

生体情報を教師データとした スマートフォンログを用いたストレス状態推定

山本直樹^{†1} 落合桂一^{†1} 稲垣章弥^{†1} 深澤佑介^{†1} 木本勝敏^{†1}
霧生和樹^{†2} 上西康平^{†2} 太田順^{†2} 沖村宰^{†3} 寺澤悠理^{†3} 前田貴記^{†3}

概要：現在、精神疾患の患者数は年々増加傾向にあり、メンタルヘルスの不調を未然に防ぐことが社会的な課題となっている。そこで本研究では、スマートフォンで取得できるセンサや端末ログ（以下、スマートフォンログ）からストレス状態を推定する手法を提案する。スマートフォンログからストレスと関連する行動の特徴量を生成し、心拍間隔から計算した LFHF を教師データとして予測モデルを学習する。提案手法の有効性を 39 名のデータで評価した。

キーワード：ストレス、メンタルヘルスクア、スマートフォンログ

1. はじめに

厚生労働省の調査[1]によると、精神疾患の患者数は年々増加傾向にあり、メンタルヘルスクアに対する社会的関心が高まっている。しかしながら、メンタルヘルスの不調は目視で確認できるものではないために、個々が自身のストレス状態を正確に把握することは非常に困難である。

自身のストレス状態を知る手法として、アンケートや生体情報を用いて、直接的にストレスを計測する方法がある。しかしながら、アンケート形式のストレス計測[2]では、連続的なストレス状態の評価ができず、また評価がユーザの主観に委ねられるために、ユーザ自身が気づいていないストレス状態を評価できないという問題がある。さらに、アンケートへの回答という行為がユーザの負担となる可能性がある。一方で、生体情報を用いる手法[7]では、血液や唾液、心拍データを用いることで客観的に自身のストレス状態を知ることができるが、生体情報を計測するための特殊な装置が必要になる。生体情報の計測や、計測装置を身体に装着することが、ユーザに苦痛や不快感を与え、ストレスの原因となる可能性がある。

そこで、ユーザに負担を与えることなくストレス状態を評価する手法として、スマートフォンを用いたストレス推定技術が存在する。スマートフォンは、個々が所持しているものであり、かつ頻りに持ち歩き、使用するものであることから、スマートフォンから取得できる各種センサや利用履歴のログは、ユーザの精神状態を推定するために有用であると考えられる。このようなスマートフォンの特性から、WHO はメンタルヘルスクアにおいて、携帯電話を用いたセルフマネジメントを促している[3]。スマートフォンから収集したログを基に、ユーザの精神状態を推定する研究として、例えば、大学生を対象として、スマートフォンに搭載されたセンサや GPS のログと、PHQ-9 スコアの関係

性を検証している研究[5]や、同データセットを用いて、GPS データを基にうつ病の重症度を特定する研究[6]が報告されている。この様に、スマートフォンログを基にストレス状態を推定することで、連続的なストレス状態の変化を評価することが可能となる。

しかしながら、これらの研究はアンケートの回答結果を教師データとしてストレス状態の推定を実施しているために、ユーザ自身が気づいていないストレスは推定することは不可能である。また、正解データがユーザの主観的なものであるために、ストレス状態を正しく捉えることができていない可能性があるという問題がある。

以上より、本研究では心拍データを基にしたストレス状態を教師データとして、スマートフォンログよりストレス状態を推定するモデルを構築することにより、特殊な装置を用いることなく、客観的かつ連続的なストレス評価が可能な手法を提案する。

以降、2 章では関連研究を述べ、3 章では提案モデルに関して説明する。4 章では対象となるデータに関して説明し、5 章で推定精度の評価方法について説明する。6 章で本論文をまとめる。

2. 関連研究

本章では、本研究に関連する既存研究について述べる。ユーザのストレス状態を評価するための技術として、主に 1)アンケートを用いる手法、2)生体情報を用いる手法の 2 通りが知られている。1)のアンケートを用いる手法は、アンケートを用いて直接的にストレスを計測する手法と、アンケートの回答とスマートフォンやセンサのログを用いて、ストレス状態を機械学習により推定する手法が報告されている。2)生体情報を用いる手法は、専用の器具を用いてユーザの血液や唾液、心拍を直接的に計測することで、ストレス状態を客観的に評価する手法である。

†1 株式会社 NTT ドコモ
NTT DOCOMO, INC.

†2 東京大学
The University of Tokyo

†3 慶應義塾大学
Keio University

2.1 アンケートを用いたメンタルヘルスの評価

メンタルヘルスの評価に用いられるアンケートとして PHQ (Patient Health Questionnaire)[8]がある。この内、大うつ病性障害に関する9つの質問を抽出したものが PHQ-9 である。PHQ-9 では、9つの各設問に対して、4択の選択肢から自分の心身の状態に該当する回答を選択する。各選択肢にはスコアが割り振られており、全設問に回答後にスコアを集計することで、その時のユーザのうつ病性障害の症状レベルを評価することが可能である。

しかしながら、アンケートを用いた手法ではストレス状態を連続的に評価することが困難であり、またユーザ自身が気づけていないストレスを評価することは不可能である。また、回答するという動作がユーザの負担となる可能性がある。

2.2 生体情報を用いたメンタルヘルスの評価

生体情報として血液や唾液、心拍を用いたストレス状態の評価手法が報告されている。生体情報を用いることで、客観的なストレス状態の評価が可能となる。Ogino らは、血液中に含まれるアルギナーゼ I と呼ばれるたんぱく質が、ストレス指標として有用であることを示している[9]。しかしながら、採血という行動がユーザに対して不安やストレス、苦痛を与える可能性が考えられる。そこで、ユーザに対して肉体的に苦痛を与えることなく、客観的にストレス状態を評価する手法として、唾液によるストレス状態の評価がある[10]。唾液腺における α -アミラーゼは、ストレスを感じている状態では通常時に比べて多く分泌されることが知られており、この特性を利用してストレス状態を評価することが可能だと言われている。しかしながら、唾液の採取を行うためには特殊な装置が必要となるために、日常的にストレス状態の評価を行うのは難しい。

精神的にストレスを感じると自律神経系にその影響が現れるため、自律神経系の変調を計測することによってストレスを解析する研究が報告されている[7]。自律神経機能の解析において多く利用されている生理指標に、心拍変動解析がある。心拍は常に一定の間隔で鼓動を打つのではなく、じっと安静にしている最中であっても周期的にゆらいでいることが知られており、この周期的な心拍のゆらぎを心拍変動と呼ぶ。心拍変動解析とは、この心拍周期の変動を解析する手法であり、解析指標としてよく知られているものに LF/HF (Low Frequency / High Frequency)がある。LF/HF とは、自律神経機能の活性度合を表す指標である。心拍データをスペクトル分析すると、心拍変動には低周波成分(LF)と高周波成分(HF)と呼ばれる2つの主要な成分を観察することが出来る。LFとHFは、それぞれ交感神経と副交感神経の活性度を表しており、交感神経が活性化している状態、すなわちストレス状態にある場合には、HF成分に対してLF成分が大きくなるため、LF/HFの値が大きくな

る。一方で、副交感神経が活性化している場合、すなわちリラックス状態である場合には、LF成分に対してHF成分が大きくなるため、LF/HFの値は小さくなる。したがって、心拍データよりLF/HFを求めることにより、客観的にストレス状態を評価することが可能である。心拍センサにより心拍データを収集することで、客観的に連続的なストレス状態を評価することが可能となるが、特殊な装置を身体に装着し続けることが、ユーザのストレスとなり得るという問題がある。

2.3 アンケートとスマートフォンログを用いたメンタルヘルスの評価

スマートフォンを用いたストレス推定に関する研究は多数報告されている[5][6][11][12][13][14][15][16]。

例えば Wang らは、アンケートを基にしたうつ病の重症度とスマートフォンを基に取得したユーザの行動傾向、大学の学業成績の3項間における関係性を報告している[5]。これらの3項間にはいくつか重要な関係性が見つかり、例えば会話の頻度と成績や、睡眠時間とうつ病の重症度合の間に関係性が発見されている。

また、Wang らは退院後の統合失調症患者21名を対象に、スマートフォンログとEMAスコアの相関関係を報告している。また、スマートフォンログを基に、ユーザの行動や周辺環境に関する特徴量を作成し、EMAを平均誤差7.6%で予測することができたと報告している[13]。

Saeb らは、学生を対象とした Wang らのデータセットを用いて、スマートフォンのGPSデータからユーザのうつ病の重症度を特定する研究を報告している[6]。Saeb らは、本研究がスマートフォンセンサを用いたうつ病の重症度を推定する可能性を示唆しているとしながらも、ユーザのストレス状態を受動的に、かつ精度よく推定する技術開発が今後の課題であるとしている。

Huang らは、18名の大学生を対象として、スマートフォンログを基にユーザの移動軌跡を解析し、滞留点を Home, Food&Leisure といった5種類のカテゴリに割り当て、各カテゴリの総滞留時間やカテゴリ間の遷移と、アンケートの相関関係を分析している[14]。

スマートフォンによりストレス状態を推定することで、特殊な装置を用いて直接的に生体情報を計測する必要がなくなり、かつスマートフォンは毎日持ち歩き、使用するものであるから、連続的なストレス状態を評価することが可能となる。しかしながら、既存のストレス推定研究では、推定モデル作成時の正解データにアンケートを用いているため、ユーザ自身が気付いていないストレスは推定することは不可能であり、またユーザの主観的なストレスしか推定できないという課題がある。

3. 提案モデル

本研究では、スマートフォンのログをより作成した行動特徴量を基に、ユーザの LF/HF を推定する手法を提案する。

本章では、まず本研究の問題設定を述べる。次にスマートフォンログを基に行動特徴量を生成する過程から、ストレス推定モデルの作成までを順に説明する。本研究におけるストレス推定の概要を図 1 に示す。提案手法では、スマートフォンログから特徴量を生成後、ユーザ毎に各特徴量の標準化を行い、データを整形する。次に、次元削減手法である t-SNE を用いて特徴量を 2 次元で表現し、2 次元の特徴量を基にクラスタリングを行うことで、行動傾向が似ているユーザ群毎に推定モデルを作成する。推定モデルには全特徴量を入力するのではなく、推定に有効な特徴量のみを選択し入力する。推定モデルは、3 種類の学習器を組み合わせたアンサンブル学習を適用しており、3 種類の推定結果を基に多数決を行った結果を最終的な推定結果として出力する。ストレス状態の指標は、生理指標である心拍変動を用いることで、客観的なストレス状態を推定することが可能となり、かつスマートフォンを用いて推定することにより、ユーザのストレス状態を非侵襲的に取得することが可能となる。また、日毎のストレスを推定することにより、連続的なストレス状態の変化を把握することが可能な技術を提案する。

3.1 問題設定

本報告では、各ユーザの LF/HF の平均値を基準として、日毎の LF/HF が基準値以上か、もしくは基準値未満かを推定する 2 値予測問題を問題として設定する。

3.2 スマートフォンログからの特徴量生成

本節では、スマートフォンログを基に生成した行動特徴量に関して述べる

3.2.1 位置情報に関する特徴量

GPS より収集した位置情報を基に、日毎の総移動距離や最大移動距離、移動距離の標準偏差、回転断面半径、自宅からの最大移動距離、訪問施設数、重要訪問施設数、新しく訪れた場所の数、位置情報の繰り返し度を求め、12 次元特徴量とした[17]。位置情報の繰り返し度は、Jaccard 係数と Dice 係数を基に、前日や前週の位置情報と、当日の位置情報がどの程度似ているかを定義し、どの程度同じ行動を繰り返しているのかを表す指標として特徴量とした。

3.2.2 センサ情報に関する特徴量

加速度、照度、気圧センサより収集したスマートフォンログは、1 日を 4 分割し、各時間帯における平均、分散と、1 日の最大値、最小値、最大値と最小値の差分を特徴量と

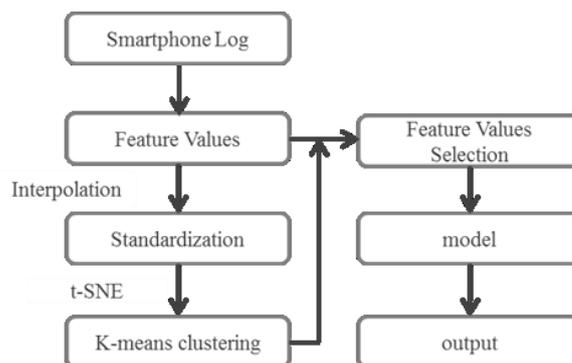


図 1 ストレス推定モデル

した。4 分割した時間帯は、1) 0 時から 6 時、2) 6 時から 12 時、3) 12 時から 18 時、4) 18 時から 24 時、である。(以下、順に 1Q, 2Q, 3Q, 4Q と表記する。) 以上より、各センサを基に 35 次元の特徴量を作成した。

3.2.3 行動特徴量

ユーザの身体的な活用の量がストレスに影響を与えるという想定で行動特徴量を利用する。具体的には歩行、走行、停止、乗り物利用の時間を利用する。装具には Wang らの研究[13]と同様に、Google が提供する Activity Recognition API^{*1}を利用した。ユーザの各行動を、センサ情報に関する特徴量と同様に 1 日を 4 分割し、各時間帯における集計値として特徴量化した。また、1 日全体での集計値を特徴量とした。以上より、40 次元の行動特徴量を作成した。

3.2.4 スマートフォン利用履歴

全てのアプリ利用履歴を取得し、センサ情報に関する特徴量と同様に 1 日を 4 分割し、各時間帯におけるアプリの利用回数を集計した。特に、業務用アプリケーションや他者とのコミュニケーションに用いる SMS、電話の起動履歴に関しては、それぞれ単独で起動回数をカウントし、特徴量として追加した。また、充電器とイヤホンの脱着回数、バッテリーとストレージの容量を取得し、特徴量に追加した。スマートフォンの利用履歴に関して、33 次元の特徴量を作成した。

3.2.5 曜日

7 次元の One-hot ベクトルを用いて曜日の情報を特徴量化した。また、祝日とそれ以外の日を 2 値で表現し、曜日に関して計 8 次元の特徴量を作成した。

*1 Activity Recognition API Google Developers
<https://developers.google.com/location-context/activity-recognition/>

3.3 ユーザのクラスタリング

スマートフォンの使用方法には個人差があると考えられる。いくつか例を挙げると、頻繁に画面の ON/OFF を繰り返す人、常時身に着け持ち歩くがほとんど使用することが無い人、夜間の使用が多い人等がある。ゆえに、スマートフォンログを基に、各ユーザのスマートフォンの使用傾向を把握する必要がある。また、行動傾向が似たユーザ毎に推定モデルを作成することで、多様な行動パターンを反映した推定モデルが構築できると考える。

そこで本研究では、まず非線形次元削減手法である t-SNE[18]によって、128 次元の毎日の特徴量を 2 次元に次元削減し、プロットすることで可視化する。そして、2 次元に次元削減した特徴量を基に、K-means 法[19]を用いてクラスタリングを行う。この時に、同一ユーザの特徴量であっても日により異なるクラスタに分類される可能性がある。この場合、各ユーザの特徴量が最も多く分類されたクラスタを、そのユーザが属するクラスタとして設定するものとした。

3.4 特徴量選択

機械学習において、予測に有効な特徴量のみを選択し、選択した特徴量のみを予測モデルに入力することで、予測モデルの汎化性能が向上すると言われている[20]。そこで本研究では、各特徴量と目的変数である LF/HF の単変量解析を行い、推定に有効であると判断された特徴量上位 k 個を用いて推定モデルを構築し、ストレス状態を推定する[21]。

3.5 アンサンブル学習

機械学習における汎化性能向上を目的としたアプローチとして、同一タスクに対して複数の学習器を個別に学習し、それぞれの出力結果を組み合わせ、最終的な出力とするアンサンブル学習と呼ばれる手法が存在する[24]。本研究では、機械学習手法である K-Nearest-Neighbor (KNN)[22], Support Vector Machine (SVM)[23], Random Forest[24]を用いて、ユーザのストレス状態を推定し、各学習器が推定したストレス状態の出力結果を組み合わせ、多数決により各日の最終的な推定結果を出力する。

4. 評価方法

本章では、スマートフォンログ及び心拍データ収集の概要と、収集したデータセットに関して述べる。

4.1 研究協力者

本研究への協力者は、株式会社 NTT ドコモの R&D 部門の社員を対象として募集した。研究への協力者は 20 代から 50 代の社員 39 名であり、男性 34 名、女性 5 名から構成さ

表 1 収集するスマートフォンログ一覧

収集したスマートフォンログ	収集間隔
3 軸方向加速度	1 [sec]
照度	5 [min]
気圧	5 [min]
位置情報	5 [min]
充電器の脱着	アクション発生時
イヤホンの脱着	アクション発生時
画面の ON/OFF 履歴	アクション発生時
アプリ起動履歴	10 [min]
バッテリー残量	5 [min]
Activity Recognition API	5 [min]
ストレージ容量	1 [回/日]

れる。研究期間は 2017 年 11 月から 2018 年 1 月の 3 か月間であり、研究参加者はこの期間内で最大 42 日間、本研究へ参加した。本研究への参加は任意であり、研究内容に同意頂けた社員にのみ協力頂いた。また、研究期間中は協力者の意思で、いずれのタイミングでも研究を中断できるものとした。研究で収集した各種データは、プライバシーの観点から最大限の配慮をして取り扱った。

4.2 研究協力の内容

研究期間中は、研究協力者の業務用スマートフォンにログ収集を目的としたアプリケーションをインストールし、各種スマートフォンログを常時収集した。収集したスマートフォンログの一覧を表 1 に示す。研究期間中、協力者は通常通りスマートフォンを利用し、特別な操作は必要としない。なお、本研究では Android5 系、6 系、7 系を搭載したスマートフォンをログ収集の対象として研究を実施した。

スマートフォンログの収集と並行し、研究協力者は胸部に心拍センサを装着し、心拍データを収集した。心拍センサの装着時間は午前 9 時から午後 8 時までの 11 時間を設定した。この時間内は、協力者は心拍センサを着用した状態で日常生活を送り、激しい運動を行うといった非日常的な動作を行う時には、一時的に心拍センサを外すものとした。本研究で用いた心拍センサはユニオンツール株式会社製の MyBeat 心拍センサ WHS-1 である。

4.3 データセット

39 名の研究参加者より、1349 日分のスマートフォンログと 724 日分の心拍データを収集した。心拍データから LF/HF を算出後、ユーザ毎に研究期間中の LF/HF の平均値を基準値として、毎日の LF/HF が基準値以上であれば 1、基準値未満であれば 0 を正解データとして設定した。

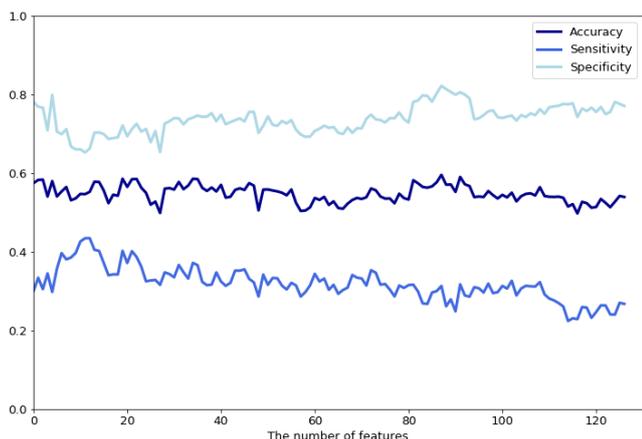


図 2 特徴量選択による精度

表 2 各 Model における推定精度

	Accuracy	Sensitivity	Specificity
Baseline Model	0.540	0.269	0.771
Feature Select Model	0.596	0.313	0.822
Clustering Model	0.607	0.292	0.831
Proposed Model	0.711	0.606	0.773

行動特徴量と正解データが紐付いたデータは 638 日分であり、35 名のユーザから構成される。正解データ中に含まれる 1 と 0 の日数は、それぞれ 285 日、353 日である。

4.4 評価指標

本研究では、スマートフォンログを基に、各ユーザの LF/HF の平均値を基準値として、日毎の LF/HF が基準値以上か基準値未満かを推定する 2 値予測を問題として設定する。

ストレス推定精度の評価は、Accuracy(正確度), Sensitivity(感度), Specificity (特異度)の 3 指標により行う。Accuracy は、ある検査について「全サンプルに対する陽性、陰性を正しく分類した割合」を表している。Sensitivity, Specificity は、臨床検査において検査の有用性を評価するために用いられている指標である。Sensitivity は、ある検査について「陽性と診断されるべきものを正しく陽性と診断した割合」を表しており、Specificity は「陰性と判断されるべきものを正しく陰性と診断した割合」を表している。

推定精度の評価は、Leave-one-user-out-cross validation を用いる。クラスタ毎に、クラスタ内全ユーザに対して、1 人分をテストデータ、他のデータを学習データとしてモデルの構築と推定を行い、推定精度を算出する。ユーザ毎に Accuracy, Sensitivity, Specificity の 3 指標を算出し、クラスタ毎にそれぞれの平均値を算出する。最後に各クラスタの 3 指標の平均値を算出し、全体における推定精度として評価する。

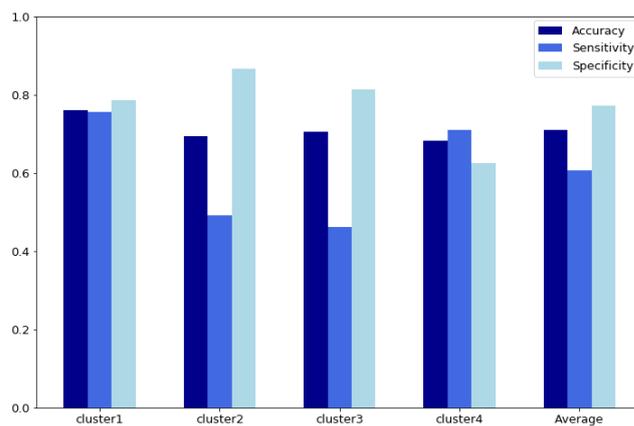


図 3 Proposed Model による各クラスタの推定精度

4.5 比較手法

本研究では、単純な機械学習モデルである KNN を用いて、全ユーザのデータに対してストレス推定を行った結果を基準とする。また、この時のモデルを Baseline Model と呼ぶこととする。次に、全ユーザのデータに対して特徴量選択を行った結果を Feature Select Model として示す。

また、Baseline Model に対して、ユーザをクラスタリングする処理を加えたモデルを Clustering Model と呼ぶこととする。最後にユーザのクラスタリング、特徴量の選択、KNN と SVM と Random Forest によるアンサンブル学習を適用した提案モデルを Proposed Model と呼ぶこととする。

5. 結果

本章では、各ストレス推定モデルにおける推定精度を評価する。特にクラスタリングと特徴量選択が推定精度に与える影響について述べ、最後に現在のストレス推定精度に言及する。

5.1 定量評価

図 2 は Baseline Model を用いてストレス推定を行った結果である。横軸は推定に用いた特徴量数を表わしており、縦軸は特徴量数を 1 次元から全特徴量数である 128 次元まで変化させた時の Accuracy, Sensitivity, Specificity の値の変化を表している。この結果より、いずれの評価指標においても特徴量選択によって約 5%から 10%程度の値の変化が確認できる。これは 128 次元の全行動特徴が各ユーザのストレス状態を反映しているのではなく、一部の行動特徴量にストレス状態が現れているためだと考えられ、推定に用いる特徴量を適切に選択することで、ストレス推定精度の向上に効果的であると考えている。

表 2 に Baseline Model, Feature Select Model, Clustering Model, Proposed Model の各モデルにおける推定精度を示す。まず、Baseline Model と Feature Select Model の推定精度を

比較すると、Accuracy, Sensitivity, Specificity において約5%の精度向上が確認できることから、特徴量選択により推定精度が向上することが示された。続いて、Baseline Model と Clustering Model の推定精度を比較すると、ユーザのクラスタリングにより Accuracy は約6%, Sensitivity は約2%, Specificity は約6%精度が向上した。したがって、クラスタリングにより、行動傾向が似たユーザ群毎に推定を行うことが重要だと示された。最後に Proposed Model の推定精度を見ると、Baseline Model と比較して Accuracy で約17%, Sensitivity で約33%の精度向上が見られ、Specificity はほぼ同程度の値となっている。Feature Select Model と Clustering Model の推定結果の様に、特徴量選択やクラスタリングをそれぞれ行うことで推定精度の向上が見られるが、これらを組み合わせることでより精度改善が図られたと考えられる。図3は Proposed Model における各クラスタの Accuracy, Sensitivity, Specificity を表している。クラスタによって、Sensitivity の値に大きな違いがあり、Cluster2, 3 では約50%程度の値となっていることから、ストレス状態をより精度よく推定するためには、これらのクラスタにおいてストレス状態を正確に推定できるよう改善が必要である。

6. おわりに

本研究では、スマートフォンログを用いてユーザのストレス状態を推定することを目指した。ユーザのストレス状態を生体情報である心拍変動を基にした LF/HF で定義し、ユーザの LF/HF が平均以上か、平均未満かを推定することを問題として設定した。推定にはスマートフォンログを用いており、スマートフォンにはユーザの行動傾向や使用方法が反映されているために、スマートフォンログを基にした行動特徴量により、ユーザ毎のストレス状態を非侵襲的、かつ連続的に評価することが可能である。

提案手法では、行動特徴量を基にユーザをクラスタリングし、クラスタ毎にストレス状態を反映した行動特徴量を選択し、アンサンブル学習を適用することで、平均して約70%の Accuracy でストレス推定が可能であることが示された。

今後の課題として、推定精度の改善と将来的なストレス状態の推定が挙げられる。より精度よくユーザのストレス状態を推定することにより、ユーザのストレス状態を正しく把握し、未然にメンタルヘルスケアの不調を防ぐことが可能になる。推定精度の向上のためには、行動特徴量の見直しが必要であると考えている。現状の行動特徴量は、各時間帯におけるスマートフォンログの平均値や集計値を基に作成したものを含んでいるが、これらの特徴量ではユーザの行動特徴を十分に反映できていない可能性がある。ユーザの日々の行動傾向の違いを明瞭に表現できる行動特徴量を作成することで、高ストレス時と低ストレス時における行動の違いを捉え、より高精度で推定を行うことが可能

だと考えている。

また、本提案手法では日毎のストレス状態を2値で定義したが、半日毎や数時間毎といったより短い時間間隔でのストレス推定や、ストレス状態を5段階で推定する、LF/HF の数値を推定するなど、推定をより細かい粒度で行うことで、ストレス状態の細かい変化を捉えることが可能になると考えている。さらに、本提案手法では当日のスマートフォンログを基に当日のストレス状態を推定しているが、今日までのスマートフォンログを基に、翌日以降のストレス状態を推定することで、メンタルヘルスケアの悪化を未然に防ぐことができる技術となると考えている。

参考文献

- [1] “厚生労働省 疾患調査”。
<http://www.mhlw.go.jp/toukei/saikin/hw/kanja/14/dl/toukei.pdf>
- [2] KROENKE, Kurt; SPITZER, Robert L.; WILLIAMS, Janet BW. The phq - 9. *Journal of general internal medicine*, 2001, 16:9: 606-613.
- [3] ANTHES, Emily. Pocket psychiatry: mobile mental-health apps have exploded onto the market, but few have been thoroughly tested. *Nature*, 2016, 532.7597: 20-24.
- [4] WANG, Rui, et al. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2014. p. 3-14.
- [5] SAEB, Sohrab, et al. The relationship between mobile phone location sensor data and depressive symptom severity. *PeerJ*, 2016, 4: e2537.
- [6] POMERANZ, Bruce, et al. Assessment of autonomic function in humans by heart rate spectral analysis. *American Journal of Physiology-Heart and Circulatory Physiology*, 1985, 248.1: H151-H153.
- [7] SPITZER, Robert L., et al. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Jama*, 1999, 282.18: 1737-1744.
- [8] OGINO, Keiki, et al. Association of arginase I or nitric oxide-related factors with job strain in healthy workers. *PLoS one*, 2017, 12.4: e0175696.
- [9] TAKAI, Noriyasu, et al. Effect of psychological stress on the salivary cortisol and amylase levels in healthy young adults. *Archives of oral biology*, 2004, 49.12: 963-968.
- [10] SANO, Akane; PICARD, Rosalind W. Stress recognition using wearable sensors and mobile phones. In: *Affective Computing and Intelligent Interaction (ACII)*, 2013 Humaine Association Conference on. IEEE, 2013. p. 671-676.
- [11] SUHARA, Yoshihiko; XU, Yinzhao; PENTLAND, Alex'Sandy'. Deepmood: Forecasting depressed mood based on self-reported histories via recurrent neural networks. In: *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2017. p. 715-724.
- [12] WANG, Rui, et al. CrossCheck: toward passive sensing and detection of mental health changes in people with schizophrenia. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016. p. 886-897.
- [13] HUANG, Yu, et al. Assessing social anxiety using GPS trajectories and point-of-interest data. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016. p. 898-903.

- [14] CANZIAN, Luca; MUSOLESI, Mirco. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing. ACM, 2015. p. 1293-1304.
- [15] BOGOMOLOV, Andrey, et al. Daily stress recognition from mobile phone data, weather conditions and individual traits. In: Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014. p. 477-486.
- [16] CANZIAN, Luca; MUSOLESI, Mirco. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing. ACM, 2015. p. 1293-1304.
- [17] MAATEN, Laurens van der; HINTON, Geoffrey. Visualizing data using t-SNE. Journal of machine learning research, 2008, 9.Nov: 2579-2605.
- [18] MACQUEEN, James, et al. Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967. p. 281-297.
- [19] CHANDRASHEKAR, Girish; SAHIN, Ferat. A survey on feature selection methods. Computers & Electrical Engineering, 2014, 40.1: 16-28.
- [20] VAPNIK, Vladimir. Statistical learning theory. 1998. Wiley, New York, 1998.
- [21] HEIMAN, Gary W. Research methods in psychology. Houghton Mifflin College Division, 1995.
- [22] SUYKENS, Johan AK; VANDEWALLE, Joos. Least squares support vector machine classifiers. Neural processing letters, 1999, 9.3: 293-300.
- [23] LIAW, Andy, et al. Classification and regression by randomForest. R news, 2002, 2.3: 18-22.
- [24] ROKACH, Lior. Ensemble-based classifiers. Artificial Intelligence Review, 2010, 33.1-2: 1-39.