

行動データに基づいた オフィスワーカーの心理的指標の分析に向けて

西村 勇亮¹ Tahera Hossain¹ 磯村 昇太² 荒川 豊³ 井上 創造¹

概要: 近年、多くの組織においてメンタルヘルス状況を認識し、改善するための取り組みが重要視されている。特にオフィスワーカーは、多くのストレスに影響を受けており、心身の疲弊が社会問題にもなっている。心理状況を改善するためには、原因を明らかにする必要がある。本稿では、リストバンド型端末とスマートフォンを用いて、センサ、行動、アンケートによる心理指標のデータを、約 100 名の知識労働を行う会社員に配布し、14 日間で約 1400 日分のデータを収集する実験を行った。そして、行動、センサ、天気などのデータからオフィスワーカーの 6 種類の心理指標の予測を行った。特に、行動（業務内容や業務環境）と心理指標との相関を分析した。結果、22 項目の心理指標の予測において 17 項目が 80% の以上の精度で予測することができた。また、行動と心理状況との相関について明らかにすることで心理状態向上のための業務改善に貢献する知見を得た。

Toward the Analysis of Office Worker's Mental Indicators Based on Activity Data

YUSUKE NISHIMURA¹ TAHERA HOSSAIN¹ SHOTA ISOMURA² YUTAKA ARAKAWA³
SOZO INOUE¹

1. はじめに

近年、メンタルヘルスの改善やウェルビーイングによって私生活や仕事を向上させることが注目されており、その実現に向けた研究や開発が盛んに行われている。特に、働き方改革に取り組む企業は増加しており、社員の労働環境の向上が求められてきている。厚生労働省によると、メンタルヘルス対策に取り組んでいる事業所の割合は 59.2 % (2018 年) となっており 2012 年と比べて約 12 % 増加している [1]。また、ストレスチェックを行い、分析結果を活用している事業所も増えている。

企業が従業員の労働におけるメンタルヘルスの向上に取り組むべき理由はいくつかある。1 つ目は、社員の退職率の低下である。仕事への満足度が低い職場では高い職場と比べて退職率が上がる傾向にある [2]。つまり、仕事への満足度が低下することで人材不足に直面する可能性が高くな

る。2 つ目は、社員の健康維持である。職場において心理的ストレスを強く受けている場合、精神的な病気だけでなく肉体的な病気へのリスクを高めている [3]。また、ウェルビーイングは仕事の生産性にも大きく影響するとされており、高い個人のグループは、低い個人のグループよりも仕事のパフォーマンスが高くなる傾向にある [4][5]。

そこで本研究では、知的労働に従事しているオフィスワーカーのアンケートデータや行動データを分析し、行動や業務の心理状況への影響に関する知見の発見に取り組む。オフィスワーカーの心理と行動について分析した知見を基に業務環境や労働内容を改善することで、上記で述べたように社員のパフォーマンスや離職率、人材不足の解決に貢献することが期待できる。

本稿では、オフィスワーカーのメンタルヘルス向上を目的として、データ分析を行う。分析方法は 2 つある。1 つ目は、近い将来のオフィスワーカーの心理状況の予測である。今回収集したデータを用いて、6 つの心理指標を予測し、検証した。結果、72~89% の精度で分類予測をすることが

¹ 九州工業大学

² NTT データ経営研究所

³ 九州大学

できた。2つ目は、行動データと心理データとの相関を分析するため、予測における特徴量重要度、SHAP Value の分析を行う。結果、業務改善に貢献する知見を得ることができた。

その結果、ウェアラブルセンサ、行動、天気データを用いてオフィスワーカーの複数の心理指標を80%以上の精度で予測することができ、オフィスワーカーの心理指標に影響を与える業務における因子を発見することができた。

2. 関連研究

近年、センサデータや人の行動に着目したデータを用いて、ストレスや気分、メンタルヘルスの指標を推定する研究は、盛んに行われてきている。例えば、Koldijkらは、コンピュータログ、webカメラやkinectを用いた顔の表情や姿勢データ、心電図などの生体データを用いて、90%の精度でストレスを検出している[6]。Alberdiらは、スマートオフィス環境で現実のストレス要因の下で働くオフィスワーカーの生理データや行動データからストレスや仕事量を予測している[7]。

ウェアラブルセンサやスマートフォンを使用したストレス予測の取り組みも関心を集めており、Akane Sanoらは、スマートフォン、ウェアラブルセンサ、アンケートデータを用いて学生の成績、ストレス、精神状態を67~92%の精度で予測分類している[8]。Pabloらは、スマートフォン、ウェアラブルセンサ、ソーシャルアプリのログデータをマイニングし、ストレスを予測可能なフレームワークを開発した[9]。雨森らの研究では、リストバンド型端末によるセンサデータとスマートフォンによる位置情報データによってHRQOLに関するアンケート項目を推定している[10]。

労働者の心理分析に関する研究も多く、Matthew Leeらは知的労働者の生理センサデータによって仕事に対する集中状態を予測している[11]。福田らは、リストバンド型ウェアラブルセンサを用いて60人のオフィスワーカーを対象に、睡眠データを収集しアンケート調査による気分指標を推定している。分類精度は60~73%であり、睡眠状態における心理予測の貢献度を示している[12]。スマートフォンを用いた労働者のストレスを予測する研究では、Garcia-Cejaらが加速度センサを用いて自己申告されたストレスレベルを分類したところユーザ固有のモデルにおいて71%の精度を達成している[13]。近年では心理指標の将来予測の研究もされており、Suharaらの研究では毎日の気分に関する自己申告の履歴によって、抑うつ気分の予測精度を向上させた[14]。Taylorらウェアラブルセンサ、スマートフォンのログ、天気から収集したデータを用いて次の日の将来の気分、ストレス、健康を予測している。提案手法は参加者を性格や性別でクラスタリングし、クラスタに対してマルチタスクラーニングを行うことで気分予測の性能を向上させている[15]。

先行研究では、生体、画像、加速度などのセンサデータによってストレス認識の向上を目指していることが多いが、人の行動の種類を説明変数として用いている研究はほとんどない。また、気分やストレスについての推定をしている研究は多いが、オフィスワーカーのパフォーマンスに重要な疲労度、生産性、ワークエンゲージメント、業務に関する自己評価などと同じ説明変数によって推定している研究は少ない。オフィスワーカーを想定した研究も行われているが、企業に勤める人を被験者として実験している例は少なく、実際の仕事におけるデータを集めることが困難だった。また、実際に働いているオフィスワーカーを対象としてストレス認識を行った研究では職場でのストレスの分析にとどまっており、行動時の環境に関する分析をしている研究はほとんどない

上記を踏まえ、本稿では、オフィスワーカーの心理状況予測を行い、心理指標と業務内容、業務環境、行動との関係性を明らかにすることを目的として分析を行う。心理状況の予測では、スマートフォンやウェアラブル端末、天気などのデータを用いて6つの心理指標の将来予測をすることで、オフィスワーカーの労働衛生の改善に向けたツールの開発に貢献できると考えられる。また、心理指標と行動などの関係性を明らかにすることで働き方を改善するための知見を発見することが期待できる。

3. データセット

本研究で用いたデータセットの種類や収集方法について説明する。

3.1 データ収集

約100名の知識労働を行う会社員から14日間に渡り、のべ1400日分のデータを収集した。参加者は、男性63名、女性37名、平均年齢42.1歳である。本実験はNTTデータ経営研究所が主導して行っており、プロジェクトに協力している複数の企業にデータ収集をしていただいた。本実験で収集したデータは主に、リストバンド型端末によるセンサデータ、時系列行動ラベルデータ、アンケートによる心理指標データである。

リストバンド型端末を用いたセンサデータは、谷らが企業のオフィスワーカーの心身と環境の分析のためのセンシングプラットフォーム[16]を使用した。本研究のセンシングデバイスにも、同様にFitbitを使用している。Fitbitデータには、消費カロリー、心拍数、睡眠特性、代謝当量、歩数、階数、活動特性のデータが含まれている。Fitbitではそれぞれのデータを1分おきに測定している。

行動データ、アンケートデータの収集には、井上らが開発した介護行動記録アプリケーションfonlog[17]を、オフィスワーカー向けに設定して使用した。行動データについては、被験者がどの時間にどのような種類の行動をしてい

るのかをラベルづけすることで収集した。行動ラベルの種類は以下のようになっている。対面会議、食事、単独作業、趣味・休憩、家事・育児、休憩（業務時間内）、移動、web会議、共同作業（コミュニケーションを伴う）、電話（会議）、業務外作業特に業務中の行動には種類だけでなく、そのタスクの詳細やタスクを行うときの環境についても被験者の入力によって記録されている。業務に関するデータについて表1に示す。

心理状況に関するアンケートでは、6種類の指標によって評価を行っている。

- DAMS[18]

DAMSは肯定的気分、抑うつ気分、不安気分に関してそれぞれの強さを7段階で評価する質問。肯定的気分は、「はつらつとした」、「嬉しい」、「楽しい」の度合いの平均値。抑うつ気分は、「暗い」、「嫌な」、「沈んだ」の度合いの平均値。不安気分は、「気がかりな」、「不安な」、「心配な」の度合いの平均値、(9問)

- 自覚症しらべ [19]

自覚症しらべは個人が自覚している疲労の種類とその度合いを評価する質問。本研究ではねむけ感、不快感、だるけ感について各5問の質問があり、それぞれの項目について5段階で評価し、それらの平均値をスコアとしている。(15問)

- リカバリー経験 [20]

リカバリー経験は、余暇の時間に個人がどのようにリラックスして仕事から回復するかを評価する質問。「仕事との心理的距離」、「リラックスできたか」、「新しいことを学んだか」、「何をするか自分で決められたか」について各1問ずつ、7段階評価の質問の平均値をスコアとしている。(4問)

- ワークエンゲージメント [21]

ワークエンゲージメントは個人がどれだけ仕事に熱中できているかを評価する質問。仕事に対する「活力」、「熱意」、「没頭」に関する質問が各1問ずつ、7段階の平均値をワークエンゲージメントのスコアとしている。(3問)

- 生産性 [22]

世界保健機関が定めた健康と労働パフォーマンスに関する質問。1日の総合的なパフォーマンスの自己評価をスコアとしている。(1問)

- 業務内容の評価

参加者がその日の業務を自己評価する質問。以下の質問について5段階評価をスコアとしている。(7問) (1)「業務に集中できた」、(2)「効率的に業務ができた」、(3)「予定通りに業務ができた」、(4)「関係者と十分にコミュニケーションがとれた」、(5)「関係者と効率的にコミュニケーションがとれた」、(6)「新たなアイデアが出た」、(7)「成果に結びついた」

表1 業務内容及び業務環境に関するレコードの種類と含まれる値
アクティビティ レコード レコードの値

アクティビティ	レコード	レコードの値
業務内容	業務内容	企画業務 開発業務 営業業務 管理業務 現場業務
	業務の位置づけ	コア業務 ノンコア業務
	業務の新規性	定型業務 非定型業務
	業務における立場	管理者 作業員/参加者 共同作業員 管理者兼作業員
業務環境	業務環境	家(書斎など) 家(リビングなど) 家(その他) 職場外ワークスペース 店/外 職場
	業務状況	一人 他者あり(無干渉) 他者あり(同僚など) 他者あり(家族など)
	業務環境評価	非常に快適 快適 どちらでもない 不快 非常に不快

4. 提案手法

4.1 センサデータの特徴量抽出

リストバンド型端末では、消費カロリー、睡眠特性、代謝当量、歩数、階数、活動特性を1分おきに記録し、心拍数を10秒を機に記録している。それらの全てのデータを参加者ごとに一日で集計している。各センサの値に対する1日における、平均値、分散値、中央値を特徴量として使用した。しかし、睡眠に関しては時間帯が限定されているため、0時から8時までのデータを集計した。睡眠データでは睡眠状態を3段階で表しているため、睡眠レベルごとの睡眠時間を集計した。

4.2 行動データの特徴量抽出

行動データは、3.1で説明した行動ラベルの種類、業務内容、業務環境の各項目を使用した。それらのデータは参加者ごとの一日で集計し、値はその行動の継続時間を分単位で表している。また、特徴量となっている行動を行っていない場合は、0の値が入力されている。

4.3 アンケートデータの事前処理

アンケートによって集計した各心理指標は、それぞれのアルゴリズムによって連続値のスコアになっている。今回は、各心理的傾向が高いか低いかの二値分類を行うため、データの下位 40%と上位 40%の 2 クラスとする。中間スコアは心理的傾向が強いか弱いか判断できないため、中間の 20%のデータは削除する。このラベリング手法は [15] を参考している。クラスラベリングを朝アンケートにおける肯定的気分のスコアに対して行ったときを図 1 に示す。

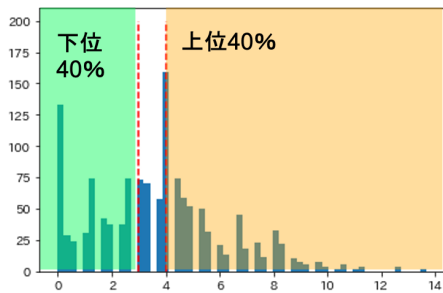


図 1 朝アンケートにおける肯定的気分のヒストグラム。x 軸：肯定的気分のスコア，y 軸：データ数

4.4 モデル構築

分類モデルには LightGBM を使用した。この機械学習モデルは Gradient Boosting Decision Tree (GBDT) の 1 つであり、従来の GBDT モデルに対して、Gradient-based One-Side Sampling (GOSS) と Exclusive Feature Bundling (EFB) を使用したモデルとなっている [23]。これによって、数が少ないデータに対して、正確な推定が可能となり、スパースなデータに対して有効なアルゴリズムとなっている。各分類モデルのパラメータはグリッドサーチを用いて、設定した。

モデルの評価には、k-fold 交差検証を行った。学習用のデータセットを 5 分割し、そのうちの 4 つのデータセットで学習し、1 つのデータセットで評価して、5 回繰り返した平均値を分類器の精度として評価する。

5. 評価実験

以下の 2 点を目的として分析する。

- (1) 今回の実験で収集したデータを用いて、各心理指標の予測精度を把握する。
- (2) 参加者の行動、業務内容や業務環境がどのように心理に影響しているのかを確認する。

2 つの評価実験を行う。1 つ目は、各心理指標の分類予測を行い、結果を示す。2 つ目は 3.1 で記述した DAMS アンケートの肯定的気分、抑うつ気分、不安気分の分類モデルを構築することによって、それぞれの目的変数と行動、業務内容との関係性を明らかにする。

5.1 分析 1：心理指標の予測結果

本節では、4.3 節で示した方法で 2 値クラスにした各アンケートデータに対して、分類予測をした結果を示す。説明変数には、3.1 節で述べたセンサデータと行動データの他に、参加者の個人データと天気データ、曜日データを追加して予測を行った。天気情報は人の心理に直接的な影響を与えるため重要であるだけでなく、参加者がいつ運動や勉強を行うのかに影響するため、心理予測に対して大きく貢献すると考えられる [15]。個人データには、参加者ごとの年齢、性別が含まれている。また、天気データは参加者の住居地域に基づいて、気象庁の web サイトからデータを収集した。天気データは 1 日ごとに集計された値であり、特徴量は以下に示す。平均気温、最高気温、最低気温、降水量、日照時間、平均風速、最高風速、平均気圧、雲量。

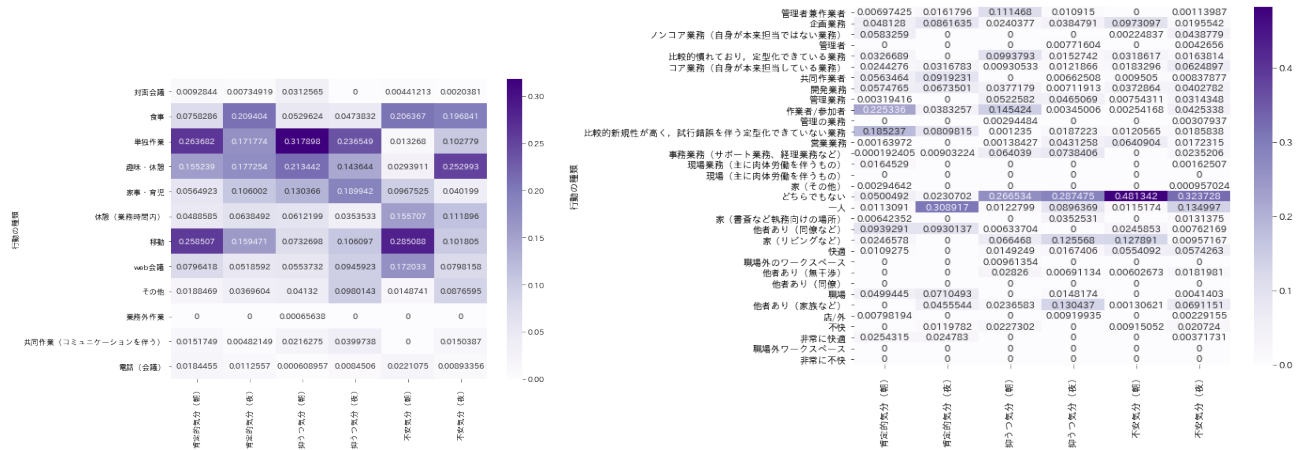
朝のアンケートと平日夜のアンケートを目的変数として分類予測をした。朝のアンケート予測を行うとき、説明変数には前日に集計したデータを用い、睡眠データのみ昨日の晩から当日の朝にかけて集計したデータを用いた。また、夜のアンケート予測を行うとき、説明変数は当日に集計したデータを用いており、睡眠データは、昨晚から当日朝にかけて集計したデータを用いた。以上のような条件で予測することによって、リアルタイムな心理状況ではなく、少し先の将来の心理状況を予測できるかを評価している。分類器の性能評価には Accuracy と F1 Score を用いた。Accuracy とは、テストを行った全ての回数のうち、正解と同じように分類した割合である。F1 Score とは Precision と Recall をバランスよく持ち合わせているかを示しており、それらの調和平均で求められる。今回、全ての 75 個の特徴量（行動：40 個、個人：2 個、天気：9 個、センサ：24 個）を目的変数とした予測を示し、その結果を基に特徴量選択をして目的変数 30 個で予測した結果も示す。特徴量選択をして予測する利点は以下の 2 点である。

- (1) ノイズになってしまう特徴量を削除でき、予測精度の向上が期待できる。
- (2) 学習に必要なメモリ・時間を削減する。

以上の条件で、評価した朝のアンケートの予測結果を表 2 に、平日夜のアンケートの予測結果を表 3 に示す。

表 2 全データを説明変数とした、朝のアンケート予測結果。左側が全データを説明変数とした予測精度、右側が特徴量選択により 30 個のデータを説明変数とした予測精度

心理指標	目的変数	accuracy	F1 score
DAMS	肯定的気分	82.5% / 83.9%	86.2% / 87.2%
	抑うつ気分	87.1% / 88.1%	86.2% / 87.2%
	不安気分	86.0% / 89.0%	85.7% / 88.7%
自覚症しらべ	ねむけ感	84.9% / 86.3%	84.9% / 86.2%
	不快感	87.1% / 88.4%	86.4% / 87.8%
	だるさ感	84.6% / 85.3%	82.4% / 83.2%



(a) 業務の種類を説明変数としたときの変数重要度

(b) 業務内容の詳細情報を説明変数としたときの変数重要度

図 2 業務の詳細，環境データを説明変数として，肯定的気分，抑うつ気分，不安気分を予測分類したときの Summary Plot

表 3 全データを説明変数とした，夜のアンケート予測結果. 左側が全データを説明変数とした予測精度，右側が特徴量選択により 30 個のデータを説明変数とした予測精度

心理指標	目的変数	accuracy	F1 score
DAMS	肯定的気分	72.1% / 72.1%	83.8% / 83.8%
	抑うつ気分	85.9% / 86.8%	87.8% / 88.4%
	不安気分	83.8% / 85.0%	88.4% / 89.1%
自覚症しらべ	ねむけ感	76.8% / 77.2%	72.4% / 72.1%
	不快感	83.2% / 82.7%	76.1% / 74.7%
	だるさ感	81.4% / 82.9%	75.6% / 78.0%
生産性	平均スコア	80.3% / 80.3%	79.2% / 79.2%
リカバリー	平均スコア	80.4% / 79.7%	80.4% / 79.6%
WE	平均スコア	79.5% / 79.4%	84.1% / 84.6%
業務の自己評価	(1)	81.3% / 81.1%	75.6% / 74.5%
	(2)	82.6% / 83.3%	86.2% / 87.1%
	(3)	80.2% / 81.5%	82.5% / 83.6%
	(4)	82.3% / 81.8%	85.4% / 85.2%
	(5)	80.2% / 81.8%	82.0% / 85.2%
	(6)	80.3% / 80.2%	86.3% / 86.5%
	(7)	80.2% / 79.7%	86.5% / 81.4%

5.2 分析 2：心理指標と行動との関係性分析

本研究では，各心理指標とオフィスワークの行動との関係性を明らかにすることを目的としており，本節では，行動データを説明変数として参加者の気分を予測し，特徴量重要度や相関からデータ間関係性を示す．今回，特徴量重要度はジニ重要度を使用した．ジニ重要度とは，決定木モデルで使われる特徴量重要度であり，ノードの不純度の減少分の重み付き和を決定木全体で平均した値となっている．つまり，その特徴量を用いたことで不純度をどれだけ改善できたかを表す指標である，特徴量重要度は各予測モデルごとに正規化して示している．

今回，朝と平日夜の DAMS（気分）を目的変数として予測したときの特徴量重要度を示す．DAMS の指標は「肯定

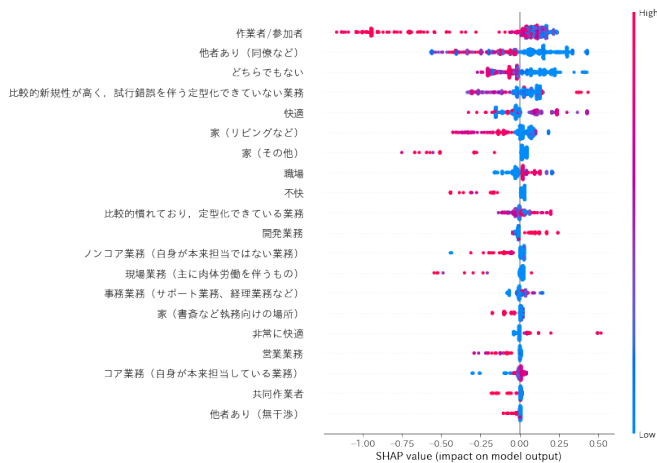
的気分」，「抑うつ気分」，「不安気分」であるため，同じ説明変数に対して 6 度の予測をした．図 2 では，横軸が目的変数，縦軸が説明変数を示し，ヒートマップで特徴量重要度の大きさを表している．また，(a) は参加者の行動の種類を説明変数とし，(b) では業務内容と業務環境を説明変数として用いている．

次に，各説明変数の予測に対する貢献度や相関を理解するために，Summary Plot を用いて特徴量ごとに SHAP Value を一軸の散布図として示す．SHAP Value とは協力ゲーム理論によって求められる各特徴量の貢献度である．ある特徴量を追加した場合の予測精度の増加分を限界貢献度と呼び，決定木による分岐の順番に依存している．そのため，全ての順番に対する限界貢献度の平均をとった値を SHAP Value と呼ぶ [24]．図 3 では，参加者の行動の種類を説明変数としたときにおける，SHAP Value の散布図を示す．

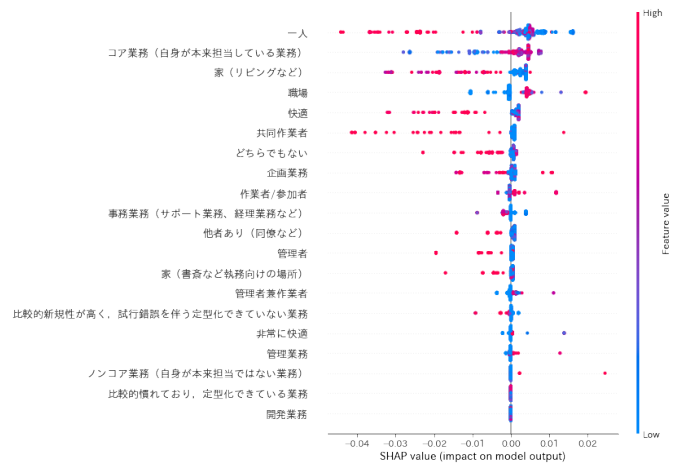
6. 考察

6.1 分析 1 についての考察

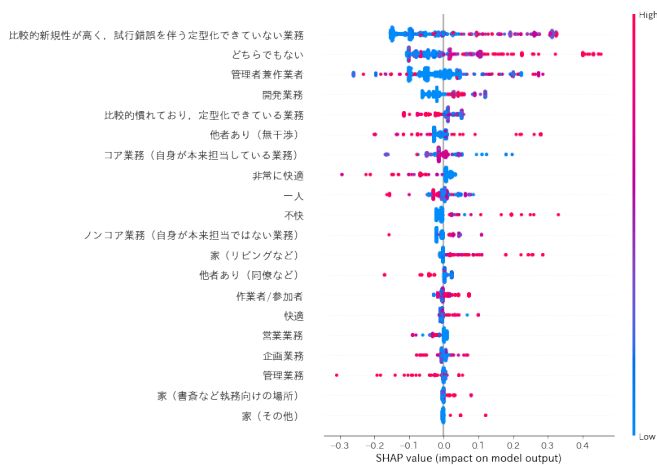
全心理指標の予測について，予測精度は 72~89% であり全体的に安定して高い精度となった．自覚症しらべについては，朝のアンケートの予測結果より夜のアンケートの予測結果が全体的に低い結果となっている．夜の自覚症しらべの予測における特徴量重要度には，業務環境や休憩時間が強く影響している．そのため，行動の入力データが少ないユーザに関しては，予測精度が低下したのではないかと考えられる．また，全ての心理指標の予測において，行動，天気，センサーデータがそれぞれ特徴量重要度が高く，予測に貢献していることがわかった．今回は，個人を考慮せずに機械学習モデルを構築したが，個人ごとに毎日の業務や心理状況に偏りがあるため，ユーザ個人を考慮したモデルにした場合さらに予測精度の向上が期待できるのではない



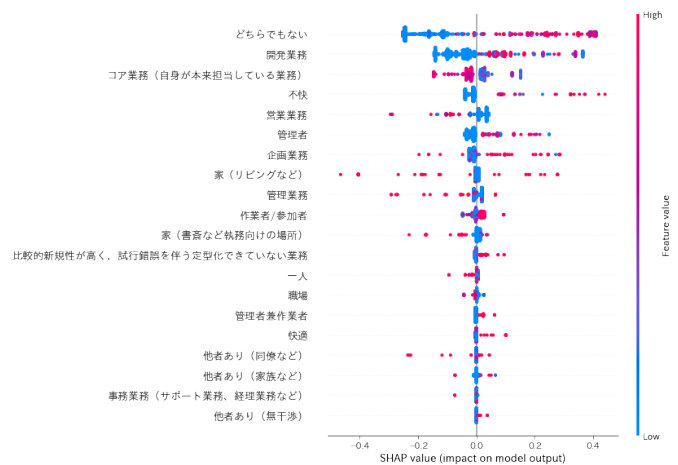
(a) 朝アンケートの肯定的気分を予測したとき



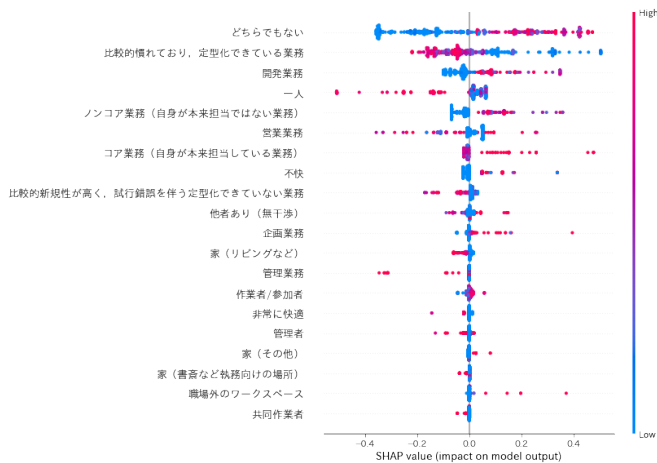
(b) 夜アンケートの肯定的気分を予測したとき



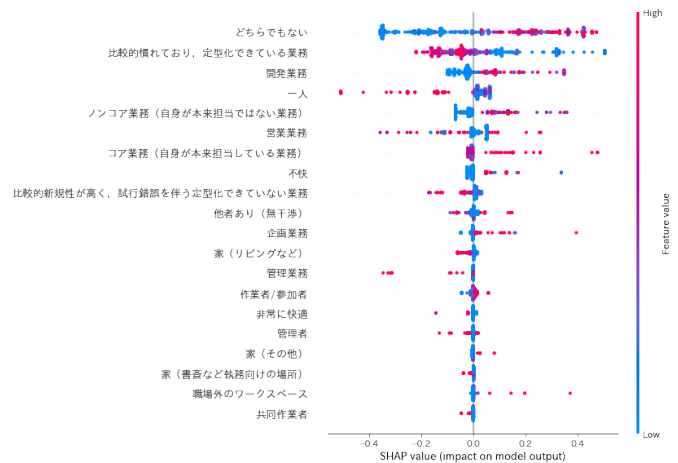
(c) 朝アンケートの抑うつ気分を予測したとき



(d) 夜アンケートの抑うつ気分を予測したとき



(e) 朝アンケートの不安気分を予測したとき



(f) 夜アンケートの不安気分を予測したとき

図3 業務の詳細、環境データを説明変数として、肯定的気分、抑うつ気分、不安気分を予測分類したときの Summary Plot

かと考えられる。

6.2 分析2 についての考察

図2に示した特徴量重要度と、図3から各心理指標と説明変数との関係性の強さや相関を理解することができる。

業務の種類とのDAMSとの関係性に関して、肯定的気分と抑うつ気分は朝と夜で同じ特徴量が予測に貢献しており、肯定的気分は「単独作業」、「趣味・休憩」、「移動」、抑うつ気分は「単独作業」、「趣味・休憩」、「家事・育児」が高い傾向にある。図3の(c),(d),(e),(f)において、業務内容の

中で業務環境評価の「どちらでもない」の特徴量の貢献度が高く、いずれも正の相関を示しており、業務環境に関して「どちらでもない」と評価している参加者は、抑うつ、不安傾向が高いことがわかる。また、図3の(c),(e),(f)より「比較的慣れており、定型化できている業務」（定型業務）と朝の抑うつ気分、朝と夜の不安気分を負の相関が見られる。このことから、業務が定型化できているかどうかは社員の不安や抑うつの心理に強い影響を与えており、定型化されているほど心理的負荷が少ないことがわかる。

在宅ワークにおける業務環境に関して、図3の(a),(c)より、「家（リビングなど）」と夜の肯定的気分との間で負の相関があり、朝の抑うつ気分と正の相関がある。つまり、在宅の業務環境も心理的な影響が強く、特にリビングのような場所での業務は、うつ傾向になりやすくなる。または、うつ傾向にある参加者が「家（リビングなど）」での勤務を選択しやすい可能性が考えられる。

行動の種類についても Summary Plot を用いて相関分析を行ったところ、「web 会議」が肯定的気分において負の相関、抑うつ気分において正の相関があることがわかった。つまり、web 会議の時間が長ければオフィスワークの心理に悪影響を及ぼす可能性が高い。「休憩時間（業務時間内）」と各気分との間でも特徴的な相関関係が見られ、朝の肯定的気分、朝と夜の不安気分において負の相関があり、朝の抑うつ気分と正の相関となった。

7. 結論

本論文では、行動データ、センサデータ、天気データにより機械学習予測モデルを構築することで、オフィスワークの6つの心理指標において72~89%の精度で分類予測をすることができた。本論文での実験及び考察において得られた知見を以下に示す。

- 今回の実験で収集した全データを用いて心理指標の二値分類を行った場合、22項目の心理指標の予測において17項目のF1 scoreが80%以上の精度となり、先行研究[6][10][15]と比べても高い水準の精度を得ることができた。
- 朝のアンケートと夜のアンケートの予測精度を比べると、夜の自覚症しらの精度が著しく低下していることがわかった。

今後は以下について調査する。

- 機械学習モデルの個人ごとの学習、またはユーザのクラスタリングを行って同じクラスタ内での学習により予測精度の向上するかどうかの検証。
- 今回は、行動データについて、行動の継続時間のみを特徴量として用いたため、行動の時刻や頻度の特徴量への追加。
- 他の機械学習モデルやアンサンブル学習による予測精度向上の検討。

8. 謝辞

本研究の実験は、一般社団法人応用脳科学コンソーシアムが事務局を務める、「2020年度センシング&トランスフォーメーション研究会」に参画している企業・大学と連携して実施したものである。

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. https://www.mhlw.go.jp/toukei/list/dl/h30-46-50_gaikyo.pdf, 2018.
- [2] Margaret Coffey, Lindsey Dugdill, and Andy Tattersall. Stress in Social Services: Mental Wellbeing, Constraints and Job Satisfaction. *The British Journal of Social Work*, Vol. 34, No. 5, pp. 735-746, 07 2004.
- [3] Maria S Kopp, Adrienne Stauder, György Purebl, Imre Janszky, and Árpád Skrabski. Work stress and mental health in a changing society. *European Journal of Public Health*, Vol. 18, No. 3, pp. 238-244, 08 2007.
- [4] Peter Warr and Karina Nielsen. *Wellbeing and work performance*. 02 2018.
- [5] S. Sharief. Job stress and its impact on employees' performance a study with reference to employees working in bpos m.kotteeswari research scholar, the new collage, chennai. 2014.
- [6] S. Koldijk, M. A. Neerincx, and W. Kraaij. Detecting work stress in offices by combining unobtrusive sensors. *IEEE Transactions on Affective Computing*, Vol. 9, No. 02, pp. 227-239, apr 2018.
- [7] Ane Alberdia, Asier Aztiriaa, Adrian Basarabb, and Diane J. Cook. Using smart offices to predict occupational stress.
- [8] Akane Sano, A. Phillips, A. Z. Yu, A. W. McHill, Sara Taylor, Natasha Jaques, C. Czeisler, E. Klerman, and Rosalind W. Picard. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 1-6, 2015.
- [9] Pablo Robles-Granda, Suwen Lin, Xian Wu, Sidney D'Mello, Gonzalo J. Martínez, Koustuv Saha, Kari Nies, Gloria Mark, Andrew T. Campbell, Munmun De Choudhury, Anind D. Dey, Julie M. Gregg, Ted Grover, Stephen M. Mattingly, Shayan Mirjafari, Edward Moskal, Aaron Striegel, and Nitesh V. Chawla. Jointly predicting job performance, personality, cognitive ability, affect, and well-being. *CoRR*, Vol. abs/2006.08364, , 2020.
- [10] Amir M. Rahmani, Nima TaheriNejad, and Paolo Perego. A method for simplified hrqol measurement by smart devices. *Wireless Mobile Communication and Healthcare*, pp. 91-98, 2017.
- [11] Matthew Lee. Detecting affective flow states of knowledge workers using physiological sensors. *CoRR*, Vol. abs/2006.10635, , 2020.
- [12] 福田修之, 谷優里, 松田裕貴, 荒川豊, 安本慶一. オフィスワークの睡眠状況と労働衛生指標の関係性分析. Technical Report 22, 奈良先端科学技術大学院大学, 九州大学/JST さきがけ, nov 2019.
- [13] Enrique Garcia-Ceja, Venet Osmani, and Oscar Mayora. Automatic stress detection in working environments

- from smartphones' accelerometer data: A first step. *IEEE Journal of Biomedical and Health Informatics*, Vol. 20, No. 4, p. 1053–1060, Jul 2016.
- [14] Yoshihiko Suhara, Yinzhan Xu, and Alex 'Sandy' Pentland. Deepmood: Forecasting depressed mood based on self-reported histories via recurrent neural networks. In *Proceedings of the 26th International Conference on World Wide Web, WWW '17*, p. 715–724, Republic and Canton of Geneva, CHE, 2017. International World Wide Web Conferences Steering Committee.
- [15] Sara Taylor, Natasha Jaques, Ehimwenma Nosakhare, Akane Sano, and Rosalind Picard. Personalized multi-task learning for predicting tomorrow's mood, stress, and health. *IEEE Transactions on Affective Computing*, Vol. 11, No. 2, pp. 200–213, 2020.
- [16] Yuki Matsuda Sozo Inoue Yuri Tani, Shuichi Fukuda and Yutaka Arakawa. Workersense: Mobile sensing platform for collecting physiological, mental, and environmental state of office workers. In *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 2020.
- [17] Sozo Inoue, Paula Lago, Tahera Hossain, Tittaya Mairittha, and Nattaya Mairittha. Integrating activity recognition and nursing care records: The system, deployment, and a verification study. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, Vol. 3, No. 3, 2019.
- [18] L. Clark. The anxiety and depressive disorders: Descriptive psychopathology and differential diagnosis. 1989.
- [19] 久保智英, 城憲秀, 武山英磨, 榎原毅, 井上辰樹, 高西敏正, 荒薦優子, 村崎元五, 井谷徹. 「自覚症しらべ」による連続夜勤時の疲労感の表出パターンの検討. *産業衛生学雑誌 = Journal of occupational health*, Vol. 50, No. 5, pp. 133–144, sep 2008.
- [20] Akihito Shimazu, Sabine Sonnentag, Kazumi Kubota, and Norito Kawakami. Validation of the japanese version of the recovery experience questionnaire. *Journal of occupational health*, Vol. 54, , 03 2012.
- [21] Maria S Kopp, Adrienne Stauder, György Purebl, Imre Janszky, and Árpád Skrabski. Work stress and mental health in a changing society. *European Journal of Public Health*, Vol. 18, No. 3, pp. 238–244, 08 2007.
- [22] Catherine MPA; Beck Arne PhD; Berglund Patricia M.BA; Cleary Paul D. PhD; McKenas David MD; Pronk Nico PhD; Simon Gregory MD; Stang Paul PhD; Ustun T. Bedirhan MD; Wang Phillip MD ScD Kessler, Ronald C. PhD; Barber. The world health organization health and work performance questionnaire (hpq). *Journal of Occupational and Environmental Medicine*, Vol. 45, pp. p 156–174, 02 2003.
- [23] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, Vol. 30. Curran Associates, Inc., 2017.
- [24] Scott Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *CoRR*, Vol. abs/1705.07874, , 2017.