

勉強中における心拍数を用いた精神状態度合の推定

川崎 勇佑¹ ロペズギヨーム¹ 横窪 安奈¹

概要：本研究では心拍センサ搭載のスマートウォッチを用いて、勉強中の精神状態を連続的に推定し、短期的なアクチュエーションおよび、長期的なフィードバックにより、パフォーマンス向上を支援するシステムの開発を目的としている。スマートウォッチの心拍センサから心拍数関連の特徴量を用いて、集中、眠気、疲労の度合いが十分に検出可能かどうかを検証した。結果、それぞれの精神状態度合いの平均が78%以上であることから、心拍数のみでもある程度の精神状態度合いを推定することができることが示唆された。

Estimating the Degree of Mental State using Heart Rate while Studying

YUSUKE KAWASAKI¹ GUILLAUME LOPEZ¹ ANNA YOKOKUBO¹

1. 現代の勉強における課題

日常生活において、勉強や仕事、スポーツをするときなど集中が必要になる場面が様々ある。特に勉強というのはどの年齢になったとしても必ずすることであり、非常に重要な要素の一つである。勉強は学問や技芸などを学ぶことであり、勉強の効果を高めるためには高い集中が必要になってくる。例えば同じ時間勉強をしたとして、集中が高い方が集中が低い時よりも効率的に勉強ができる。しかし、高い集中が重要であるにもかかわらず、集中の高さを定量的に評価する手法は日常生活には普及していない。

また、現代では勉強の質よりも勉強時間に重きを置いている。文部科学省の第18回21世紀出生児縦断調査の結果を見ると、大学または大学院に進学しようとしている高校3年生相当の学生は、休日における勉強時間が6時間以上の割合が多い(大学進学希望者は26.7%、大学院進学希望者は49.5%) [1]。また、調査票を見てみると勉強時間に対する質問しかない。このようなことから現代では勉強時間を評価の一つにしているのではないかと考えられる。しかし、一度も休憩を挟まない長時間学習に比べ、一定時間ごとに短い休憩を挟む積み上げ学習の方が作業効率が向上することが示されている [2]。

ベネッセコーポレーションの調査によると、15分×3

(計45分)の積み上げ学習と60分の長時間学習を比較してみると、積み上げ学習の方が勉強時間が少ないにもかかわらず長時間学習よりも良いスコアを得ていることが分かっている [3]。また、集中力に関与している前頭葉のガンマ波が40分以降に急激に低下しており、集中力は40分程度しか持続しない可能性を示唆していた。

本研究では、図1に示すように、ウェアラブル端末を用いて、勉強中のメンタル状態を連続的に推定し、短期的なアクチュエーションおよび長期的なフィードバックにより、パフォーマンス向上を支援するシステムを実現することを目的としている。

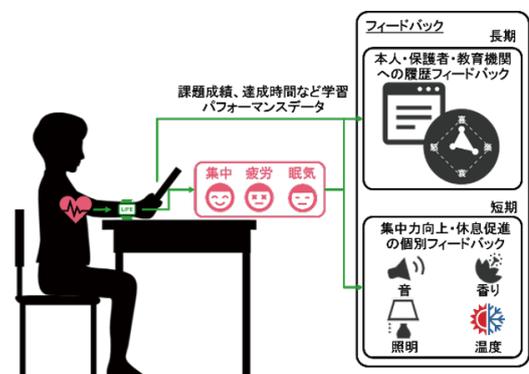


図1 心拍数を用いた勉強中のメンタル状態推定とフィードバックシステムの概念図

¹ 青山学院大学
Aoyama Gakuin University

2. 精神状態推定に関する先行研究

2.1 加速度センサを用いた集中状態の推定

大久保ら [4] は椅子に加速度センサを付けて、集中状態を評価するシステムを提案した。実験は被験者 3 人に 30 分間タイピング作業をさせ、その間の加速度のパワースペクトルの和と単位時間当たりのキーボードの打鍵数の関係について検討している。結果は、パワースペクトルの和と単位時間当たりの打鍵数には強い負の相関が存在することが分かり、パワースペクトルの和が小さくなると集中していると推定することの妥当性が明らかとなった。しかし、実験で使用した椅子は、回転椅子のみであり限定的な環境でしか実験されていない。また、ユーザの利用する椅子にセンサを正しく装着するのは、ユーザ本人のみでは難しいという問題点があった。

植田ら [5] は、胸ポケットにあるスマートフォンから加速情報を取得し、集中度を評価した。また、加速度データを取得すると同時に、ビデオカメラで被験者の様子を記録することによって、被験者がどのような行動をとっていたか確認している。結果は適合率が 74.31% となっており、概ね良好であると結論付けていた。しかし、使用したデータの長さが 144 分しかなく、データ数がとても少ないことや評価指標が適合率しかないといった問題点がある。

2.2 脳波センサを用いた精神状態の推定

脳波の波形は精神状態の指標として用いることができるということは、様々な心理学や脳科学の研究結果からわかっている [6][7]。人間の精神状態を推定するには、得られた脳波に対して離散フーリエ変換することで得られた α 波や β 波のパワースペクトルや、 α 波や β 波の脳は全体に対する割合、 α 波と β 波の比率を測ることが有効とされている [8][9]。一般的に α 波は安静時・覚醒時ともに見られる波形であり、リラックス状態時には α 波の振幅が大きくなる。反対に緊張時には α 波の振幅が小さくなり β 波の出現がみられるなどの特徴がある。

平井ら [10] は、簡易脳波計を用いた学習状態中における思考、および記憶の相関分析のため、 α 波・ β 波・ low_y 波のパワースペクトルおよび β/α を実験により比較分析した。結果 β/α 成分を観察することによってストレスの度合や、瞬間的に思考をするための集中度を測る指標となり得ることが示された。また low_y においては記憶しているかどうかの度合を測る判断基準として有効であると考えられると述べられていた。

しかし、デバイスの大きさおよびデバイスを本人のみで装着することの難しさ、電磁ノイズの影響に敏感なことから日常生活において利用することが困難であるといった問題点がある。

2.3 心電位センサを用いた精神疲労状態の推定

心拍変動とは、心臓の拍動を測定した心電位に表れる上向き幅の狭い波 (R 波) の間隔である RRI (R-R Interval) や、末梢血管の血流量を測定した脈波のピークの間隔である PPI (Pulse-to-Pulse Interval) の変動を意味する。RRI の変動は自律神経系との関連が示唆されており、個人差があるものの、RRI の変動が小さい時に自律神経系の交感神経が活性化しており、RRI の変動が大きい時は副交感神経が活性化しているとされる [11]。

角田ら [12] および、東海林ら [13]、Nkurikiyeyezu ら [14] のように、心拍変動の特徴量と精神疲労水準の間に相関があることを確認し、精神疲労における逐次推定は可能であることが多くの研究にて示唆されている。しかし、心拍変動を計測するには連続的に心拍間隔を高精度に計測する必要があり、日常生活で用いられるデバイスでは、未だに不可能である。

3. 心拍数を用いた精神状態判別の検証

3.1 心拍数と自律神経系活動

2.3 節で紹介したように、精神状態を反映している自律神経系の活動を定量評価するため、心拍変動解析が良く用いられている。多くのスマートウォッチは心拍センサを搭載していても、心拍間隔データを取得する仕組みは提供されていない。一方、自律神経系の活動を評価するには心拍数や脈拍数のみでは不十分とされているため、既存のスマートウォッチでは評価できないとされている。また、ほとんどの先行研究では集中、眠気、疲労いずれか一つの精神状態しか評価していない。しかし、勉強のパフォーマンスを決定するのは集中のみではなく、同時に発生している眠気と疲労もかかわっていると考える。そこで、今回の目標はスマートウォッチの脈拍センサから心拍数関連の特徴量を用いて、集中、眠気、疲労の度合が十分に検出可能か検証し、その有効性を示すことである。

心拍数のみを用いて、精神状態を判別するのは不十分とされており、ほとんど研究が行われていない。そこで、精神状態を判別可能であると示唆されてある、脈拍変動 (PRV) と心拍数を比較することで、心拍数のみでも精神状態を判別可能かどうかの検証を行った。PPI の周期的な変動の特徴は、心拍間隔 (RRI) のと完全に一致しないものの、情動や感情等の精神状態を表していることが報告されている [15] ことから、本研究ではより非侵襲な方法で手首から計測できる PPI の周期的な変動 (PRV) を自律神経系活動指標として用いる。

3.2 実験方法

被験者は健康な成人男性 2 名とした。図 22 のように左右どちらかの腕に心拍数を測定するために光電式容積脈波

センサ搭載スマートウォッチ（FitBit Versa2）を装着してもらい、もう片方の腕に心拍変動を測定するためのスマートウォッチ（Polar M600）を装着してもらった。前述の通り、通常のスマートウォッチでは心拍数しか取得できないが、Polar M600 では、Wear OS API の一拍毎心拍間隔データ取得メソッドの実装や、Polar 社独自のセンサ API により脈波の生データも取得可能なため、心拍変動の算出は可能になっている。手を激しく動かすと正しく測定ができない恐れがあるため、被験者には日常的に読んでいる勉強本を読んでもらった。集中、眠気、疲労のそれぞれの度を低・中・高の3段階に設定し、勉強開始時と勉強中にいずれかの精神状態ラベルの度が変わったときに図3のアプリに被験者が手動で記録した。以下の図4の手順に従い勉強中の心拍データと脈拍変動データと精神状態ラベルの度合データを取得した。



図2 装着したスマートウォッチ（左腕が FitBit Versa2, 右腕が Polar M600）



図3 作成した精神状態度合記録アプリケーションソフトウェア（Polar M600 版）

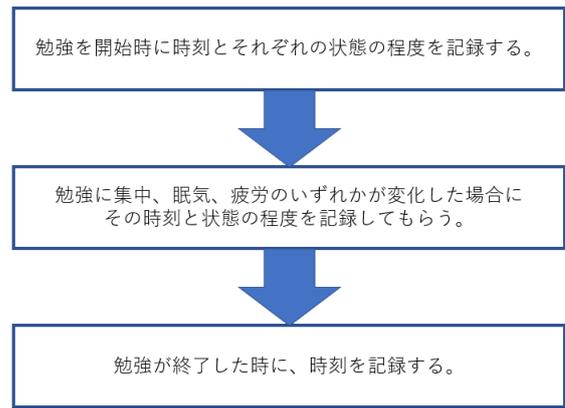


図4 実験手順

3.3 データセットの作成方法

計測した心拍データ（HR）と脈拍変動データ（PPI）はある程度の時間でまとめてその間の変化を特徴として、精神状態と関連付けることが必要になってくる。その最適な時間を特定するために3つの分け方を試した。第1に1分窓で10秒シフト、第2に5分窓で30秒シフト、第3には10分窓で60秒シフトである。

脈拍変動に関してはスマートウォッチという形上、少しの動作によってノイズが混入することがある。ノイズを放置したまま、分析を進めると期待通りの結果が得られだけでなく、想定と真逆の結果を得ることもある。このようなことから、脈拍変動データにはノイズ除去処理を行った。時間情報に基づくPPI異常値除去の手法は、下限値 250[ms] 以下、上限値 1500[ms] 以上のPPIをまず除外する。残ったPPIのうち、平均±3×標準偏差の範囲（3σ）に収まらないものを異常値としてさらに除外し[16]、残ったPPIのみを脈拍変動解析の対象とした。

心拍と脈拍変動の精神状態度合の教師ラベルの設定方法は特定の区間のデータセット内で一番多かったラベルをその区間の教師ラベルにした。精神状態度合ごとに収集した心拍と脈拍変動とラベルデータ数を表1にまとめた。表のPPIの1分窓の行にある空白は、解析窓が1分であると脈拍変動は周波数解析をすることができず、データを得ることができないからである。

表1 心拍と脈拍変動とラベルデータ数

状態	解析窓（分）	HR			PPI		
		低	中	高	低	中	高
集中	1	778	2046	453			
	5	238	635	141	255	892	137
	10	298	98	58	104	410	60
眠気	1	1179	1255	243			
	5	557	383	74	709	506	69
	10	254	168	32	319	225	30
疲労	1	1597	1579	101			
	5	505	480	29	650	607	27
	10	231	210	13	287	275	12

3.4 分析手法

Python (v.3.7.6) を使って分析をした。脈拍変動データから抽出した特徴量は、「LF, HF, LF/HF, 最大値, 最小値, 最大値-最小値, 平均値」の7つである。LF, HF, LF/HF とは脈拍変動を周波数解析して得ることができる指標である。低周波領域 (0.004~0.15[Hz]) パワー値 (LF) は交感神経と副交感神経の両方の活動を反映し、LF 値が高い場合副交感神経が増大しリラックスしている状況と言える。高周波領域 (0.15~0.40[Hz]) パワー値 (HF) は副交感神経の活動を反映している。LF/HF とは、LF と HF のパワー比率であり、交感神経と副交感神経の全体のバランスを表し、数値が高いと交感神経優位を、低い場合は副交感神経優位を示している [17]。周波数解析をする場合に1分間のデータでは解析することができない。脈拍変動の分析には5分窓で30秒シフトと10分窓で60秒シフトの2つのデータセットを使用した。心拍数から抽出した特徴量は「最大値, 最小値, 平均値, 標準偏差」の4つである。こちらは、1分窓でシフト10秒と5分窓で30秒シフトと10分窓で60秒シフトの3つのデータセットを使用した。

ランダムにデータセットの75%を学習用に利用した。Support Vector Machine (SVM), K近傍法 (KNN) の2つのモデルを心拍数データと脈拍変動データのそれぞれで作成し、データセットの残りの25%で精度評価した。そのプロセスを10回繰り返して、各精神状態度合のF1値 (%) の平均を算出した。

3.5 検証結果と考察

結果の表2を見てみると、それぞれの解析窓でHRとPPIそれぞれで精度を比較してみた時に1分窓のデータセットが著しく精度が低いのにに対し、10分窓のデータセットが最も良かった。また、HRとPPIで同じ解析窓で比較してみると、PPIの方が精度が良かった。しかし、10分窓のHRのデータセットの精度を見てみると、差が平均で8.5%しかなく、多くの精神状態度合では差が10%前後で収まっている。このことから、HRのみでもある程度の精神状態を判別することは可能だということが示唆された。しかし、眠気の高ラベルに関してはSVMもKNNも精度がとて低く、ラベル数を増やしたり、実施する分類器を増やしたりするなどして改善を進めていく。

4. 心拍数を用いた精神状態度合の推定

前述の実験により、心拍数のみを用いて、精神状態を分析が可能だということが示唆された。今回の実験は実際の勉強においてデータを収集し、より詳しく分析を進めた。

4.1 実験方法

実験参加者の心拍数を抽出するために左右いずれかの腕

表2 HRとPPIを使った分析結果

状態	解析窓 (分)	HR			PPI		
		F1 (%)			F1 (%)		
		低	中	高	低	中	高
集中 (SVM)	1	31	79	36			
	5	72	89	79	83	92	67
	10	74	90	80	90	95	81
眠気 (SVM)	1	70	61	5			
	5	86	81	46	87	84	56
	10	90	86	38	94	91	68
疲労 (SVM)	1	61	67	11			
	5	81	81	53	89	86	65
	10	83	83	85	93	92	87
集中 (KNN)	1	46	74	47			
	5	80	90	81	85	93	68
	10	81	90	73	92	96	83
眠気 (KNN)	1	72	61	41			
	5	89	83	68	87	84	59
	10	89	87	55	94	92	73
疲労 (KNN)	1	64	63	32			
	5	88	88	70	87	86	66
	10	88	87	75	93	92	89

に、光電式容積脈波センサ搭載スマートウォッチ (FitBit versa2) を装着してもらった。集中、眠気、疲労のそれぞれの度合を低・中・高の3段階に設定し、勉強開始時と勉強中にいずれかの精神状態ラベルの度合が変わったときに記録した。勉強に極力影響しないように、スマートウォッチの画面をタッチするだけで変更された精神状態のラベルと度合と時刻が記録できるようなアプリケーションソフトウェアを開発した。

塾に通っている高校生男女5名に協力してもらい、自宅学習の勉強中の連続心拍データと精神状態ラベルの度合データを取得した。前回の実験と同様 (図4) の手順に従い、被験者に通常通り勉強をしてもらった。

4.2 心拍と精神状態の度合のデータセットの作成方法

前回の実験から、1分窓シフト10秒のデータセットの精度が著しく低かった。そこで今回は、以下のデータセットで比較する。

- 3分窓シフト20秒のデータセット
- 5分窓シフト30秒のデータセット
- 10分窓シフト60秒のデータセット

精神状態度合ごとに収集した心拍とラベルデータ数を表3にまとめた。

4.3 分析手法と結果

2種類の分析手法を実施した。??に窓時間と学習モデルの評価、4.3.2に未知のユーザに対するロバスト性評価を述べる。

表 3 心拍とラベルデータ数

精神状態	解析窓 (分)	精神状態度合			
		低	中	高	全体
集中	3	1240	11568	5899	18707
	5	800	7455	3816	12071
	10	366	3399	1760	5525
眠気	3	12766	4626	1315	18707
	5	8221	2998	852	12071
	10	3737	1387	401	5525
疲労	3	9484	7966	1257	18707
	5	6109	5149	813	12071
	10	2776	2376	373	5525

4.3.1 窓時間と学習モデルの評価

分析をするために心拍数データから次の七つの特徴量を抽出した: 平均値, 最大値, 最小値, 最大値 - 最小値, 標準偏差, 尖度, 歪度. データセットからランダムに選択した 75% を学習用に利用した. Support Vector Machine (SVM), K 近傍法 (KNN), Random Forest (RF), Gradient Boosted Regression Trees (GBRT), Extra Trees Classifier (ET) の 5 つのモデルを学習, 残りの 25% のデータで精度評価した. いずれの分類器もパラメータは初期値である. そのプロセスを 10 回繰り返し, 各精神状態度合の F1 値 (%) と標準偏差を算出した.

表 44, 表 5, 表 6 は分析の結果である. 集中・眠気・疲労それぞれ一番高い精度は 10 分窓のデータセットで分類器が ET のモデルであり, いずれの精神状態を 78% 以上の平均精度で推定できた (表 7). また, どの分類器においても 10 分窓のデータセット, 5 分窓のデータセット, 3 分窓のデータセットの順に精度が落ちていっており, 心拍数を使用した精神状態度合推定は窓時間が広いほど精度が良くなっている. 分類器同士で比較してみると, GBRT はどの窓時間においても他の分類器と比べて平均の精度が低かった.

しかし, 中集中, 低眠気に関してはほかの分類器と同程度の精度があった. 中集中と低眠気に関してはそれぞれラベル数が一番多いものになっていた. 10 分窓のデータセットの SVM や KNN を使ったモデルに関しては, ET を使ったモデルには及ばないが 70% 前後の平均精度で推定できた.

4.3.2 未知のユーザに対するロバスト性評価

次に, 前述の心拍数特徴量と分類器の学習モデルを用いて, 未知のユーザデータに対する精度評価を行った. 前述 5 名のデータのうち, 1 名のデータを順番に除いてモデルの学習を行ったうえで, 除いたユーザのデータでテストを行った. 評価方法としてテストデータの正解ラベルとモデルが予測した予測ラベルの F1 値を導き出した.

表??, 表??, 表?? は分析の結果である. 明らかに 1 つ目の結果と比べて精度が落ちていることがわかる. 60% を

表 4 窓時間と学習モデルの評価の結果 (3 分窓のデータセット)

	低	中	高	平均
集中 (SVM)	25.5 ± 3	76.0 ± 0	47.0 ± 1	49.5
眠気 (SVM)	80.5 ± 1	37.7 ± 1	25.1 ± 3	47.8
疲労 (SVM)	69.6 ± 1	65.0 ± 1	27.6 ± 2	54.1
集中 (KNN)	36.9 ± 2	73.5 ± 0	51.7 ± 1	54.0
眠気 (KNN)	77.8 ± 0	47.6 ± 1	34.0 ± 2	53.1
疲労 (KNN)	69.6 ± 1	63.8 ± 1	38.3 ± 1	57.2
集中 (RF)	40.1 ± 2	79.5 ± 0	55.3 ± 1	58.3
眠気 (RF)	83.6 ± 0	49.2 ± 1	36.9 ± 2	56.6
疲労 (RF)	75.5 ± 1	70.5 ± 1	43.6 ± 2	63.2
集中 (GBRT)	11.7 ± 1	76.7 ± 0	21.3 ± 2	36.6
眠気 (GBRT)	81.2 ± 0	10.7 ± 1	9.4 ± 2	33.8
疲労 (GBRT)	66.4 ± 1	61.6 ± 1	13.6 ± 2	47.2
集中 (ET)	50.7 ± 2	80.6 ± 0	60.8 ± 1	64.0
眠気 (ET)	84.3 ± 0	54.8 ± 1	45.3 ± 2	61.5
疲労 (ET)	77.2 ± 0	71.1 ± 0	50.0 ± 2	66.1

表 5 窓時間と学習モデルの評価の結果 (5 分窓のデータセット)

	低	中	高	平均
集中 (SVM)	40.8 ± 2	79.2 ± 1	58.4 ± 1	59.5
眠気 (SVM)	82.9 ± 0	52.0 ± 1	33.5 ± 2	56.1
疲労 (SVM)	74.6 ± 1	70.9 ± 1	40.7 ± 3	62.1
集中 (KNN)	50.5 ± 1	78.6 ± 0	62.2 ± 1	63.8
眠気 (KNN)	81.7 ± 1	57.2 ± 1	46.3 ± 2	61.7
疲労 (KNN)	75.5 ± 1	70.3 ± 1	51.3 ± 3	65.7
集中 (RF)	53.7 ± 2	83.5 ± 0	65.9 ± 1	67.7
眠気 (RF)	86.3 ± 0	60.7 ± 1	49.1 ± 3	65.4
疲労 (RF)	79.6 ± 1	75.7 ± 1	55.8 ± 2	70.4
集中 (GBRT)	17.0 ± 3	77.5 ± 0	31.1 ± 2	41.9
眠気 (GBRT)	81.7 ± 0	18.6 ± 1	13.8 ± 3	38.0
疲労 (GBRT)	69.5 ± 1	63.2 ± 1	20.3 ± 5	51.0
集中 (ET)	61.2 ± 3	85.3 ± 1	71.6 ± 1	72.7
眠気 (ET)	87.8 ± 0	67.5 ± 1	59.2 ± 3	71.5
疲労 (ET)	81.8 ± 1	77.4 ± 0	63.7 ± 3	74.3

表 6 窓時間と学習モデルの評価の結果 (10 分窓のデータセット)

	低	中	高	平均
集中 (SVM)	55.2 ± 4	82.5 ± 1	68.1 ± 2	68.6
眠気 (SVM)	85.9 ± 1	63.6 ± 2	50.4 ± 5	66.6
疲労 (SVM)	79.8 ± 1	76.4 ± 1	54.9 ± 5	70.4
集中 (KNN)	64.0 ± 3	82.9 ± 1	70.4 ± 2	72.4
眠気 (KNN)	86.0 ± 1	67.7 ± 2	54.0 ± 3	69.2
疲労 (KNN)	81.1 ± 1	77.3 ± 1	60.6 ± 5	73.0
集中 (RF)	66.7 ± 5	86.6 ± 1	74.0 ± 2	75.8
眠気 (RF)	89.2 ± 1	70.4 ± 2	59.7 ± 6	73.1
疲労 (RF)	84.9 ± 1	81.7 ± 1	64.8 ± 3	77.1
集中 (GBRT)	27.2 ± 4	79.2 ± 1	47.9 ± 4	51.4
眠気 (GBRT)	83.0 ± 1	34.7 ± 2	26.6 ± 3	48.1
疲労 (GBRT)	73.5 ± 2	68.2 ± 1	24.1 ± 6	55.3
集中 (ET)	74.2 ± 3	88.8 ± 0	79.4 ± 1	80.8
眠気 (ET)	91.2 ± 1	77.7 ± 1	67.5 ± 5	78.8
疲労 (ET)	87.7 ± 1	84.3 ± 1	73.7 ± 3	81.9

表 7 窓時間ごとの最も良いモデルとその平均精度

窓時間	3	5	10
モデル	ET	ET	ET
集中	64.0	72.7	80.8
眠気	61.5	71.5	78.8
疲労	66.1	74.3	81.9

超えていたラベルが、10%を下回る結果になり、80%を超えていたラベルも大幅に精度を落としている。時間窓同士で比較しても10分窓、5分窓、3分窓いずれもはっきりとした差がなくなっている。

表 8 未知のユーザに対するロバスト性評価の結果（3分窓のデータセット）

	低	中	高
集中 (SVM)	10.1	54.9	34.5
眠気 (SVM)	58.8	32.5	2.2
疲労 (SVM)	37.5	55.1	8.7
集中 (KNN)	9.0	54.8	34.9
眠気 (KNN)	57.5	32.0	3.4
疲労 (KNN)	37.4	53.1	9.3
集中 (RF)	6.6	55.9	33.2
眠気 (RF)	60.1	30.2	2.6
疲労 (RF)	37.8	56.6	7.0
集中 (GBRT)	12.4	48.2	39.2
眠気 (GBRT)	36.5	39.3	8.7
疲労 (GBRT)	19.8	59.7	10.3
集中 (ET)	2.8	61.1	30.1
眠気 (ET)	66.0	24.2	0.1
疲労 (ET)	41.2	54.4	7.0

表 9 未知のユーザに対するロバスト性評価の結果（5分窓のデータセット）

	低	中	高
集中 (SVM)	8.6	57.7	31.7
眠気 (SVM)	62.6	30.1	1.1
疲労 (SVM)	39.9	53.8	8.1
集中 (KNN)	9.1	53.1	35.8
眠気 (KNN)	58.9	29.8	1.1
疲労 (KNN)	38.1	53.3	8.8
集中 (RF)	3.8	57.5	33.0
眠気 (RF)	62.9	26.2	0.8
疲労 (RF)	36.9	54.1	6.3
集中 (GBRT)	11.6	46.4	39.2
眠気 (GBRT)	47.1	35.7	4.2
疲労 (GBRT)	31.3	60.4	10.8
集中 (ET)	1.7	62.1	30.9
眠気 (ET)	66.7	20.1	0.5
疲労 (ET)	40.9	53.8	1.1

表 10 未知のユーザに対するロバスト性評価の結果（10分窓のデータセット）

	低	中	高
集中 (SVM)	5.3	57.8	34.8
眠気 (SVM)	66.2	18.2	1.9
疲労 (SVM)	42.9	48.8	1.2
集中 (KNN)	9.6	53.6	38.7
眠気 (KNN)	64.5	27.2	2.5
疲労 (KNN)	33.7	52.1	7.4
集中 (RF)	1.8	60.8	37.2
眠気 (RF)	65.4	16.2	3.7
疲労 (RF)	36.6	53.8	2.0
集中 (GBRT)	7.1	54.2	39.4
眠気 (GBRT)	57.5	26.5	5.5
疲労 (GBRT)	29.3	57.5	5.9
集中 (ET)	1.0	62.2	32.4
眠気 (ET)	68.8	15.0	2.5
疲労 (ET)	39.9	51.5	1.7

5. 考察

個人によらないデータセットで学習したモデルでは、精神状態の度合に関して、一番精度が良かった10分窓のデータセットでETモデルでは、70%を下回る項目はあったものの、各状態度合のデータ数と照らし合わせると、400レコード程度の低集中、高疲労と、高眠気がそれぞれ一番低い精度となり、レコード数の多かった中集中、低疲労と、低眠気はそれぞれ87%を超える高い精度となった。そのことから、更なるデータを収集し、ラベルごとのアンバランスを縮小すれば、精度の向上が期待できると考えられる。このことから、心拍数のみでも、ある程度の精神状態度合推定ができることが示唆された。

また、個人によらないデータセットで学習したモデルで分析した結果が時間窓や分類器によって大きな差が生まれていたのに対し、個人によらないデータセットで学習したモデルで未知のデータを分析してみるとGBRT以外に大きな差が生まれず、40%から50%の間に収まっている。それぞれのモデルの学習が不十分であったから、すべてのモデルで低い精度になってしまったのではないかと考えられる。このことから、未知のデータに対しての推定は機能しないことが示された。しかし、個人ごとのキャリブレーションによって改善が期待できると考えている。

6. 結論と今後の展望

本研究では、日常生活で再現が容易なスマートウォッチの心拍数のみを用いて勉強中の複数の精神状態度合の推定を試みた。最初は、精神状態推定が可能と示唆されている脈拍変動と比べて、心拍数のみで十分な精度得られそうかを調べた。結果、脈拍変動を用いたモデルの方が精度が良かったが、差が平均で8.5%と差が小さいことから、心拍数

のみでもある程度の精神状態を判別することは可能であるということが示唆された。

次に、心拍数のみで被験者を増やし再度行った分析では、2つの分析手法を行った。第1に窓時間と学習モデルの評価であり、一番精度が良かった時間窓とモデルでは71%の平均精度で推定することができた。第2に未知のユーザに対するロバスト性評価であり、窓時間と学習モデルの評価の結果と比べて明らかに精度が落ちてしまっており、未知のユーザに対しては心拍数のみでは使い物にならない可能性が示唆された。しかし、先行研究で示されたように個人ごとのキャリブレーションによって改善が期待できると考えている [18]。

今後の展望として、データをより多く集めることや、学習モデルの最適化、特微量エンジニアリングの実施などをして改善を進めていく。

謝辞

本研究の実験データは株式会社 JSOL と提携の学習塾にて収集され、株式会社 JSOL と著者らの共同研究の一環として解析を行っている。

参考文献

- [1] 第18回21世紀出生児縦断調査(平成13年出生時)の結果, 入手先 (https://www.mext.go.jp/content/20200902-mxt_chousa01-000009574_1rrrrr.pdf) (参照 2021-05-04).
- [2] Y. Watanabe and Y. Ikegaya. Effect of intermittent learning on task performance: a pilot study. *NeuroNet*, Vol. 38, pp. 1-5, 2017.
- [3] 学習時間を細かく分けた「45分」で「60分」と同等以上の学習効果を発揮 “長時間学習”よりも短時間集中の“積み上げ型学習”が有効であった, 入手先 (<https://prtimes.jp/main/html/rd/p/000000562.000000120.html>) (参照 2021-01-09).
- [4] 大久保雅史, 藤村安耶. 加速度センサを利用した集中度合い推定システムの提案. *WISS*, 2008.
- [5] 植田智明, 杉村博, 松本一教, 一色正男. センサデータからの人間の行動推定. 2013年度人工知能学会全国大会, pp. 1-4, 2013.
- [6] 趙丹俊, 松田香菜, 片山喜規, 伊良皆啓治. 脳波・mirs・心電図による集中状態の評価, 電子情報通信学会技術研究報告. 電子情報通信学会技術研究報告, Vol. 111, No.423, pp.117-20, 2012.
- [7] S. Tanabe, R. Nagase, K. Fujita, H. Touyama, and Y. Tobe. Providing the attention level of a remote user using eeg. *IWSSS*, 2017.
- [8] 上田, 石田, 松田, 福嶋, 中道, 大平, 松本, 岡田. 脳波を利用したソフトウェアユーザビリティの評価. *ヒューマンインタフェース学会論文誌*, pp. 233-242, 2008.
- [9] K. Yoshida, Y. Sakamoto, I. Miyaji, and K. Yamada. Analysis comparison of brain waves at the learning status by simple electroencephalography. *Knowledge-Based Intelligent Information and Engineering Systems*, pp. 1817-26, 2012.
- [10] 平井章康, 吉田幸二, 宮地功. 簡易脳波計による学習時の思考と記憶の比較分析. *DI-COMO*, pp. 1441-1446, 2013.
- [11] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, Choo Min Lim, and Jasjit S. Suri. Heart rate variability: a review. *Medical and Biological Engineering and Computing*, 44(12):1031-51, 2006.
- [12] 角田啓介, 千葉昭宏, 千明裕, 吉田和広, 渡部智樹, 水野理. 心拍変動を用いた認知能力水準の逐次推定. *情報処理学会論文誌*, Vol. 58, No. 2, pp. 568-79, 2017.
- [13] 東海林可奈, 石井峻, 横窪安奈, ロベズギヨーム. 生体信号を用いたデスクワークにおけるメンタル疲労度推定. *DICOMO* 2019.
- [14] Nkurikiyeyezu, K.; Shoji, K.; Yokokubo, A. and Lopez, G. (2019). Thermal Comfort and Stress Recognition in Office Environment. In Proc. of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies - HEALTHINF, pp. 256-63. doi: 10.5220/0007368802560263.
- [15] Yuda, E., Shibata, M., Ogata, Y. et al. Pulse rate variability: a new biomarker, not a surrogate for heart rate variability. *J Physiol Anthropol* 39(21) (2020). doi: 10.1186/s40101-020-00233-x
- [16] 横田康成, 河村洋子, 松丸直樹, 白井邦博. 心拍変動系列変化を用いた敗血症の前駆症状モニタリング. 第54回自動制御連合講演会, pp. 1258-1261, 2011.
- [17] 心拍変動パラメータ用語集入手先 (<https://www.trytech.co.jp/checkmyheart/glossary.html>) (参照 2021-05-04).
- [18] Kizito Nkurikiyeyezu, Anna Yokokubo, and Guillaume Lopez. The influence of person-specific biometrics in improving generic stress predictive models. *Sensors and Materials* 32, pp. 703-722, 2020.