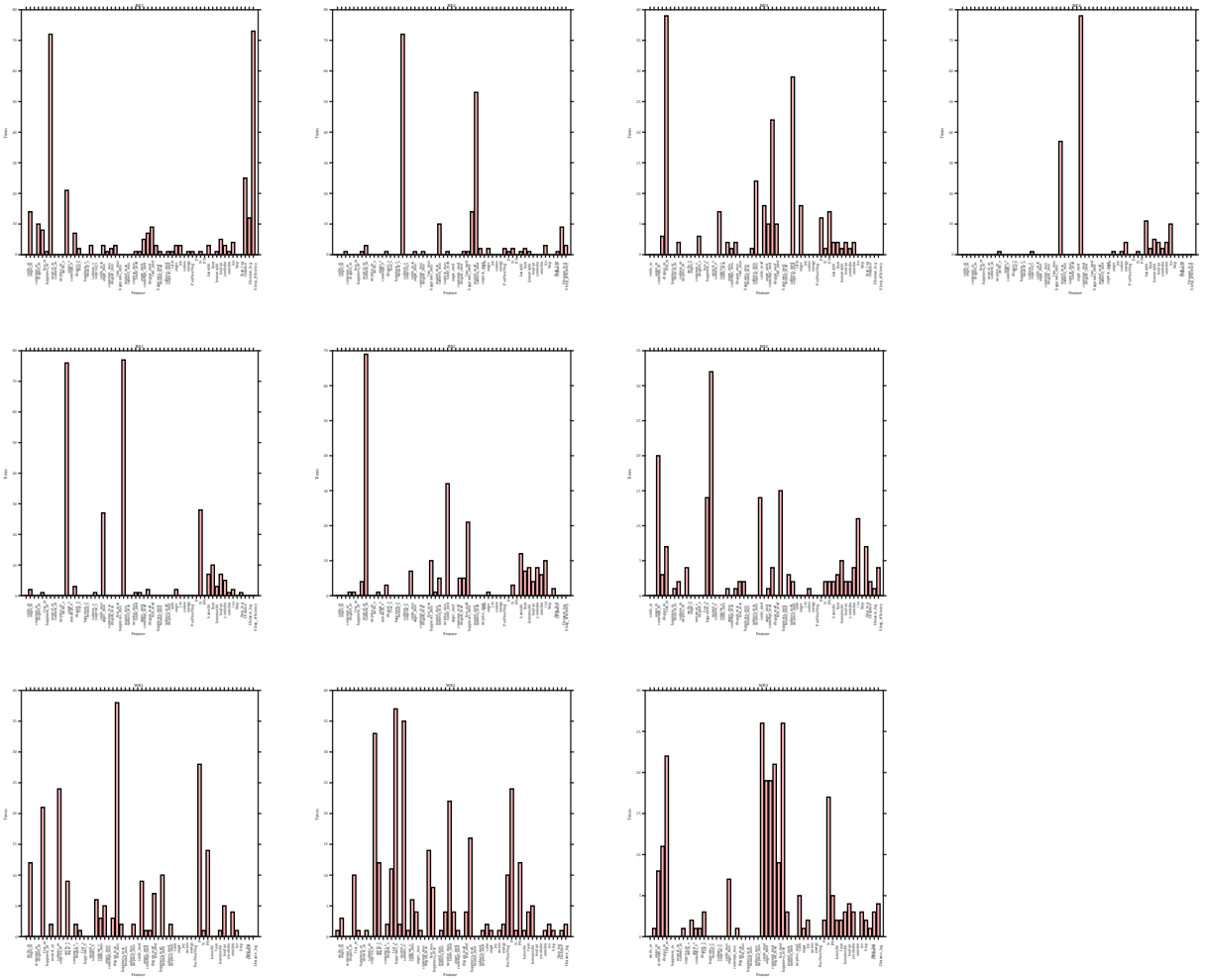

図 7　各項目の推定で選択された特徴量の度数分布

6.　考察

Work Attitude の推定においては，３値分類（RS3 は２値問題）で特徴量選択を行った場合も，行わなかった場合も，全体的に 50%程度の正答率を得られた．しかし，F1 スコアはいずれの場合も 37%程度であり，各クラスのサンプル数の偏りが推定結果も偏らせていると考えられる．このことは，特徴量選択なしの RE4 と WE2 の推定において，特定のクラスが全く選ばれなかったことにも表れている．

また，特徴量選択を行った場合と行わなかった場合を比較すると，全体では正答率と F1 スコアに大きな変化は見られなかった．

ただし，各クラスのサンプル数が最も均等である RE1 の推定においては，特徴量選択を行った場合に推定性能の向上が見られ，チャンス・レベルが 33%程度であるのに対し，正答率も F1 スコアも 50%程度であった．

以上から，他の項目についても，さらに大規模なデータ収集を行い，各クラスのサンプル数を充実させることで，有効な推定できるようになる可能性がある．

また，特徴量選択の結果を見ると，選択された特徴量数，特徴量が選択された度数分布は，項目ごとに異なる傾向が見られた．このことから，Work Attitude 推定には，マルチモーダル情報の特徴量を用いる我々の提案手法が適していることが示唆される．充分なサンプル数を確保すれば，各項目と特徴量の関係を明らかにし，推定に必要なデータを削減できる可能性がある．

なお，最も充実したサンプルを得た RE1 においては，毎回の推定で選択された特徴量が比較的多かった．また，選択された特徴量の度数分布においても，特定の特徴量（特に，Sleep_efficiency, neutral_m）が特に多かった．ただし，度数分布についての傾向は，他の項目にも同様に見られた．

7.　おわりに

本稿では，労働者の日常的な精神衛生をモニタリングを可能にするため，既存の複数の心理尺度を複合した Work Attitude と，スマートフォンによって取得できるデータを用いて Work Attitude を推定・蓄積・フィードバックする「Work Attitude PLR 収集基盤」について述べた．また，これを実現するために，機械学習を用いた推定モデルが必要であるため，その推定性能の評価や，推定に用いる特徴

© 2023 Information Processing Society of Japan

量の調査のための実験を行い，その結果を示した．本稿で述べた実験は，大規模な実証実験に先立つ予備的なものである．

実験の結果，充分にサンプル数を確保できた項目においては，特徴量選択が推定性能を向上させることを示し，3値分類で50%程度の推定性能を得ることができた．ただし，サンプルに偏りの大きいデータセットに対しては，有効な推定を行うことができなかった．また，特徴量選択を行った推定において，使用された特徴量数やその度数分布には推定対象の項目ごとに異なった傾向が見られることを示した．

今後の展望としては，有効な推定結果を得るためにはサンプル数の偏りを減らす必要があると考えられるので，大規模なデータ収集実験を行い，充分なサンプル数を確保した上で，ダウン・サンプリングを行うなど処理を施したデータセットを用いた実証実験を実施するつもりである．

謝辞　本研究活動は，文部科学省による Society 5.0 実現化研究拠点支援事業 (グラント番号: JPMXP0518071489) によって行われたものである．

参考文献

[1] 厚生労働省. 平成 30 年「労働安全衛生調査 (実態調査)」の概況. 2018.

[2] 厚生労働省. 労働安全衛生法に基づくストレスチェック制度実施マニュアル. 2019.

[3] V.P. Cornet and R.J Holden. Systematic review of smartphone- based passive sensing for health and well-being. *Journal of Biomedical Informatics*, Vol. 77, pp. 120–132, 2018.

[4] Mark R Serper Philip D Harvey, Barbara R Greenberg. The affective lability scales: development, reliability, and validity. *Journal of clinical psychology*, Vol. 45, No. 5.

[5] RICHARD LI SIDNEY K. D' MELLO MUNMUN DE CHOUDHURY GREGORY D. ABOWD THOMAS PLÖTZMEHRAB BIN MORSHED. Prediction of mood instability with passive sensing. *Interact. Mob. Wearable Ubiquitous Technol*, Vol. 3, No. 3.

[6] C. Amenomori, T. Mizumoto, and et al. A method for simplified hrqol measurement by smart devices. *Wireless Mobile Communication and Healthcare*, Vol. 247, pp. 91–98, 2018.

[7] Jaques Natasha and et al. Predicting students' happiness from physiology, phone, mobility, and behavioral data. *Int Conf Affect Comput Intell Interact Workshops*, 2015.

[8] A.J. Sano, A.Z. Phillips, and et al. Recognizing academic per- formance,sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. *Proc. IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 1–6, 2015.

[9] M. Boukhechba, A. Daros, and et al. Demonicsalmon: Monitoring mental health and social interactions of college students using smartphones. *Smart Health, CHASE 2018 Special Issue*, Vol. 9-10, pp. 192–203, 2018.

[10] Y. Fukazawa, T. Itoh, and et al. Predicting anxiety state using smartphone-based passive sensing. *Journal of Biomedical Informatics*, Vol. 93, pp. 103–151, 2019.

[11] 前川哲志, 水本旭洋, 諏訪博彦, 島津明人, 安本慶一. 労働者の精神衛生モニタリングに向けたデータ収集と分析. 行動変容と社会システム vol.07, Vol. 2021, , 2021.

[12] 島津明人. ワーク・ライフ・バランスと メンタルヘルス-共働き夫婦に焦点を当てて. 日本労働研究誌, No. 653, pp. 75–84, 2014.

[13] W.B. Schaufeli, M. Salanova, V Gonzalez-Romá, and et al. The measurement of engagement and burnout: A two sample confirmative analytic approach. *Journal of Happiness Studies*, Vol. 3, pp. 71–92, 2002.

[14] W. B. Schaufeli and A.B. Bakker. Job demands job resources and their relationship with burnout and engagement :a multi-sample study. *Journal of Organizational Behavior*, Vol. 25, pp. 293–315, 2004.

[15] W. B. Schaufeli and A.B. Bakker. The conceptualization and measurement of work engagement:a review. *Work Engagement: A Handbook of Essential Theory and Research*, pp. 10–24, 2010.

[16] A. Shimazu, W.B. Schaufeli, S. Kosugi, and et al. Work engagement in japan:validation of the japanese version of utrecht work engagement scale. *Applied Psychology: An International Review*, Vol. 57, pp. 510–523, 2008.

[17] S. Sonnentag and C. Fritz. The recovery experience questionnaire: Development and validation of a measure for assessing recuperation and unwinding from work. *Journal of Occupational Health Psychology*, Vol. 12, No. 3, pp. 204–221, 2007.

[18] A. Shimazu, S. Sonnentag, K. Kubota, and N. Kawakami. Validation of the japanese version of the recovery experience questionnaire. *Journal of Occupational Health*, Vol. 54, pp. 196–205, 2012.

[19] C. Binnewies, S. Sonnentag, and E. J. Mojza. Daily performance at work: feeling recovered in the morning as a predictor of day-level job performance. *Journal of Organizational Behavior,J. Organiz. Behav*, Vol. 30, pp. 67–93, 2009.

[20] S. Sonnentag, C. Binnewies, and E. J. Mojza. "did you have a nice evening?" a day-level study on recovery experiences, sleep, and affect. *Journal of Applied Psychology*, Vol. 93, No. 3, pp. 674–684, 2008.

[21] D. J. Buysse, C. F. Reynolds, T. H. Monk, S. R. Berman, and D. J. Kupfer. Pittsburgh sleep quality index (psqi). *Psychiatry Research*, Vol. 28, pp. 193–213, 1989.

[22] Y. Fukazawa, N. Yamamoto, T. Hamatani, and et al. Smartphone-based mental state estimation: A survey from a machine learning perspective. *Journal of Information Processing*, Vol. 28, pp. 16–30, 2020.

[23] F. Noroozi, M. Marjanovic, A Njegus, and et al. Audio-visual emotion recognition in video clips. *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING*, Vol. 10, No. 1.

[24] A.A. Farhan, J. Lu, J. Bi, A. Russell, B. Wang, and A. Bamis. Multi-view bi-clustering to identify smartphone sensing features indicative of depression. *IEEE 1st International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pp. 264–273, 2016.

© 2023 Information Processing Society of Japan