

NAIST-IS-MT1651067

修士論文

生産性の低下抑止のための ウェアラブル機器による体調の推定と予測

高橋 雄太

2018 年 3 月 15 日

奈良先端科学技術大学院大学
情報科学研究科 情報科学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学) 授与の要件として提出した修士論文である。

高橋 雄太

審査委員：

安本 慶一 教授	(主指導教員)
池田 和司 教授	(副指導教員)
荒川 豊 准教授	(副指導教員)
水本 旭洋 特任助教	(副指導教員)

生産性の低下抑止のための ウェアラブル機器による体調の推定と予測*

高橋 雄太

内容梗概

体調の不具合は誰しもが経験することがあり，生活習慣，寝不足，食生活やストレス等様々な要因で生じる．疾患に比べれば体調は軽視されることが多いが，日々の生活での体調の不具合は活力の低下や集中力を妨げる要因となり，生産性の低下を引き起こす．1日の体調の推定や次の日の体調の予測ができれば，体調のモニタリングや体調不良の抑止などに利用できる．しかし，体調のような内的な状態は，動作などの外的な状態とは異なりセンサでの推定が難しいとされている．体調が良い状態，悪い状態共に生活習慣が影響していることから，体調の推定や予測には生活周期を取れるようなセンサが必要となる．人の生活を常時センシングできるデバイスとしてウェアラブル機器が利用できる．歩行，心拍，睡眠のような日々の活動であるライフログをウェアラブル機器から収集し，解析することで生活周期を抽出できると考える．ウェアラブル機器から得られるライフログから体調の推定と予測が行えるのかを検証・評価することが本研究の目的である．本研究では，まず労働生産性に影響の大きい体調を過去の研究から調査・選定し，選定した体調が既存の技術で測定・推定可能か調査した．調査結果から，体調の推定には心拍センサが有用であることを確認し，実験では心拍センサを搭載したウェアラブル機器を利用した．体調の推定・予測手法として，ライフログから時間，生活周期に関する特徴量を自己相関関数などによって抽出する．自己相関関数は同じ波形を畳み込みさせながら相関を算出する関数で，波形のおお

*奈良先端科学技術大学院大学 情報科学研究科 情報科学専攻 修士論文, NAIST-IS-MT1651067, 2018年3月15日.

よその周期を求めることができる。特徴量を説明変数，主観的なメンタルヘルス（3クラス：悪い，普通，良い），フィジカルヘルス（3クラス：悪い，普通，良い），体調不良の度合いを目的変数とした推定モデルと次の日の体調の予測モデルを Support Vector Machine を含む3種類の機械学習アルゴリズムで構築する。提案手法の有用性を評価するため，10名の実験協力者を募り，ライフログと主観的な体調に関するアンケートの回答を1ヶ月半に亘り収集した。実験のデータを用いてモデルを構築し推定・予測能力の評価を行ったところ，フィジカルヘルスは，88.1%の精度で推定，84.0%の精度で予測でき，メンタルヘルスは，93.2%の精度で推定，94.1%の精度で予測できた。体調不良の度合いに関しては，データ量が十分でないなどの理由で，適切な推定・予測モデルの構築できなかった。これらの結果から，ウェアラブル機器のみでフィジカル・メンタルの調子を個人依存のモデルで高い精度で推定・予測できることを明らかにした。

キーワード

健康，内的状態，ウェアラブル機器，機械学習，自己相関関数

Estimation and prediction of physical and mental health using a wearable device for preventing loss productivity.*

Yuta Takahashi

Abstract

Everyone has an experience of becoming bad physical condition or bad mental condition which is caused by various factors such as lifecycle, lack of sleeping, eating habit, or stress. Such conditions are often neglected, but they can lead to the decreased activity and distraction, and decrease of the people's productivity. If there is a technology that everyone can estimate or predict his/her (physical/mental) conditions on the current day or the next day, he/she can cope with or prevent bad conditions. However, estimating inner states like physical condition or mental condition using sensors are much more difficult than estimating outer states like activities. Both good and bad conditions are considered affected by a life habit. Therefore, we need a sensor for recording a life habit for estimating and predicting conditions. We can use a wearable device as a device that can sense our life for a whole day. It is very probable to extract a life habit from life logs such as walking steps, heartbeats and sleeping time/duration obtained by analyzing data from a wearable device. Our object in this thesis is to investigate and evaluate whether it is possible or not to estimate and predict physical and mental conditions using life logs from the wearable device. For this purpose, First, we surveyed existing studies and selected a kind of condition which affects

*Master's Thesis, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT1651067, March 15, 2018.

work productivity. We also surveyed whether the conditions can be estimated or measured using current technologies. From a result of the survey, we confirmed that a heartbeat sensor is effective for estimating each condition. Then, we determined to employ a wearable device with a heartbeat sensor in our study. We propose a method for estimating and predicting physical and mental conditions consisting of two steps. The first step extracts features related to time and life habit by using autocorrelation function and so on. The autocorrelation function is the function which calculates the correlation while convolute the same waveform to determine the period of the waveform. Then, we build a model for estimating (for the current day) and predicting (for the next day) physical/mental conditions by using three machine learning algorithms including Support Vector Machine. In the model, features are explanatory variables, and three classes (bad, normal, good) of mental health or physical health condition are objective variables. To evaluate usefulness of our proposed method, we conducted an experiment where ten participants collected life logs and answered subjective physical and health conditions (ground truth) every day during one month and a half. We built a model using the collected data and evaluated the accuracy of estimation and prediction. As a result, we confirmed physical health could be estimated and predicted with accuracy of 88.1% and 84.0%, respectively. Moreover, mental health could be estimated and predicted with accuracy of 93.2% and 94.1%, respectively. We could not build a proper model for estimating and predicting physical and health conditions in finer level. However, our results revealed that physical and mental conditions can be estimated and predicted with reasonable accuracy by using only a wearable device.

Keywords:

Health, Inner State, Wearable Device, Machine Learning, Autocorrelation Function

目 次

1. 序論	1
2. 関連研究	4
2.1 内的状態の推定	4
2.2 本研究の立ち位置	7
3. 労働生産性に影響のある体調とその計測・推定技術の調査	8
3.1 体調	8
3.2 プレゼンティーズム	8
3.3 労働生産性に影響のある体調不良の選定	10
3.4 体調の計測・推定技術	11
3.4.1 不安・抑うつ	12
3.4.2 肩こり・腰痛	12
3.4.3 鼻炎	13
3.4.4 疲労	13
3.4.5 睡眠障害	14
3.4.6 腹痛	14
3.5 体調の計測・推定技術のまとめ	14
4. ウェアラブル機器による体調の推定と予測	17
4.1 ウェアラブル機器によるライフログ収集	17
4.2 主観的な体調に関するアンケート	17
4.3 アンケート収集システム	20
4.3.1 アンケート回答アプリ「Anketo!」	22
4.3.2 ライフログの収集	22
4.4 体調の推定・予測手法	23
4.4.1 ライフログからの特徴量抽出	25
5. ライフログの収集実験と推定・予測評価	27

5.1	ライフログと体調の収集実験	27
5.2	フィジカルヘルスの推定・予測結果	27
5.3	メンタルヘルスの推定・予測結果	28
5.4	体調不良の度合いの推定と予測評価結果	30
6.	結論	31
	謝辞	33
	参考文献	34

図 目 次

1	SmartBand2	18
2	アンケート収集システムの概要	20
3	アンケート回答アプリ「Anketo!」	21
4	アンケートの質問項目の種類	22
5	体調の推定・予測に利用するデータの違い	24

表 目 次

1	プレゼンティーズムとワーク・エンゲージメントの代表的な指標	10
2	労働生産性に影響のある体調不良	11
3	体調の測定方法の調査結果	15
4	体調に関する質問項目	19
5	フィジカルヘルスの推定結果	28
6	フィジカルヘルスの予測結果	28
7	メンタルヘルスの推定結果	29
8	メンタルヘルスの予測結果	29

1. 序論

近年の AI (artificial intelligence) 技術の発達により、今まで人手で行われてきた仕事が計算機やロボットに置き換わることが期待されている。画像認識技術の向上による画像の精査や人や動植物のモニタリング、自動運転技術、産業用ロボットの高性能化、音声認識や対話技術の向上によるテレオペレータや店の受付のエージェント化など、実用ベースで研究開発が行われている。このような技術は、かつて手洗いによる洗濯が洗濯機による自動化に変わったように、時間がかかる作業を機械が代わりに行うことで、人の労働力を確保し、よりよい社会へ効率よく発展することへと繋がる。

AI によって、労働力を確保すると同時に、人間自身の労働生産性を向上する技術も期待される。ヒューマン・インタフェースの分野では、これまで人の知力や知覚能力、身体能力を増強させる研究が行われてきた。人とコンピュータの距離がより近くなることで、知覚的・身体的な作業の効率化が期待できる。例えば、エンジニアリングの労働生産性の尺度として、スピード、正確性、注意要求 [1] があげられる。このような尺度に作用する因子として、人の内外面の状態と環境が考えられる。ユビキタス・コンピューティングの分野では人の状態のことを内的コンテキスト (internal context) [2] という。内的コンテキストに関してこれまで様々な研究が行われており、行動のようなフィジカルな人間の状態の計測・推定は行えるようになってきた [3, 4, 5, 6]。そして、近年ではストレス [7, 8, 9, 10]、感情 [11, 11]、メンタルヘルス [7, 8, 10]、アカデミックパフォーマンス [7, 8]、幸福度 [7, 10, 12] などの人間の内面的な状態の計測・推定を目的とした研究が行われている。このような内面的な状態は直接計測できないため、行動のような外的な状態に比べて計測・推定が困難とされている。

労働生産性に直結する内面的な状態として、健康状態があげられる。疾患が原因の体調不良はもちろん、日々の生活習慣が起因する体調不良は、人の集中力を妨げ、労働生産性を低下させる。特に後者の体調不良においては、健康な身体であっても誰しもが患う可能性がある。このような体調不良を計測できれば、労働生産性の予測ができるとともに、不調を改善するようなアドバイスができ、労働生産性の低下を防ぐことができる可能性がある。さらに、次の日の体調を予測す

ることができれば、次の日の体調に合わせたスケジューリングや不調を防ぐような行動の提示も行える。しかしながら、体調は人間が内的に認知することができても、直接センサから測定・推定する技術は存在しない。

スマートフォンの普及、デバイス・センサの小型化・低電力化、無線通信の低電力化、バッテリーの小型化・大容量化によりリストバンド型ウェアラブル機器が製品化され普及しつつある。このようなウェアラブル機器は、加速度センサが内蔵されており、日々の行動をモニタリングし、歩いた歩数や距離、運動量を記録できる。そして、光学式の心拍センサが搭載されることで、人の睡眠状態を記録できるようになった。このような日々の活動のログをライフログと呼び、ウェアラブル機器を装着するだけで様々な種類のライフログを記録できる。ウェアラブル機器で記録できるライフログは人の生活を反映しており、生活習慣の情報を含んでいる。そのため、ウェアラブル機器で記録したライフログを用いることで人の体調を推定・予測できる可能性ある。

本研究では、既製品のウェアラブル機器から得られるライフログを用いて、人の体調の推定・予測を行う。まず、労働生産性を下げるような体調不良を既存研究から選定し、ライフログとアンケートを収集するプラットフォームを構築した。ウェアラブル機器から得られるライフログからフィジカルヘルス（3クラス：悪い、普通、良い）、メンタルヘルス（3クラス：悪い、普通、良い）、体調不良の度合いを、幾つかの機械学習アルゴリズムを使ってモデルを構築し、推定・予測する。予備実験を通して、3日分のデータから時間、相関や自己相関関数を用いて生活周期に関する特徴量を780種類抽出し、推定モデル、予測モデルの構築を行うこととした。自己相関関数は同じ波形を畳み込みさせながら相関を算出する関数で、波形のおおよその周期を求めることができる。

提案手法の有用性を評価するため、10名の実験協力者にウェアラブル機器を装着してもらい、ライフログを1か月半に渡り収集した。実験中は、体調に関するアンケートを毎日実施し、これを正解値とした。実験で得られたデータを元に(1)フィジカルヘルス、(2)メンタルヘルス、(3)体調不良の度合いを抽出する。(1)、(2)は良い/普通/悪いの3クラスの分類問題であり、(3)は労働生産性に影響が強い体調不良の度合のスコアであり実数値の回帰問題となっている。(1)、(2)に関

する推定モデル，予測モデルを構築し，Leave-One-Out cross-validation にて評価したところ，フィジカルヘルスは，最大 88.1%の精度で推定，最大 84.0%の精度で予測でき，メンタルヘルスは，最大 93.2%の精度で推定，最大 94.1%の精度で予測できた．推定モデル，予測モデルとして Random Forest[13], AdaBoost[14], Support Vector Machine[15] の 3 種類を評価したところ Support Vector Machine が最も精度の高いモデルであった．そして，フィジカルヘルスよりメンタルヘルスの方が 10%も推定，予測の精度が高いことも明らかとなった．(3) についての回帰モデルを構築し，Leave-One-Out 交差検証により，決定係数 R^2 と Explained Variance Score (EVS) で評価したところ，良い回帰モデルの構築は行えなかった．モデルの構築がうまくいかない原因として，データの数や評価値の偏り，複数の体調不良を 1 つのスコアで表現したことがあげられる．しかしながら，ウェアラブルのみでフィジカルヘルス，メンタルヘルスの良い状態・普通の状態・悪い状態を高い精度で推定，予測できることが明らかにした．

2 章では，関連する研究をあげ，本研究の位置づけについて明確化する．3 章では，体調と労働生産性に関する既存研究をもとに本研究で対象とする体調不良について選定し，これらの体調不良が既存の技術で推定・測定できるか調査しまとめた．4 章では，3 章で得られた知見を元に必要なウェアラブル機器を選定し，データの収集方法とデータの処理の手法，体調の推定，予測手法について述べる．5 章では，実験について述べ，体調を推定と予測した結果とその考察について述べる．6 章では，実験結果を踏まえ，提案手法の限界と今後の展開について述べ，まとめる．

2. 関連研究

2.1 内的状態の推定

これまで、スマートフォンやウェアラブル機器を用いて行動や睡眠などの認識が活発に行われてきている。機械学習技術の普及と共に十分な精度で認識できることがわかり、行動認識は、Google の Activity Recognition API¹や睡眠は Fitbit 社や Garmin 社のウェアラブルデバイスの睡眠認識検知機能のように社会実装されている。一方、アカデミックの分野では、より認識が難しいとされる内的状態を認識するような研究が増えつつある。このような研究では、ストレス [7, 8, 9, 10], 感情 [11, 11], メンタルヘルス [7, 8, 10], アカデミックパフォーマンス [7, 8], 幸福度 [7, 10, 12] などが対象とされ、人の内的状態とスマートフォンやウェアラブルセンサから得られるデータとの関連性が少しづつ明らかになっている。

Sano ら [7] は、スマートフォンとウェアラブルセンサを用いて、アカデミックパフォーマンスと睡眠の質とストレスレベルとメンタルヘルスの推定を行っている。スマートフォンからは SMS, Email, 電話の着信, スクリーンの状態など, ウェアラブルセンサは, Affectiva 社の Q-Sensor を用いており, 加速度, スキンコンダクタンス, 温度などのデータを収集している。そして, 朝と夜にアンケートを行い, 睡眠や幸福度など様々な項目の情報を収集し, これも推定に利用している。正解値となる指標として, アカデミックパフォーマンスには, 授業の成績の平均を表す Grade Point Average (GPA), 睡眠の質は Pittsburgh Sleep Quality Index (PSQI)[16], ストレスは, perceived stress scale (PSS)[17], SF-12 Physical and Mental Health Composite Scale (MCS)[18] を用いている。実験では, 66 人の実験協力者から 30 日分のデータをしている。4 つの指標の推定では, 上位 20% と下位 20% の二つのグループを作成し, これらの 2 値分類を行っている。全体の結果としては, 67-92% の精度が得られ, スマートフォンよりウェアラブルセンサの方が推定精度が高いことを報告している。

Wang ら [8] は, スマートフォンから得られるデータを用いてメンタルヘルス,

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

アカデミックパフォーマンス，行動の傾向の評価を行っている．スマートフォンの加速度，マイクロフォン，照度センサ，GPS からアクティビティ，会話，睡眠，位置を測定・推定し，収集している．そして，1日に数回，通知を行い，ストレスに関する質問，16枚の写真から写真を選択してもらい今の感情抽出する PAM[19]によりデータの収集を行っている．メンタルヘルスは鬱，活発度，ストレス，孤独の4つの軸で評価している．それぞれの評価方法として，鬱は patient health questionnaire (PHQ-9)[20]，活発度は，flourishing scale[21]，ストレスは PSS[17]，孤独は UCLA loneliness scale[22] を用いている．アカデミックパフォーマンスは GPA で評価している．実験は，スマートフォンのプログラミングの授業に参加する生徒 75 名に対して 10 週間行われた．実験のインセンティブとして，回答率の高い生徒をランダムに選びスマートフォンやウェアラブルデバイスを渡している．相関分析の結果，鬱は発話や睡眠，ストレスは発話，孤独は活動量，アカデミックパフォーマンスは室内の移動量や発話に関しての相関を報告している．活発度に関しては有意確率 $p < 0.05$ を満たすような相関が得られていない．なお，この研究で開発された StudentLife システムのデータセットは公開されている．

2017 年に公開された最近の研究で，Taylor[10] らは，ウェアラブルセンサ，スマートフォン，天気を用いて，次の日の感情，ストレス，健康の予測を行っている．ウェアラブルセンサは Affectiva 社の Q sensor を使用しており，スキンコンダクタンス，加速度，温度を 8Hz で収集し，スマートフォンからは電話，SMS，スクリーンの状態，位置情報を収集している．正解値として，感情 (sad/happy)，ストレス (stressed out/calm)，健康 (sick/healthy) を朝と夜に毎日収集している．実験協力者は，それぞれに対し主観的に 100 点満点でスコアをつける．実験は 104 名に対して 30 日間行われた．彼女らの手法の特徴として，マルチタスク学習を使用していることがあげられる．マルチタスク学習は転移学習の手法の一つであり，複数のタスクを一つのモデルに行わせることで，ニューラルネットの中間層のような共通因子を持たせ，精度を向上させる手法である．この研究では，感情，ストレス，健康の3つの予測タスクを行わせている．予測結果として，性別や Big Five trait taxonomy[23] を元に似通ったパーソナリティを持つ人をクラスタリングし，クラスターで分けて学習させた結果が最もよく，それぞれ 70-80%の精

度で予測することができている。予測モデルとしては、Multi-Task Learning したニューラルネット、Multi-Task Multi-Kernel Learning をしたニューラルネット、Hierarchical Bayesian Logistic Regression を使用している。

LiKamWa[24] らはスマートフォンを用いて一日の感情の推定に取り組んでいる。スマートフォンからはメール、SMS、電話、インターネットやアプリの使用状況を収集し、専用のアプリケーションを用いて一日の感情を記録する。感情のモデルは、Circumplex mood model[25] を用いており、喜びと活力の2軸に対して推定を行う。推定には、Multi-Linear Regression を使用し、Sequential Forward Selection[26] によって特徴量選択を行っている。実験は32名に対し、2か月間行われ、スマートフォンのデータは全てインターネットを通じて集約され、オンライン学習によって、毎日学習が行われる。結果として、個人依存の推定では、実験開始初期の精度が63%であったのにも関わらず、最終的には93%の精度で推定している。

Bogomolov ら [12] はスマートフォンから得られるデータを用いて幸福状態の認識を行っている。スマートフォンからは電話、SMS、Bluetooth の電波を収集し、7段階のスケールで主観的な幸福度を評価してもらっている。実験は、117名に対し行われ、1年半の期間行われた。幸福度は”not happy”, ”neutral”, ”happy” の三つのクラスに分割し、ランダムフォレストを用いて識別したところ、80%の精度が得られている。

Ciman ら [9] はスマートフォンを用いたストレスの評価を行っている。論文では、2つの評価方法を提案している。1つ目は、スマートフォンのアプリケーション用いた方法で、スマートフォン上で簡単なタスクをいくつか行うことでストレスの評価を行う。タスク中のスワイプやスクロールのようなタッチ操作の特徴量を抽出し、評価に使用している。13名に対して実験を行い、様々なモデルで5段階評価のストレスの推定をしたところ、スクロールの特徴量では、ユーザ依存で77-81%、ユーザ非依存で67-78%の精度が得られ、スワイプの特徴量では、ユーザ依存で79-87%、ユーザ非依存で75-92%の精度が得られた。識別モデルとしては、決定木、k近傍法、SVM、ニューラルネット、ベイジアンネットワークを使用、比較している。2つ目の方法は日々のスマートフォンの使用状況からストレ

スを評価する方法である。1つ目の方法のタッチ操作は専用のアプリケーションでないと詳細なデータの取得が難しく、AndroidではRoot化しなければ常時収集することができないため、実用性が低く、こちらの方法は実用性の高さを重視している。特徴量として、アプリケーションの使用履歴、画面のタッチ情報、ユーザの行動、スクリーンの状態を収集している。実験は、42名に対して4週間行われた。5段階のストレス値を3段階に落とし、識別モデルを構築したところ、ユーザ依存で86-88%、ユーザ非依存で63-88%の精度が得られている。情報ゲインから識別に有用な特徴量を求めると、使用したアプリの種類と頻度が有用となっている。

2.2 本研究の立ち位置

本研究も関連研究と同じく内的状態を推定する研究であるが、本研究は、コンシューマ向けのウェアラブル機器のみ用いている点、睡眠状態を収集できる点、推定・予測対象を体調としている点が異なっている。類似している研究として、Taylor[10]らの研究に近いが、それぞれ2値分類の問題である点やスマートフォンを用いている点などが異なる。また、メンタルヘルスの予測には取り組んでいない。本研究の新規性として、ウェアラブル機器のみでの体調の推定・予測、時間や周期を表現した特徴量を利用した内的状態の推定があげられる。

3. 労働生産性に影響のある体調とその計測・推定技術の調査

3.1 体調

体調は体の調子を意味しており，頭痛，腰痛，倦怠感などの自覚可能な不調が生じたら「体調を崩す」，「体調が優れない」と表現される．体調不良は人のやる気を低下し，集中力を妨げ，人の生産性を低下させる原因となる．そのため，体調を良い状態に保つことは，生産性の低下を防ぐことへとつながる．体調不良が生じる原因としては，睡眠，食生活のような生活習慣など様々な要因が考えられる．また，ストレスのようなメンタルの不調が原因となっている場合も多い．ウェアラブル機器を使用することで，睡眠や生活習慣，ストレスを収集することができ，人の体調を把握できる可能性がある．食生活を把握するために，食べたものを撮影し記録するような研究 [27, 28] やサービス [29] もあるが，量や精度がまだ十分でないため，本研究では取り扱わない．本研究で取り扱う体調として，元々持っていた疾患が原因となるものは，ウェアラブル機器によつての把握が難しいと考えられるため除く．そして，日々の生活習慣が原因となるような，入院や診察が必要なく，かつ，フィジカルヘルスやメンタルヘルスのような自覚可能な体調の状態や体調不良を対象とする．具体的な体調不調の種類に関しては，3.3 節にて過去の研究結果から選定を行う．

3.2 プレゼンティーイズム

体調と労働生産性に関してはこれまで産業分野において研究が進められてきた．たとえば，健康リスク要因（栄養失調，肥満，運動不足など）が多いほど，労働生産性の低下が大きいということが判明しており [30]，健康と労働生産性には密接な関係がある．米国の労働生産性と健康に関する調査では，体調が優れない状態での労働生産性の低下による損失は，従業員の医療費・薬剤費より大きいことが報じられている [31]．体調が優れない状態で労働することをプレゼンティーイズム (Presenteeism) といい，体調が優れなく欠勤してしまうことをアブセン

ティーイズム (Absenteeism) という。プレゼンティーイズムでの損失は、アブセンティーイズムと医療費の合計よりも多いと試算されており [32]，経営者の立場からすると，潜在的な会社の損失を減らすためには，従業員の健康状態を維持・増進する必要がある。このような考えに基づき，従業員の健康状態を把握・支援し，会社全体の生産性を上げる経営戦略を健康経営という。健康経営は経済産業省を中心に推進されており，近年，週休3日制や6時間労働を導入する企業が増えてきている。ヒューマンキャピタルテクノロジー社の Geppo[33] というサービスでは，従業員の満足度，人間関係，健康の情報を収集することで，従業員のコンディションを逐次把握でき，組織運営に役立てることができる。

従業員の立場からすると，体調が良い状態を維持することでより活発的に仕事に取り組みやすくなる。仕事に熱中し，熱意を持ち，仕事から活力を得ていきいきしている心理状態のことをワーク・エンゲージメントという。そして，仕事には熱中しているが，過度に脅迫的に仕事に取り組む状態をワーカホリックという。日本での調査において，ワーク・エンゲージメントは不健康に対し負の相関を示し，ワーカホリックは不健康に対し強い正の相関があることが報告されている [34]。また，生産性は，ワーク・エンゲージメントに正の相関があり，ワーカホリックに負の相関がある。つまり，従業員がいきいきとやりがいを持って生産的に仕事をするためには，健康は重要な要因の一つとなっている。

このように，労働生産性はプレゼンティーイズムやワーク・エンゲージメントを元に研究・調査が行われている。これらの実態を調査する手法として，質問紙調査票ベースの指標が使用されてきた。表 1 にプレゼンティーイズムとワーク・エンゲージメントの調査で使用される代表的な指標の一覧を示す。この中でも Stanford Presenteeism Scale (SPS)[35] と The Health and Work Performance Questionnaire (HPQ) [36]，health problem を自由に記述することができるため，どの体調がプレゼンティーイズムに影響し，労働生産性を低下させているのかを分析することができる。

表 1 プレゼンティーイズムとワーク・エンゲージメントの代表的な指標

調査対象	指標の種類
プレゼンティーイズム	Work Limitations Questionnaire[37]
	Stanford Presenteeism Scale[35]
	The Health and Work Performance Questionnaire[36]
	Work Productivity and Activity Impairment Questionnaire[38]
ワーク・エンゲージメント	Maslach Burnout Inventory-General Survey [39]
	Oldenburg Burnout Inventory [40]
	The Gallop Q ¹²
	Utrecht Work Engagement Scale [41]

3.3 労働生産性に影響のある体調不良の選定

体調不良にも様々な種類があるため、まず、取り扱う体調不良を選定する必要がある。プレゼンティーイズムとアブセンティーイズムの既存研究では、体調と労働生産性への影響に関して、調査しているものがある。本節ではその結果を用いて、労働生産性に影響のある体調不良を取り扱う。

Koji WADA らはSPSを用いて、19企業から計6,777人の回答を集計した[42]。この調査から、肩こり・腰痛による経済的損失が大きいことが判明した。また、症状別の一人当たりの損失は抑うつ・不安・感情の障害が最も経済的損失が大きいことを報告している。Collins,J.J. らは、SPSを用いて、1つの企業から計7,797人の回答を取得し、集計を行った[32]。この調査では、抑うつ、アレルギー、関節症、肩こり・腰痛が労働生産性に大きく影響していることが報告されている。アレルギー、肩こり・腰痛はSPSのスコア値が高かったが、状態ごとの経済的損失を試算した結果、他の症状に比べて有意な差は現れなかった。しかし、アレルギー、肩こり・腰痛を患っている従業員は多いため、1人あたりの経済的損失は小さくても、全体としての損失が大きい。Loeppke,R. らは、HPQを用いて、10企業を対象に51,648人から回答を取得、集計を行った[31]。プレゼンティーイズムでの影響は、鬱、不安、疲労、睡眠障害、過敏性腸症候群、頭痛が大きいこと

表 2 労働生産性に影響のある体調不良

労働生産性への影響の度合い	体調不良の種類
とても大きい	不安・抑うつ 肩こり・腰痛 鼻炎
大きい	疲労 睡眠障害 腹痛 頭痛

が報告されており，どの症状も年間 10 営業日以上分の経済的損失が試算されている。

これらの調査結果から労働生産性への影響がある体調の不具合に，不安・抑うつ，肩こり・腰痛，アレルギーがあり，影響が大きい体調の不具合に疲労，睡眠障害，過敏性腸症候群，頭痛があることが判明した．結果を表 2 にまとめる．アレルギーは，食べ物によるものではなく，花粉から生じるアレルギー性鼻炎を指すため，鼻炎に置き換えている．過敏性腸症候群は，体調不良ではなく疾患の意味合いが強いため，広義の腹痛に置き換えている．この中でも，不安・抑うつが最も生産性を下げる症状となっている．肩こり・腰痛，鼻炎は，他の症状とあまり大きな差はないが，患っている人数がとても多いため，全体の損失から見ると労働生産性へ大きく影響しているという理由から影響度合いを「とても大きい」に分類している。

3.4 体調の計測・推定技術

3.3 節では，過去のプレゼンティーイズム・ワークエンゲージメントの研究結果から労働生産性を下げる体調不良を調査し，不安・抑うつ，肩こり・腰痛，鼻炎，疲労，睡眠障害，腹痛，頭痛が労働生産性を下げることが判明した．次の手順として，これらの体調不良が既存技術でどの程度推測，推定できるかの調査を

行い、体調推定、体調予測に必要となるセンサ技術を明らかにする。頭痛に関しては、有効な既存技術を発見、考案することができなかったため、本研究では取り扱わない。

3.4.1 不安・抑うつ

不安・抑うつ状態になると、気分が落ち込み、仕事への活力がなくなってしまう。この状態は他の体調に比べ労働生産性を最も下げる要因となっている。不安・抑うつはストレスが原因であり、長時間労働、上司・部下との職場関係、家庭環境等と様々な要因がこの状態を引き起こす原因となっている。

不安・抑うつのようなメンタルヘルスの不調が生じると、活動量の低下、精神活動の低下のような変化が現れる。活動量の低下は、ウェアブル機器やスマートフォンのGPSや加速度計などを用いて測定することが容易であり、この手法によるメンタルヘルスの推定が行われている[43, 44]。また、SNSでのアクティビティを取り入れた推定も行われている[45]。

メンタルヘルスの不調の原因となるストレスを定量化する代表的な手法は複数あり、(1) 心電計から心拍間隔の時系列データを測定し、自律神経の乱れであるLF/HFを算出する手法、(2) 生理的なストレス量であるアミラーゼを唾液から測定する方法[46]、(3) 皮膚の電気抵抗を測定し、発汗量から緊張度を推定する手法[47]があげられる。(1)と(3)の手法によるストレス値の測定は、既にウェアブル機器での測定が可能である。間接的な推定方法として、PCの操作ログを用いてストレス量の推定を行う手法[48]や画像を選ぶだけでストレス量が推定できる手法[49]が提案されている。

3.4.2 肩こり・腰痛

肩こり・腰痛が生じる原因は単純ではなく、複雑な要素から生じている。大きく分けると、骨や筋肉から生じる器質的な痛み、神経の異常から生じる神経痛、神経的なストレスから生じる非器質的な痛みの3つに分類される。慢性的な肩こり・腰痛は、労働中の集中力を妨げる原因となり、生産性の低下に繋がると考える。

肩こり・腰痛を測る指標としては、筋肉の硬さを表す筋硬度が用いられているが、筋硬度は性別・運動経験の有無等の様々な要因により個人の特性を大きく反映するため、定期的に計測を行わなければ肩こり・腰痛を測定することはできない。パソコンを長時間使用すると、肩こり・腰痛が生じるリスクが高まることがわかっており [50]、症状の改善・防止として、姿勢を正すことを促すシステムが提案されている [51, 52]。

3.4.3 鼻炎

アレルギーや風邪が原因で生じる鼻炎は、労働生産性を大きく下げる原因とはならないが、患者が多く総合的な労働生産性の損失が大きい。鼻炎から頭痛が生じることもあり、集中力の欠如の原因となる。

鼻炎状態を検知する手法が見つからなかったため、間接的な推定方法を提案する。鼻炎になると、くしゃみの回数やちり紙で鼻をすする回数が増えるため、この音をマイクで検出することで、鼻炎かどうかの推定が行えると考ええる。Chenらは、スマートフォンのマイクからくしゃみと鼻をすする音を検出する技術を開発している [53]。また、くしゃみの動作をウェアラブル機器のIMU センサから識別することができれば、間接的に鼻炎状態を推定できる可能性がある。

3.4.4 疲労

十分な休息が取れない、または、睡眠の質が低いと身体が疲労しやすくなる。疲労状態になると、「反応速度が鈍くなる」、「思考力・集中力が低下する」、「行動力が低下する」、「自律神経に異常が生じる」といった状態に陥る。

疲労の計測は自動車の分野で研究が盛んに行なわれている。手法としては、(1) 脳波から推定する方法 [54]、(2) 外部のカメラから被験者の顔をトラッキングし、瞬きの回数、眼を閉じている長さ (PERCLOS)、頭の動き、視線の方向を計測し、疲労具合の推定を行う方法 [55]、(3) 心拍の変動 (HRV) から推定する方法 [56] があげられる。運動における疲労度であれば、足に IMU センサを取り付けて活動量を測定し推定する方法が提案されている [57]。

3.4.5 睡眠障害

睡眠障害を患うと、疲労の十分な回復が行えず、労働中に眠くなり、集中力が保てず労働生産性への低下へと繋がる。睡眠障害には不眠症、過眠症など様々な症状があるが、睡眠状態と睡眠の質の推定が行えれば、睡眠障害の検出は容易であると考えられる。睡眠状態の推定には、(1) 心拍から推定する方法、(2) 寝室の環境音から推定する方法、(3) 体動センサから推定する方法などがあり、それぞれ既存製品として広まっている。近年では、心拍計を搭載しているウェアラブル機器が増え、睡眠のロギング機能がついている。

3.4.6 腹痛

腹痛は、食が原因で生じるが、ストレスが原因で生じる場合もある。ストレスが原因の腹痛を特に過敏性腸症候群という。腹痛が生じれば、短時間ではあるが集中力が大いに欠如し、労働生産性を大きく下げることとなる。

腸の状態をモニタリングするデバイスとして、トリプル・ダブリュー・ジャパン株式会社の DFree[58] が利用されている。DFree は超音波を用いることで、腸の状態をセンシングすることができ、腹部の状態や排便のタイミングを推定することができる。この技術を用いれば、腸の状態の異常を検知し、腹痛の予測や検知が可能だと考えられる。

3.5 体調の計測・推定技術のまとめ

不安や抑うつには推定に取り組んだ事例を多く、既存の技術で測定できる可能性が十分高い。GPS のような移動軌跡やウェアラブルセンサの加速度から活発度を取ることができ、推定に役立てることができる。さらに、様々な手法によって鬱の原因となるストレスは高い精度で推定可能である。肩こり・腰痛に関しては、まだ研究が進んでおり、技術的な課題がある。筋硬度のような指標により、ウェアラブル機器以外で測定はある程度可能だが、より定量的で利用しやすい指標や手法が求められる。鼻炎に関しては、マイクによってくしゃみや鼻をすする音を収集することで判別することができそうであるが、そこまで踏み込んだ研究はま

表 3 体調の測定方法の調査結果

体調	ウェアラブル機器での測定	それ以外での測定
不安・抑うつ	○	○
肩こり・腰痛	×	△
鼻炎	△	△
疲労	○	○
睡眠障害	○	○
腹痛	△	△
頭痛	×	×

だない。しかしながら、マイクはスマートフォンやウェアラブル機器に装着してあるため、センシングは難しくなさそうである。疲労に関しては、既存技術が多く、様々な方法で推定できそうである。特に、HRV はウェアラブル機器でも収集することができ、ウェアラブル機器でもそれ以外の方法でもセンシングすることが十分可能であると考える。睡眠障害に関しては、睡眠計測の既存研究が多く、既にコンシューマ向けに様々なデバイスが販売されており、睡眠パターンを分析することで睡眠障害の判別は容易であると考え、腹痛に関しては、超音波系によるお腹のセンシングが必要となる。しかしながら、近年、お腹のセンシングが可能な小型デバイスが登場したことから、今後推定できる可能性がある。頭痛に関しては、センシングする技術がほとんどなく、現段階の技術では計測することは難しい。調査結果を表 3 に示す。「○」は測定・推定可能、「△」は測定・推定ができる可能性がある、「×」は測定・推定が難しい、または、既存研究がないを表す。

7 種類の体調のうち、不安・抑うつ、疲労、睡眠障害の 3 種類は心拍が計測できれば推定可能である。そのため、ウェアラブル機器に必要なセンサとして、心拍センサは重要となってくる。しかしながら、HF/LF 比のような詳細な心拍情報を収集することができるウェアラブル機器はあまり多くなく、コストが高い。例えば、*empatica* 社の E4 wristband[59] は反射型光電脈波センサ (PPG センサ) がついており、血流計測した生の値を分析できるが、一台\$1,690 (現在の日本円で約 18 万円) ととても高価である。研究の第一歩として、本研究では、1 万円程度

のより安価であり一般に普及しているウェアラブル機器を用いることにする。

4. ウェアラブル機器による体調の推定と予測

4.1 ウェアラブル機器によるライフログ収集

3章での調査結果から本研究では、心拍センサが装着されている市販のリストバンド型ウェアラブル機器を用いることにした。市販のウェアラブル機器を用いる理由としては、実用ベースでの検証が進められること、睡眠の認識機能があることがあげられる。市販のウェアラブル機器にも様々な種類があるが、本研究ではSONY社のSmartBand2²を使用した。

SmartBand2は、ライフログを目的としたウェアラブル機器で、心拍数、ストレスレベル、消費カロリーや歩数情報をユーザに提示してくれる。図1にSmartBand2を装着した様子を示す。Android, iOS両方のスマートフォンに対応しており、AndroidであればGoogle Fit³, iOSであればHealthKit⁴にライフログは蓄積され、取り出すことができる。以前は、公式のサーバにライフログは全て蓄積され、APIが提供されていたが、2017年にサーバが停止されてしまったため、別のサービス経由でないとライフログを取り出せなくなってしまった。睡眠の認識精度や電池の持ち具合(2日)、充電端子が独自の規格でない、防水防塵性などの理由から実験で使用することにする。SmartBand2には、加速度センサと光学式の心拍センサが内蔵されており、10分ごとに心拍の計測を行う。

4.2 主観的な体調に関するアンケート

体調の定量的な評価というのは難しく、現時点では自己の状態を自覚し主観的に評価するしかない。しかしながら、医者が患者と対話することで病を特定するように、主観的な健康というのは重要である。アメリカの18万人規模の調査[60]では、ストレスが健康に影響を与えると認識している人の方が、ストレスが健康に影響がないと認識している人と比べて、健康やメンタルヘルスの状態が悪く、

²SmartBand 2 (スマートバンド 2) SWR12 — スマートプロダクト — ソニーモバイル公式
サイト : <http://www.sonymobile.co.jp/product/smartproducts/swr12/>

³Fit Google : <https://www.google.com/intl/en/fit/>

⁴HealthKit - Apple Developer : <https://developer.apple.com/healthkit/>



図 1 SmartBand2

ストレスが実際に健康に大きく影響したときの早期死亡のリスクが43%も高いことが報告されている。このように、健康の認知が自身の健康に影響を及ぼすほど主観的な健康状態というのはセンシティブであるが、逆にいえば、主観的な健康状態というのは信頼性が高いといえる。

本研究のように体調に着目した研究はあまりなく、体調に関する質問指標がなかったため、本研究では体調の報告を行うアンケートを独自に作成した。表4に作成したアンケートを示す。

表 4 体調に関する質問項目

質問内容	評価方法
(a-1) 本日は健康的に過ごせた	5 段階
(a-2) 本日は規則正しく過ごせた	5 段階
(a-3) 前日の疲れが残っている	5 段階
(a-4) 本日の身体の調子はよかった	5 段階
(a-5) 本日は幸福に過ごせた	5 段階
(a-6) 本日のメンタルの調子はよかった	5 段階
(b-1) 本日は 4 時間以上作業した ¹	はい/いいえ
(b-2) 作業中の生産性は高かった	5 段階
(b-3) 作業中の集中力は高かった	5 段階
(b-4) 作業中に眠くなった	5 段階
(b-5) 作業中に疲れ・倦怠感を感じた	5 段階
(b-6) 作業中に不安になった	5 段階
(b-7) 作業中に抑うつ状態になった	5 段階
(b-8) 作業中に腹痛が生じた	5 段階
(b-9) 作業中に頭痛が生じた	5 段階
(b-10) 作業中に肩こりを感じた	5 段階
(b-11) 作業中に腰痛を感じた	5 段階
(b-12) 作業中に鼻炎になった	5 段階

¹ 「いいえ」を選んだ場合、アンケートは終了となる

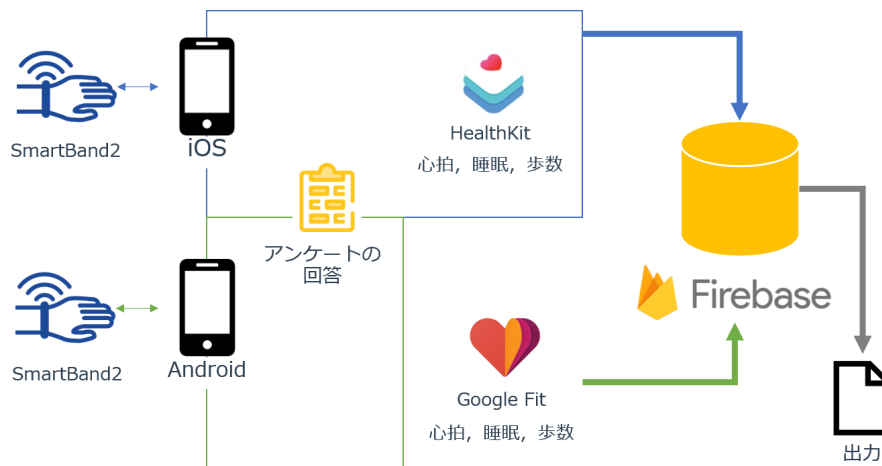


図 2 アンケート収集システムの概要

このアンケートは (a) と (b) の二つの質問項目群から構成される。(a) に該当する質問項目は、労働や作業に関係のない質問項目で一日の所感についての項目である。一方、(b) は労働や作業を行った日に回答する質問項目である。アンケートでは、一日のうち、作業を 4 時間以上行った場合のみ (b) の質問についての回答を行う。日本の一般的な労働時間は 8 時間であり、半日休暇のような就業規則が普及していることから、半分の 4 時間以上を作業した日とみなすことにした。全ての質問項目は 5 段階の Likert Scale で回答することができ、1 が「全く同意できない」、2 が「同意できない」、3 が「どちらともいえない」、4 が「同意できる」、5 が「非常に同意できる」と対応している。本研究ではこのアンケートを用いて実験協力者の主観的な体調データを収集する。

4.3 アンケート収集システム

毎日の体調データの収集は実験協力者への負荷が大きい。体調データを紙や表計算ソフトで収集した場合、回答という行為そのものがストレスになることは勿論、回答結果の管理も実験協力者に求めなければならない。そこで、全てのアンケートをスマートフォン上で完結させることで、回答の煩雑さと回答結果の管理のコストを減らすことができ、実験協力者への負荷を下げるができる。図 2 に開発したシステムの概要を示す。アンケートの回答アプリ「Anketo!」は、でき



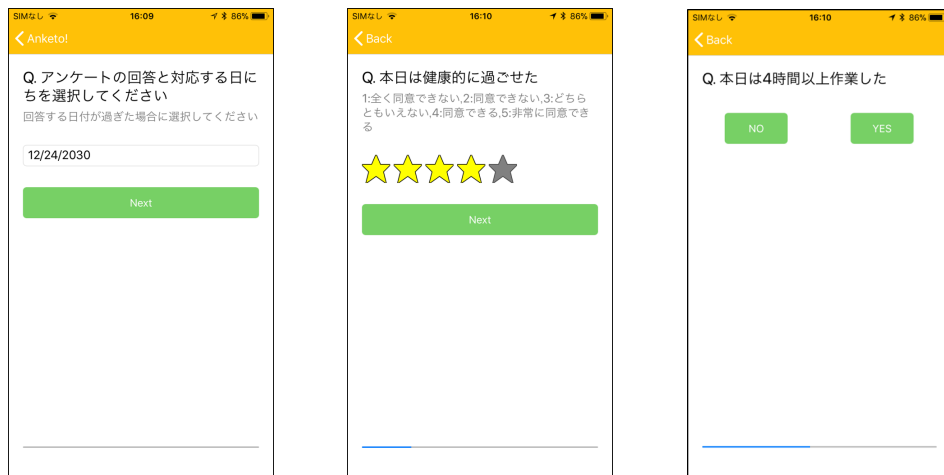
(a) Android 版



(b) iOS 版

図 3 アンケート回答アプリ「Anketo!」

るだけ多くのユーザに対応させるため、Android, iOS どちらでも動作するように開発した。そして、アンケートの回答結果は、インターネットを通じて DBaaS (Database as a Service) である Google 社の Firebase⁵ に送信される。アンケートの回答結果はセンシティブな情報であることから、セキュリティ対策が必要である。そこで、システムでは Firebase の認証機能を利用することでデータベースのアクセス権を制限しており、クライアントはセッショントークンを用いた通信をしなければならない。データベースでは、ユーザの認証情報、ユーザのプロファイル (割り当てられたアンケートや OAuth 認証した外部サービスの refresh token など)、アンケートのコンテンツ、アンケートの回答結果、iOS 版のアプリから送信されたライフログを管理している。



(a) 日付

(b) リッカート尺度

(c) 二者択一

図 4 アンケートの質問項目の種類

4.3.1 アンケート回答アプリ「Anketo!」

Anketo!はアンケートの回答を目的としたスマートフォンアプリケーションである。アンケートの回答結果は全てインターネットを介してクラウドデータベースに送信される。Anketo!はクロスプラットフォームのネイティブアプリ環境である Xamarin⁶ で開発されているため、iOS、Android 共に動作可能である。図 3 に両 OS で動作させたときのトップ画面を示す。ネイティブの UI の仕様により、パーツの細かい点は異なるが、同じ機能を提供していることがわかる。

次に、本アプリケーションで利用可能なアンケートの種類の一部を図 4 に示す。研究で用いるアンケートでは、日付、リッカート尺度、はい・いいえの二者択一の 3 種類を使用しているが、他に数値、文字などの質問項目も利用できる。

4.3.2 ライフログの収集

SmartBand2 は直接データを取り出せる API を公式的に提供していないため、他のサービスを経由して取り出す必要がある。Android の場合では、Google Fit と

⁵Firebase : <https://firebase.google.com/>

⁶Mobile App Development & App Creation Software - Xamarin : <https://www.xamarin.com/>

いう Google 社が提供するヘルスケアアプリと SmartBand2 を連携させることで、SmartBand2 のデータが Google Fit のサーバへ送信されるようになる。そして、Google Fit と OAuth 認証することにより、サーバのデータへとアクセスできる。この OAuth 認証は、Anketo! の設定画面で行える。iOS の場合では、SmartBand2 のデータは iPhone の HealthKit というヘルスデータ管理システムへと送信される。Anketo! がヘルスデータにアクセスすることを許可すれば、生のライフログを取り出せる。iOS の HealthKit は Google Fit と異なり、ローカルにヘルスデータが保持されている。iOS を所持する実験協力者には、実験終了後ヘルスデータのアップロードを行ってもらい、Firebase にデータを収集する。

4.4 体調の推定・予測手法

収集した体調アンケートとライフログを用いて、一日の体調の推定と次の日の体調の予測を行う。アンケートは、毎日 20 時を目安に回答してもらうため、一日の区切りを 20 時にする。当日ではなく数日間 of 行動や睡眠量が体調へ影響することが考えられるため、3 日分データから時間、生活周期に関する特徴量を抽出し推定・予測に使用する。図 5 に推定と予測に使用するデータの時間の違いをまとめた。現在を d_0 日とすると、推定に使用するデータは $d_{-2} \sim d_0$ の 3 日分のデータとなる。一方、予測は一日後の体調の予測を行うことから使用するデータは $d_{-3} \sim d_{-1}$ の 3 日分のデータとする。

推定・予測にはそれぞれ機械学習を用いる。まず、前処理として、データの特徴量 X は平均 μ 、分散 σ から以下の式で標準化を行う。

$$z = \frac{X - \mu}{\sigma} (\sigma > 0) \quad (1)$$

$\sigma = 0$ のときは、 $z = 0$ と操作する。

推定・予測を試みる体調の項目として、「フィジカルヘルス」、「メンタルヘルス」、「体調不良の度合い」の 3 種類を対象とする。「フィジカルヘルス」は、表 4 の質問項目の「(a-4) 本日の身体の調子はよかった」、「メンタルヘルス」は「(a-6) 本日のメンタルの調子はよかった」の回答結果である。(a-4) は 1 日のフィジカルの状態、(a-6) は 1 日のメンタルの状態を総合的に判断した結果である。5 段階のリッ

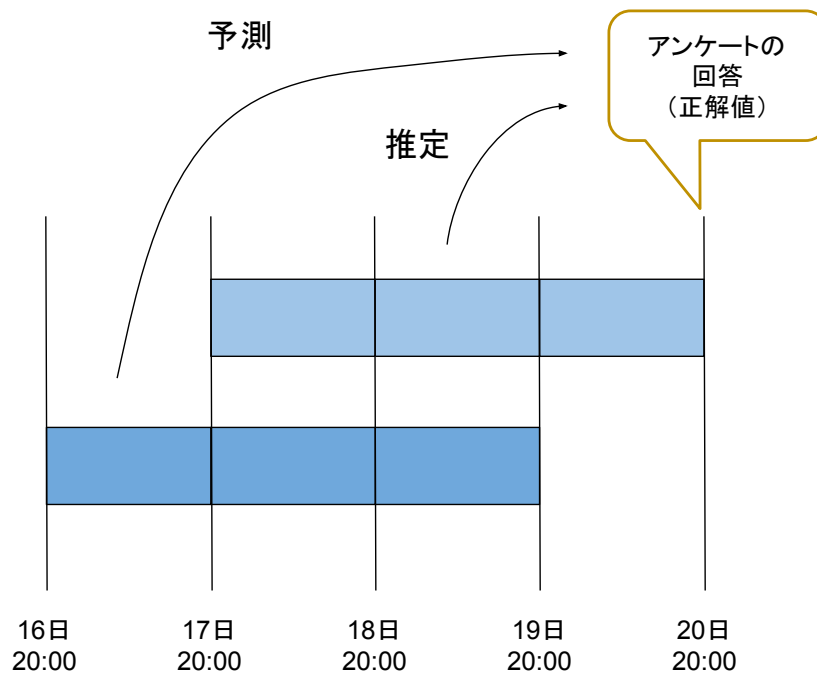


図 5 体調の推定・予測に利用するデータの違い

カート尺度の 1-2 の値を「悪い」、3 の値を「普通」、4-5 の値を「良い」とした分類の問題として置き換えることにした。なお、目的変数の特徴量は、決定木による特徴量選択により、情報量の高い特徴量のみで学習モデルを構築する。体調が優れない日が頻繁に起こることは多くなく、「悪い」のラベルのデータが少なくなる。ラベルが不均一なデータはモデルの誤学習を引き起こしやすいと共に、正確な識別性能の評価が難しくなる。そこで、オーバーサンプリングを行うことによりラベルの数が不均一な状態から均一な状態へと整える。オーバーサンプリングには Synthetic Minority Over-sampling Technique (SMOTE)[61] を使用する。

「体調不良の度合い」では、「(b-4) 作業中に眠くなった」、「(b-5) 作業中に疲れ・倦怠感を感じた」、「(b-6) 作業中に不安になった」、「(b-7) 作業中に抑うつ状態になった」、「(b-8) 作業中に腹痛が生じた」、「(b-9) 作業中に頭痛が生じた」、「(b-10) 作業中に肩こりを感じた」、「(b-11) 作業中に腰痛を感じた」、「(b-12) 作業中に鼻炎になった」の労働生産性を下げる体調のアンケートに結果から、体調不良のスコアを算出し、回帰によってその値の推定と予測を行う。スコアとしては、(1) 平

均のスコア と (2) 労働生産性への影響度合の重みを利用したスコアの2種類を検証する。(2)に関しては、目的変数を「(b-2) 作業中の生産性は高かった」のスコアとし、説明変数を8つのスコアとして、重回帰モデルを構築し、得られた重みを使うことで、個人の体調と生産性の影響度を考慮したスコアにしている。学習の際は、まず、決定木による特徴量選択によって、重要な特徴量を選択し、出力が実数である回帰モデルを構築する。

4.4.1 ライフログからの特徴量抽出

収集したライフログには心拍、歩数、睡眠が含まれている。心拍は1秒間当たりの脈拍数、歩数は1分間あたりの歩数、睡眠は開始時刻と終了時刻が記録される。ここから推定・予測に使用する特徴量を3日分のデータから780種類抽出する。

まずは、20時を基準とし、3時間ごと(8パターン)、6時間ごと(4パターン)、12時間ごと(2パターン)、24時間(1パターン)ごとに区切ったデータから、心拍は平均、標準偏差、最小、最大、第一四分位数、中央値、第三四分位数の7種類、歩数は心拍と同じ7種類と歩数の合計の計8種類、睡眠は睡眠時間の中央値[分]の1種類の特徴量を抽出する。さらに、1日前の同じ特徴量、2日前の同じ特徴量も含める。ウェアラブル機器を装着していない時間は、他の同時刻のデータ全てを用いてを代わり算出する。ウェアラブル機器を装着しているかどうかのデータは得られないため、心拍データの有無によって判別する。機器を装着している場合、およそ10分に1回心拍をセンシングすることから、ある時間幅の装着時間が50%未満であれば、ウェアラブル機器を装着していないと判断する。例えば、3時間の幅では9回未満、6時間であれば18回未満であれば未装着とする。長い時間を見た特徴量として、1日前から基準の日、2日前から1日前、2日前から基準日までの3パターンに対しても同様の特徴量を抽出する。

生活周期に関する特徴量で、データの基準日と1日前、基準日と2日前、1日前と2日前の心拍、歩数、睡眠の相関を求め、得られた9つの相関係数を特徴量とする。相関係数が高いほど同じ生活周期であることを意味する。それと、自己相関によって、3日間の生活周期に関する特徴量の抽出も行う。自己相関は時系列データ自身を畳み込みさせ、相関を算出していく手法で、おおよその時系列の

周期を求めることができる．データ数が N であり 1 時間ごとの離散系列データ x_i の自己相関関数は

$$R_{xx}(i) = \frac{1}{N/2} \sum_{j=0}^{N/2} x_{i-j} x_j \quad (2)$$

で表される．12 時間以内の周期は無視したいため， $i > 12$ においての $R_{xx}(i)$ の極大を求める．値が最も大きな極大値の位置 k が求めたい生活周期であり，この k を特徴量とする．これについて，心拍，歩数，睡眠においてそれぞれ求める．なお，極大が見つからない場合は $k = 0$ とする．

5. ライフログの収集実験と推定・予測評価

5.1 ライフログと体調の収集実験

ウェアラブル機器のライフログと体調の評価データを得るための実験を行った。実験では、Ecological Momentary Assessment (EMA)[62] を採用し、実験協力者にウェアラブル機器の SmartBand2 を装着してもらい、普段の生活からデータを収集する。4章で述べたスマートフォンアプリを各被験者自身が所有するスマートフォンにインストールし、毎日、20 時以降アンケートを回答してもらう。実験には、10 名 (男性, 24.2 ± 2.0 歳) が参加し、Android ユーザは 4 名、iPhone ユーザは 6 名であった。実験期間は、2017 年の 11 月 23 日から 12 月 10 日までの約 1 ヶ月半行われた。本実験は、本学の倫理審査会の承認を得ている。

実験から、320 日分のライフログ (7 名分) が収集され、310 日分のアンケート回答 (10 名分) が得られた。アンケートのうち、4 時間以上作業した日のデータが 192 件、作業していない日が 118 件であった。3 人の実験協力者に関しては、誤ったアプリの削除、スマートフォンの破損によりライフログの収集ができなかった。

5.2 フィジカルヘルスの推定・予測結果

表 4 の質問項目の「(a-4) 本日の身体の調子はよかった」の評価結果を、1-2 の値を「悪い」、3 の値を「普通」、4-5 の値を「良い」とした 3 クラスに分割し、正解値として推定モデルと予測モデルの構築を行った。SMOTE フィルタによってオーバーサンプリングし、ラベルの数を均一にし、学習・評価している。3 クラスとも最低 3 つ以上のデータがない実験協力者のデータは、オーバーサンプリングしても酷似した値が生成されることから評価から除外する。ここでは 2 名のデータを除外した。分類モデルとしては、分類性能が高いとされる Random Forest, AdaBoost, Support Vector Machine の 3 種類を用いた。Leave-One-Out cross-validation で検証し、Precision, Recall, F-score を算出した。フィジカルヘルスの推定結果を表 5、予測結果を表 6 に示す。この結果は被験者の推定精度を平均したものである。

表 5 フィジカルヘルスの推定結果

分類機	Precision	Recall	F-score
Random Forest	78.2%	77.7%	77.6%
AdaBoost	65.9%	67.0%	65.2%
Support Vector Machine	89.3%	88.1%	88.1%

表 6 フィジカルヘルスの予測結果

分類機	Precision	Recall	F-score
Random Forest	73.9%	74.4%	74.0%
AdaBoost	67.2%	66.0%	65.0%
Support Vector Machine	85.4%	83.8%	84.0%

フィジカルヘルスの推定結果の F-score は、Random Forest が 77.6%，AdaBoost が 65.2%，Support Vector Machine が 88.1% の精度となっており，分類機としては，Support Vector Machine が最も良く．対して，予測結果の F-score は，Random Forest が 74.6%，AdaBoost が 65.0%，Support Vector Machine が 84.0% の精度となっており，こちらも Support Vector Machine での分類精度が最も高い．個別にデータを見ると，AdaBoost では，38% の識別率しかない場合があり，一部の実験協力者のデータを上手く学習できていないため，識別率が低くなっていた．また，推定精度の最大が 88.1% で，予測精度の最大が 84.0% であり，予測より推定の精度の方が良い．それぞれのモデルの F-score を平均しても 7 割後半の精度があり，3 クラスの分類において，良い精度で推定・予測できている．

5.3 メンタルヘルスの推定・予測結果

フィジカルヘルスと同様に表 4 の質問項目の「(a-6) 本日のメンタルの調子はよかった」の 1-2 の値を「悪い」，3 の値を「普通」，4-5 の値を「良い」とした 3 クラスに分割し，正解値として推定モデルと予測モデルの構築を行った．使用した分類モデルと検証方法，オーバーサンプリング処理もフィジカルヘルスの場合

表 7 メンタルヘルスの推定結果

分類機	Precision	Recall	F-score
Random Forest	85.2%	84.8%	84.6%
AdaBoost	65.3%	61.7%	61.3%
Support Vector Machine	94.0%	93.2%	93.1%

表 8 メンタルヘルスの予測結果

分類機	Precision	Recall	F-score
Random Forest	88.5%	87.4%	87.5%
AdaBoost	69.2%	69.3%	68.2%
Support Vector Machine	94.6%	94.2%	94.1%

と同様であり、2名のデータを除外している。フィジカルの調子の推定結果を表 7、予測結果を表 8 に示す。この結果は被験者の推定精度を平均したものである。

メンタルヘルスの推定結果の F-score は、Random Forest が 84.6%、AdaBoost が 61.3%、Support Vector Machine が 93.1% の精度となっており、分類機としては、Support Vector Machine が最も良い。予測結果の F-score は、Random Forest が 87.5%、AdaBoost が 68.2%、Support Vector Machine が 94.1% の精度となっており、フィジカルヘルスと同様に Support Vector Machine での分類精度が最も高く、AdaBoost が最も低い。また、推定精度の最大が 93.1% で、予測精度の最大が 94.1% であり、予測より推定の精度が若干良い。それぞれのモデルの F-score を平均しても 8 割程度の精度があり、3 クラスの分類においてはある程度高い精度で推定・予測できるといえる。

フィジカルヘルスの結果と比べると、最大精度の Support Vector Machine においては、メンタルヘルスの精度の方が 10% 高くなっている。精度が高くなった理由として考えられるのが、3.4.1 項で取り上げたように、メンタルヘルスの不調と活動量に関係があることから、歩数の特徴量と、睡眠障害は鬱との関係性がある [63] ため、睡眠の特徴量がフィジカルに比べ精度の向上に寄与したと考えられる。

5.4 体調不良の度合いの推定と予測評価結果

4.4 節で述べたように、3 章で選定した体調不良に関するアンケート結果を (1) 平均のスコア、(2) 労働生産性の影響度合の重みを利用したスコアの 2 種類で表現をし、特徴量選択の後に、回帰モデルの Random Forest, AdaBoost, Support Vector Machine で学習させ、One-Leave-Out cross-validation を使い決定係数 R^2 と Explained Variance Score (EVS) で評価したところ、ほぼすべての場合において、回帰モデルに当てはまらず、推定や予測が全く行えなかった。(1) のスコアの推定においては R^2 と EVS が 0.37, (1) のスコアの予測で R^2 と EVS が 0.32, (2) のスコアの推定においては R^2 と EVS が 0.20, (1) のスコアの予測で R^2 と EVS が 0.25 が最大であった。なお、これらは全て同じ実験協力者である。

モデルの当てはまりがうまくいかない原因としては、(a) 作業をした時にしか体調不良は評価されないため全体データが少ないこと、(b) 全ての体調不良を一つのスコアとして表現していること、(c) 体調不良が生じることが少ないため過学習していることがあげられる。

6. 結論

本研究では、労働生産性の低下の抑止に向けて、ウェアラブル機器によって体調の推定・予測が行えるかどうか検証・評価した。まず、プレゼンティーイズムの観点から体調と労働生産性との関連について調査し、心拍センサが必要であること確認し、心拍センサが取り付けられているウェアラブル機器の選定を行った。ウェアラブル機器から得られる心拍、睡眠、歩数のライフログを収集する基盤と、体調の正解値とするためのアンケートとアンケートをスマートフォンで回答できるシステムを開発した。ライフログから時間や生活周期に関する特徴量を3日分抽出し、アンケートの回答項目であるフィジカルヘルス（3クラス：悪い、普通、良い）、メンタルヘルス（3クラス：悪い、普通、良い）、体調不良の度合いに対して、それぞれの推定・予測モデルを構築した。構築したモデルを評価した結果、フィジカルヘルスは、88.1%の精度で推定、84.0%の精度で予測でき、メンタルヘルスは、93.2%の精度で推定、94.1%の精度で予測できた。フィジカルヘルスの精度よりメンタルヘルスの精度の方が推定・予測どちらとも10%ほど高く、ウェアラブル機器のみであっても、十分な精度で心身の状態の推定・予測できることが明らかとなった。

しかしながら、体調不良の度合いについては、データ量が少ないことやスコアの偏りによって、良いモデルの構築が行えなかった。そのため、残された課題として、ライフログ収集基盤の改良、体調不良のより細かな度合の推定と予測を挙げる。実験では、3名のライフログが誤って消去、破損されてしまったため、実験の評価に利用でき無いという問題が発生した。iOSに関しては、ローカルでデータが保持されているため、クラウド上にライフログを自動的にアップロードする仕組みを開発する必要があると考えられる。体調不良の細かな度合に関しては、体調を適切な数値で表し、回帰モデルを構築することで、細かな体調の度合いが推定・予測できるかを検証する必要がある。

今後の課題として、データ量が少ないことによる問題を解決するため、実験を続け、より多くのデータを収集して、体調の推定・予測の検証をより詳細に行うことが考えられる。また、どのような特徴量が推定・予測に有用なのか明らかにし、体調不良にならず、改善するためのノウハウを自動的に抽出し、生活改善を

促すシステムへとつなげていく必要がある。さらに、労働生産性についての推定や予測に関する技術の開発を行うことで、労働生産性の低下を抑止するためのシステムの実現に繋がるものと考えている。

謝辞

本研究を進めるにあたり、主指導教員である安本慶一教授には、熱心なご指導を賜りました。体調という難しいテーマについて、何回も協議し熱心に指導してくださいました。副指導教員である池田和司教授には、多忙の中論文審査委員を引き受けていただくとともに、中間発表や本審査での貴重な助言をくださいました。副指導教員である荒川豊准教授には、実験に必要なデバイスの準備、環境の構築のサポートしてくださいました。副指導教員である水本旭洋特任助教には、研究会論文の執筆にて細かい指導をしてくださいました。研究室の諏訪博彦助教、藤本まなと助教には、ゼミや研究発表会の練習等で助言やコメントしてくださいました。また、博士前期課程1年の渡邊洸君には、ドイツ留学中に日本での実験のサポートをしていただきました。本研究に携わってくださった皆様に心からの感謝の気持ちを込めてお礼申し上げます。

参考文献

- [1] Christopher D Wickens, Justin G Hollands, Simon Banbury, and Raja Parasuraman. *Engineering psychology & human performance*. Psychology Press, 2015.
- [2] Jacek Gwizdka. What 's in the context. In *Computer Human Interaction*, 第 2000 卷, pp. 1–4, 2000.
- [3] Ling Bao and Stephen S Intille. Activity recognition from user-annotated acceleration data. In *International Conference on Pervasive Computing*, pp. 1–17. Springer, 2004.
- [4] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, Vol. 12, No. 2, pp. 74–82, 2011.
- [5] Masashi Fujiwara, Yukitoshi Kashimoto, Manato Fujimoto, Hirohiko Suwa, Yutaka Arakawa, and Keiichi Yasumoto. Implementation and evaluation of analog-pir-sensor-based activity recognition. *SICE Journal of Control, Measurement, and System Integration*, Vol. 10, No. 5, pp. 385–392, 2017.
- [6] Eri Nakagawa, Kazuki Moriya, Hirohiko Suwa, Manato Fujimoto, Yutaka Arakawa, and Keiichi Yasumoto. Toward real-time in-home activity recognition using indoor positioning sensor and power meters. In *Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on*, pp. 539–544. IEEE, 2017.
- [7] Akane Sano, Andrew J Phillips, Z Yu Amy, Andrew W McHill, Sara Taylor, Natasha Jaques, Charles A Czeisler, Elizabeth B Klerman, and Rosalind W Picard. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones.

In *Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on*, pp. 1–6. IEEE, 2015.

- [8] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 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*, pp. 3–14. ACM, 2014.
- [9] Matteo Ciman and Katarzyna Wac. Individuals’ stress assessment using human-smartphone interaction analysis. *IEEE Transactions on Affective Computing*, 2016.
- [10] Sara Ann Taylor, Natasha Jaques, Ehimwenma Nosakhare, Akane Sano, and Rosalind Picard. Personalized multitask learning for predicting tomorrow’s mood, stress, and health. *IEEE Transactions on Affective Computing*, 2017.
- [11] Kiran K Rachuri, Mirco Musolesi, Cecilia Mascolo, Peter J Rentfrow, Chris Longworth, and Andrius Aucinas. Emotionsense: a mobile phones based adaptive platform for experimental social psychology research. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pp. 281–290. ACM, 2010.
- [12] Andrey Bogomolov, Bruno Lepri, and Fabio Pianesi. Happiness recognition from mobile phone data. In *Social Computing (SocialCom), 2013 International Conference on*, pp. 790–795. IEEE, 2013.
- [13] Leo Breiman. Random forests. *Machine learning*, Vol. 45, No. 1, pp. 5–32, 2001.
- [14] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, Vol. 55, No. 1, pp. 119–139, 1997.

- [15] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, Vol. 20, No. 3, pp. 273–297, 1995.
- [16] Daniel J Buysse, Charles F Reynolds, Timothy H Monk, Susan R Berman, and David J Kupfer. The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry research*, Vol. 28, No. 2, pp. 193–213, 1989.
- [17] Sheldon Cohen, Tom Kamarck, and Robin Mermelstein. A global measure of perceived stress. *Journal of health and social behavior*, pp. 385–396, 1983.
- [18] John E Ware Jr, Mark Kosinski, and Susan D Keller. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Medical care*, Vol. 34, No. 3, pp. 220–233, 1996.
- [19] John P Pollak, Phil Adams, and Geri Gay. Pam: a photographic affect meter for frequent, in situ measurement of affect. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 725–734. ACM, 2011.
- [20] Kurt Kroenke, Robert L Spitzer, and Janet BW Williams. The phq-9. *Journal of general internal medicine*, Vol. 16, No. 9, pp. 606–613, 2001.
- [21] Ed Diener, Derrick Wirtz, William Tov, Chu Kim-Prieto, Dong-won Choi, Shigehiro Oishi, and Robert Biswas-Diener. New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, Vol. 97, No. 2, pp. 143–156, 2010.
- [22] Daniel W Russell. UCLA loneliness scale (version 3): Reliability, validity, and factor structure. *Journal of personality assessment*, Vol. 66, No. 1, pp. 20–40, 1996.
- [23] Oliver P John and Sanjay Srivastava. The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory*

and research, Vol. 2, No. 1999, pp. 102–138, 1999.

- [24] Robert LiKamWa, Yunxin Liu, Nicholas D Lane, and Lin Zhong. Moodscope: Building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pp. 389–402. ACM, 2013.
- [25] James A Russell. Evidence of convergent validity on the dimensions of affect. *Journal of personality and social psychology*, Vol. 36, No. 10, p. 1152, 1978.
- [26] Anil Jain and Douglas Zongker. Feature selection: Evaluation, application, and small sample performance. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 19, No. 2, pp. 153–158, 1997.
- [27] Ziad Ahmad, Marc Bosch, Nitin Khanna, Deborah A Kerr, Carol J Boushey, Fengqing Zhu, and Edward J Delp. A mobile food record for integrated dietary assessment. In *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management*, pp. 53–62. ACM, 2016.
- [28] Kiyoharu Aizawa and Makoto Ogawa. Foodlog: Multimedia tool for health-care applications. *IEEE MultiMedia*, Vol. 22, No. 2, pp. 4–8, 2015.
- [29] 【カロミル】人工知能（ai）がダイエットと健康管理をサポート. <https://calomeal.com/>. [Online; accessed 29-January-2018].
- [30] E Brian Faragher, Monica Cass, and Cary L Cooper. The relationship between job satisfaction and health: a meta-analysis. *Occupational and environmental medicine*, Vol. 62, No. 2, pp. 105–112, 2005.
- [31] Ronald Loeppke, Michael Taitel, Vince Haufle, Thomas Parry, Ronald C Kessler, and Kimberly Jinnett. Health and productivity as a business strategy: a multiemployer study. *Journal of Occupational and Environmental Medicine*, Vol. 51, No. 4, pp. 411–428, 2009.

- [32] James J Collins, Catherine M Baase, Claire E Sharda, Ronald J Ozminkowski, Sean Nicholson, Gary M Billotti, Robin S Turpin, Michael Olson, and Marc L Berger. The assessment of chronic health conditions on work performance, absence, and total economic impact for employers. *Journal of Occupational and Environmental Medicine*, Vol. 47, No. 6, pp. 547–557, 2005.
- [33] Geppo - 従業員のコンディション変化発見ツール. <https://www.geppo.jp/>. [Online; accessed 29-January-2018].
- [34] Akihito Shimazu and Wilmar B Schaufeli. Is workaholism good or bad for employee well-being? the distinctiveness of workaholism and work engagement among japanese employees. *Industrial health*, Vol. 47, No. 5, pp. 495–502, 2009.
- [35] Cheryl Koopman, Kenneth R Pelletier, James F Murray, Claire E Sharda, Marc L Berger, Robin S Turpin, Paul Hackleman, Pamela Gibson, Danielle M Holmes, and Talor Bendel. Stanford presenteeism scale: health status and employee productivity. *Journal of occupational and environmental medicine*, Vol. 44, No. 1, pp. 14–20, 2002.
- [36] Ronald C Kessler, Catherine Barber, Arne Beck, Patricia Berglund, Paul D Cleary, David McKenas, Nico Pronk, Gregory Simon, Paul Stang, T Bedirhan Ustun, et al. The world health organization health and work performance questionnaire (hpq). *Journal of Occupational and Environmental Medicine*, Vol. 45, No. 2, pp. 156–174, 2003.
- [37] Debra Lerner, Benjamin C Amick III, William H Rogers, Susan Malspeis, Kathleen Bungay, and Diane Cynn. The work limitations questionnaire. *Medical care*, Vol. 39, No. 1, pp. 72–85, 2001.
- [38] Margaret C Reilly, Arthur S Zbrozek, and Ellen M Dukes. The validity and reproducibility of a work productivity and activity impairment instrument.

Pharmacoeconomics, Vol. 4, No. 5, pp. 353–365, 1993.

- [39] Christina Maslach, Susan E Jackson, Michael P Leiter, Wilmar B Schaufeli, and Richard L Schwab. *Maslach burnout inventory*, Vol. 21. Consulting Psychologists Press Palo Alto, CA, 1986.
- [40] Evangelia Demerouti and Arnold B Bakker. The oldenburg burnout inventory: A good alternative to measure burnout and engagement. *Handbook of stress and burnout in health care*. Hauppauge, NY: Nova Science, 2008.
- [41] Wilmar B Schaufeli, Marisa Salanova, Vicente González-Romá, and Arnold B Bakker. The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness studies*, Vol. 3, No. 1, pp. 71–92, 2002.
- [42] Koji Wada, Mikako Arakida, Rika Watanabe, Motomi Negishi, Jun Sato, and Akizumi Tsutsumi. The economic impact of loss of performance due to absenteeism and presenteeism caused by depressive symptoms and comorbid health conditions among japanese workers. *Industrial health*, Vol. 51, No. 5, pp. 482–489, 2013.
- [43] Sohrab Saeb, Mi Zhang, Christopher J Karr, Stephen M Schueller, Marya E Corden, Konrad P Kording, and David C Mohr. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *Journal of medical Internet research*, Vol. 17, No. 7, 2015.
- [44] Luca Canzian and Mirco Musolesi. 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*, pp. 1293–1304. ACM, 2015.
- [45] Shoya Ishimaru and Koichi Kise. Quantifying the mental state on the basis of physical and social activities. In *Adjunct Proceedings of the 2015 ACM*

International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, pp. 1217–1220. ACM, 2015.

- [46] 山口昌樹, 金森貴裕, 金丸正史, 水野康文, 吉田博. 唾液アミラーゼ活性はストレス推定の指標になり得るか. *医用電子と生体工学*, Vol. 39, No. 3, pp. 234–239, 2001.
- [47] Helen Koo, Ivan Hebrío, Megan Johnston, Nicholas Hosein, and Kris Fallon. Stresssense: skin conductivity monitoring garment with a mobile app. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pp. 487–490. ACM, 2016.
- [48] 鳥羽美奈子, 櫻井隆雄, 森靖英, 恵木正史ほか. オフィスワークのストレス量と pc 操作ログ特徴量の重回帰分析— pc 操作ログ分析サービスの応用に向けて—. *デジタルプラクティス*, Vol. 7, No. 1, pp. 71–79, 2016.
- [49] Sophia Haim, Rui Wang, Sarah E Lord, Lorie Loeb, Xia Zhou, and Andrew T Campbell. The mobile photographic stress meter (mpsm): A new way to measure stress using images. In *Adjunct proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2015 ACM international symposium on wearable computers*, pp. 733–742. ACM, 2015.
- [50] Paula T Hakala, Arja H Rimpelä, Lea A Saarni, and Jouko J Salminen. Frequent computer-related activities increase the risk of neck–shoulder and low back pain in adolescents. *The European Journal of Public Health*, Vol. 16, No. 5, pp. 536–541, 2006.
- [51] Rushil Khurana, Elena Marinelli, Tulika Saraf, and Shan Li. Neckgraffe: a postural awareness system. In *CHI’14 Extended Abstracts on Human Factors in Computing Systems*, pp. 227–232. ACM, 2014.

- [52] Hosub Lee, Young Sang Choi, Sunjae Lee, and Eunsoo Shim. Smart pose: mobile posture-aware system for lowering physical health risk of smartphone users. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pp. 2257–2266. ACM, 2013.
- [53] Nan-Chen Chen, Kuo-Cheng Wang, and Hao-Hua Chu. Listen-to-nose: a low-cost system to record nasal symptoms in daily life. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 590–591. ACM, 2012.
- [54] Chin-Teng Lin, Ruei-Cheng Wu, Sheng-Fu Liang, Wen-Hung Chao, Yu-Jie Chen, and Tzyy-Ping Jung. Eeg-based drowsiness estimation for safety driving using independent component analysis. *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 52, No. 12, pp. 2726–2738, 2005.
- [55] Hao Yin, Yuanqi Su, Yuehu Liu, and Danchen Zhao. A driver fatigue detection method based on multi-sensor signals. In *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*, pp. 1–7. IEEE, 2016.
- [56] Udo Trutschel, Christian Heinze, Bill Sirois, Martin Golz, David Sommer, and David Edwards. Heart rate measures reflect the interaction of low mental workload and fatigue during driving simulation. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 261–264. ACM, 2012.
- [57] Marcus Schmidt, Carl Christian Rheinländer, Sebastian Wille, Norbert Wehn, and Thomas Jaitner. Imu-based determination of fatigue during long sprint. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pp. 899–903. ACM, 2016.
- [58] D free — maintain human dignity. <http://dfree.biz/>. [Online; accessed 7-November-2016].
- [59] Real-time physiological signals — e4 eda/gsr sensor. <https://www.empatica.com/research/e4/>. [Online; accessed 29-January-2018].

- [60] Abiola Keller, Kristin Litzelman, Lauren E Wisk, Torsheika Maddox, Erika Rose Cheng, Paul D Creswell, and Whitney P Witt. Does the perception that stress affects health matter? the association with health and mortality. *Health Psychology*, Vol. 31, No. 5, p. 677, 2012.
- [61] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, Vol. 16, pp. 321–357, 2002.
- [62] Saul Shiffman, Arthur A Stone, and Michael R Hufford. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.*, Vol. 4, pp. 1–32, 2008.
- [63] Yoshitaka Kaneita, Takashi Ohida, Makoto Uchiyama, Shinji Takemura, Kazuo Kawahara, Eise Yokoyama, Takeo Miyake, Satoru Harano, Kenshu Suzuki, and Toshiharu Fujita. The relationship between depression and sleep disturbances: a japanese nationwide general population survey. *The Journal of clinical psychiatry*, 2006.