

早稲田大学審査学位論文  
博士（人間科学）

複合的アプローチによるヘルスデータ分析  
Health data analysis with integrated approaches

2021年 1月

早稲田大学大学院 人間科学研究科  
多胡 輝一  
TAGO, Kiichi

研究指導担当教員：金 群 教授

# 目次

第 1 章	序論	1
1.1	背景	1
1.2	目的	4
1.3	本論文の構成	5
第 2 章	関連研究	6
2.1	ヘルスデータ分析による診断推定	6
2.2	ヘルスデータ分析による異常検知	9
2.3	本研究の位置づけ	10
第 3 章	生活周期を考慮したヘルスデータ分析による診断推定	12
3.1	脈像からの脈特徴の抽出	13
3.1.1	中国伝統医学の概要	13
3.1.2	脈特徴の抽出手法	14
3.2	ランダムフォレストを用いた脈診断推定	15
3.2.1	ランダムフォレスト分類器の構築と評価指標	16
3.2.2	データセット	17
3.2.3	診断推定の結果と考察	18
3.2.4	性別に着目したクラスター分析の結果と考察	19
3.3	生活周期を考慮した脈診断推定	20
3.3.1	自己相関分析に基づく生活周期の推定手法	21
3.3.2	データセット	22
3.3.3	脈特徴量の抽出と生活周期推定の結果	24
3.3.4	診断推定の結果と考察	25

第 4 章	潜在的要因に着目したヘルスデータ分析による異常検知	31
4.1	潜在的要因の定義と定量化	32
4.2	構造方程式モデリングとクラスター分析を用いた潜在的要因定量化	37
4.2.1	ドメインモデルの作成と構造方程式モデリングの適用	38
4.2.2	活動因子に基づいたクラスター分析	38
4.2.3	データセット	39
4.2.4	実験の結果と考察	40
4.3	疲労度合い推定による異常検知	43
4.3.1	隠れマルコフモデルと HotellingT2 法による異常検知手法	43
4.3.2	データセット	46
4.3.3	疲労度合い推定の結果	46
4.3.4	実験の結果と考察	50
4.4	疾病リスク度合い推定に基づく異常検知アプローチの検証	52
4.4.1	疾病リスク度合い別の sGMRFmix アルゴリズムによる異常検知	52
4.4.2	データセット	55
4.4.3	実験データの準備と評価指標	56
4.4.4	実験の結果と考察	57
第 5 章	総合考察	62
5.1	診断推定研究と異常検知研究からの総合考察	62
5.2	本研究の特色と今後の展開	65
第 6 章	結論	67
6.1	本研究のまとめ	67
6.2	今後の課題	68
	謝辞	69
	参考文献	70

# 目次

1.1	提案する複合的アプローチ . . . . .	4
3.1	一般的な脈像 . . . . .	14
3.2	ランダムフォレストを用いた脈診機と医師診断の一致率検証実験の概要 . . . . .	16
3.3	階層クラスター分析によるヘルスデータの分類結果 . . . . .	19
3.4	診断推定のための複合的アプローチ . . . . .	20
3.5	自己相関分析結果の例 . . . . .	24
3.6	構築したディープラーニングモデル . . . . .	26
3.7	データの分割方法 . . . . .	27
4.1	ドメインモデル . . . . .	34
4.2	クラスター分析を用いた潜在的要因の分析概要 . . . . .	37
4.3	活動因子を分析するためのドメインモデル . . . . .	38
4.4	クラスタごとの活動因子の度合いの分布 . . . . .	41
4.5	クラスタごとのヘルスデータの分布 . . . . .	42
4.6	異常検知のための複合的アプローチ . . . . .	44
4.7	SEM と HMM の関係性 . . . . .	45
4.8	疲労度の要因モデル . . . . .	47
4.9	疾病リスクを潜在的要因としたドメインモデル . . . . .	53
4.10	因子得点の推移 . . . . .	54

# 表目次

1.1	代表的なウェアラブルデバイスと取得できるデータ . . . . .	3
3.1	先行研究で示されている脈特徴量の重要度 . . . . .	15
3.2	ランダムフォレストによる一致率検証結果 . . . . .	17
3.3	実験協力者の概要 . . . . .	23
3.4	シングルラベル分類での GBDT による分類精度 . . . . .	26
3.5	マルチラベル分類の分類精度 . . . . .	26
3.6	ADASYN を適用した後のマルチラベル分類精度 . . . . .	29
3.7	6 つの機械学習手法による分類精度比較 . . . . .	30
4.1	要因の種類 . . . . .	33
4.2	ウェアラブルデバイスとスマートフォンによって取得されたデータ . . . . .	40
4.3	モデル適合度の比較 . . . . .	41
4.4	異常検知の結果 (男性 A) . . . . .	48
4.5	異常検知の結果 (女性 E) . . . . .	49
4.6	男性 A のドメインモデル適合度 . . . . .	50
4.7	女性 B のドメインモデル適合度 . . . . .	50
4.8	ウェアラブルデバイスと HMM による疲労度推定数の比較 (男性 A) . . . . .	51
4.9	ウェアラブルデバイスと HMM による疲労度推定数の比較 (女性 E) . . . . .	52
4.10	疾病リスク定量化のためのドメインモデル適合度 . . . . .	53
4.11	実験で利用した異常データ . . . . .	54
4.12	ベースラインとなる異常検知手法と本アプローチとの比較 . . . . .	56
4.13	異常検知の結果 . . . . .	58
4.14	5 交差検証による検証結果 . . . . .	59

5.1	診断推定における関連研究との比較 . . . . .	63
5.2	異常検知における関連研究との比較 . . . . .	65

# 第1章

## 序論

---

この章では、背景と本研究の目的について述べる。現代の社会がかかえる健康に関する問題を明らかにし、健康状態の把握が重要であることを示す。さらに、ヘルスデータを取得することが容易になった現代において、個人のヘルスデータを用いた分析やサービスについて紹介し、それらが抱える課題について述べる。最後に、指摘した課題への取り組みとして、従来ではあまり焦点を当てられていなかった人間が持つ特徴や潜在的要素についても焦点を当てて分析を行い、これまでの手法と組み合わせる複合的アプローチを提案することを示す。

---

### 1.1 背景

現代においては、寿命を延伸させるだけではなく、生活の質（QoL: Quality of life）を向上させ Well-being に生きることが非常に重要であり、その考えとともに健康寿命という言葉も普及している。QoL を向上させるためには、運動や食事に気を配るとともに、自身の健康がどのような状態であるかを日頃から把握することが大切である。しかし、心疾患や脳血管疾患など自分が気づかない間に進行している疾病も多くあり、迅速な発見や予防は健康寿命を伸ばす上で大きな鍵となっている。また、医療費の増大や介護需要の増加といった健康に関する社会の課題も年々深刻さを増しており [1]、社会問題としても健康管理の重要性は日々増している。健康管理を日々の生活の中で行うためには、主観的な体調の評価だけではなく、自分の健康状態を連続的に記録して分析するためのデバイスやシステムなどが必要となる。

センサーやバッテリーの小型化や性能向上に伴い、健康管理を行うことができるデバイスとして、近

年ウェアラブルデバイスが普及してきた。以前から歩数を計測する万歩計等は市販されているが、近年では血圧や心拍数を計測できる多機能な腕時計型デバイスや、集中力や視線を計測できるメガネ型デバイスなど、多くの種類のウェアラブルデバイスが販売されている。ウェアラブルデバイスが一般的になるにつれ、ユーザーは自身のヘルスデータを容易に記録できるようになった。ウェアラブルデバイスは健康管理や医療費の削減を目的として作成されており [2]、生活の中でほぼ意識すること無く装着して過ごすことができる。スマートフォンやパソコンに簡単にデータを蓄積して確認できるため、特別なセンサーやシステムなしで、誰でも血圧や心拍数の把握を行えるようになった。これらを活かし、個人の健康増進と生活の質の向上に関する研究プロジェクトもヨーロッパやアメリカなどで大規模に実施されており\*1、健康に関する問題意識はますます高まっている。複数のセンサーやデバイスを使用した人間の行動検出などの個人データ分析も近年研究されており [3]、人間に注目したヘルスケア研究も盛んになってきている。これらの研究でもセンサー付きリストバンドやスマートウォッチなどのデバイスが使用されているため [4, 5]、その活用方法については多くの可能性があるといえる。

ウェアラブルデバイスなどを通じて生活する中で得られた個人の健康と密接に関係するデータとして、ヘルスデータがある。ヘルスデータに関してはパブリック・ヘルスや統計処理されたデータとしてのヘルスデータの言葉もあるが、本研究でのヘルスデータは、個人から取得された健康に関わるデータを指す。ヘルスデータは、心拍数や血圧だけでなく、年齢や身長、心電図の波形、ゲノム情報などその種類は多岐にわたっており [6]、膨大な量となるため一種のビッグデータとも考えられる。市販されているウェアラブルデバイスも近年増えてきており、表 1.1 に示すように、デバイスの種類や取得できるデータの種類も多く、日常的に様々なデータを取得できるようになった。また、歩数や運動の活動記録など、直接ではないが健康に関わるデータも含むヘルス関連データも取得することができる。そのようにデータが簡単に記録できるようになった現在においては、どのように健康管理のために活用するかが次の課題として浮かび上がってくる。血圧や歩数といったヘルス関連データが日々どのくらいであったのか、週単位平均や簡単なグラフなどの分析結果を見るサービスはすでに提供されている\*2。しかし、個人の持つ特徴や個人差、コンテキストまで考慮し、本当にその個人に合わせたサービスやデータの分析まで行われているとは言い難い。例えば、急激な心拍数の上昇があった際、どのタイミングで警告を出すべきか、出すとしてもどのような警告の種類にするかなど、その個人の持つ特徴やコンテキストによって変えることが望ましい。

異常検知に関するサービスとしては、Apple Watch による転倒検知による異常検知がある。傾きを検

---

\*1 Horizon2020, <https://www.ncp-japan.jp/about> (最終確認日 2020/12/15)

\*2 Fitbit アプリ, <https://www.fitbit.com/jp/app> (最終確認日 2020/12/15)

表 1.1 代表的なウェアラブルデバイスと取得できるデータ

デバイス名	デバイスの種類	取得できるデータ
Apple watch	腕時計型	心拍数, 歩数, 睡眠時間, 血中酸素, 心電図, 呼吸数, GPS, 高度
Fitbit	腕時計型	心拍数, 歩数, 睡眠スコア, 血中酸素, 呼吸数, 血圧
JINS MEME	メガネ型	集中力 (瞬き, 視線移動, 姿勢から計算), 活動量
Owlet smart sock	靴下型	幼児の心拍数, 体温, 血中酸素, 睡眠データ
COCOMI	T シャツ型	心拍数, 呼吸, 筋肉活動量

知るセンサーが内蔵された腕時計型デバイスにより転倒を検知し、1 分間応答がなければ緊急通報を行うというものである。また、Fitbit では運動量が少ない場合に振動によってアラートを出してくれるなど、健康促進のためのサービスもすでに存在する。しかし、これらのデバイスでは人間の内部に潜む疾病リスクや状態までは考慮されていない。診断推定に関するサービスでは、東芝から 2020 年に「疾病リスク予測 AI サービス」という提供が開始された\*3。このサービスでは、糖尿病や生活習慣病のリスクを個人の検診データから予測し、フィードバックを行うことができる。しかし、検診の機会は多くないことや、日々気軽に最新のデータを用いて分析することができないなど、未だ課題も多いと言える。そのため、ヘルスデータ分析は人間の状態を考慮しない画一的な計算論的アルゴリズムによって分析できるようなものではなく、人間を中心に考えて多角的に個人を分析していく必要がある。しかし、未だ個人の特徴や状態まで考慮されているとは言い難く、人間中心の観点からヘルスデータ分析を行うためには多くの課題が残されている。

人間科学的な視点からヘルスデータ分析を行うためには、そのための分析モデルや従来手法の拡張が必要である。本論文における人間科学的視点とは、これまで情報科学分野で考慮されることが少なかった人間自体に焦点を当てて分析を行うだけでなく、人間が直面する問題に対して、分野を横断して学際的に研究を行うことによりさらなる改善や新たな知見を得るための視点である。しかし、人間が持つ様々な要素に対してどのようにして情報科学とヘルスケアの分野を横断して分析すればよいか、そしてどのように組み合わせればよいかについては、未だ多くの課題が残っている。人間中心でヘルスデータ分析を行う際には人間がもつ特徴や要素を分析するための新しいモデルや分析手法が必要であるため、これまでの画一的な閾値や計算によるものではなく、複数の手法を組み合わせた統合的アプローチが求められている。

\*3 東芝 AI 技術カタログ, <https://www.toshiba.co.jp/tech/ai/catalog004.htm> (最終確認日 2020/12/15)

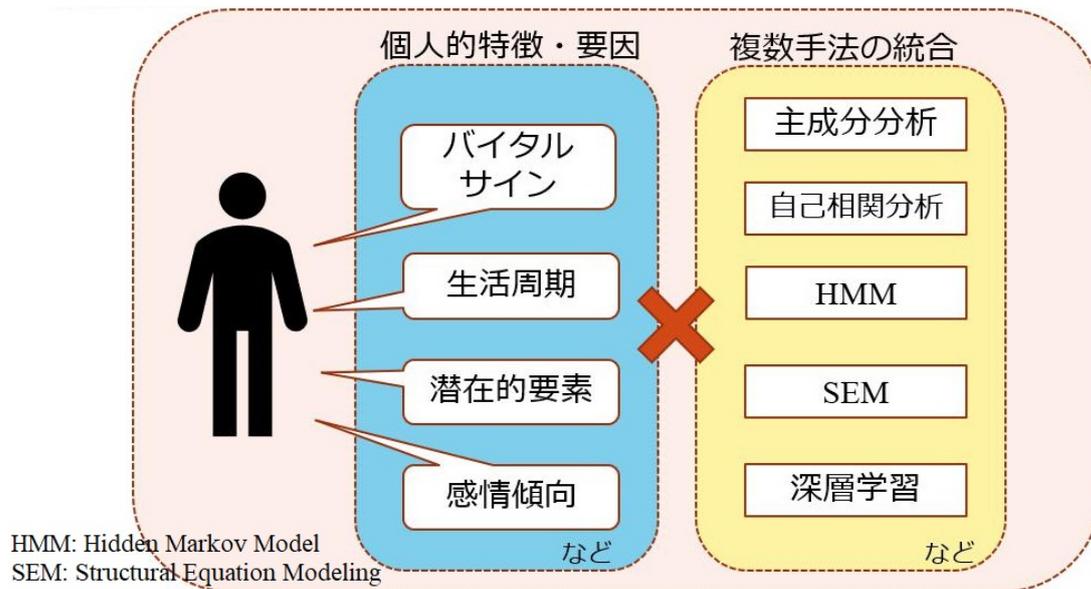


図 1.1 提案する複合的アプローチ

## 1.2 目的

上述したように、デバイスから得られたヘルスデータに対しては、人間中心の視点から人が持つ特徴や要素に対しても分析を行い、その個人に合わせて結果をフィードバックすることが望ましい。個人が持つ特徴を従来の手法と組み合わせて評価する場合、これまでの手法だけでは十分であるとはいえず、別途従来の手法と組み合わせるなどして統合的な手法を用いる必要がある。例えば、病気を引き起こす要因の度合い定量化、それを用いた異常検知手法の拡張、健康状態の推定など多くの分析が必要であり、それらを組み合わせた新たな分析アプローチが求められる。

そこで本研究では、人間科学的視点から人が持つ生活周期や個人の特徴等を踏まえ、複数の手法を組み合わせて分析を行う複合的アプローチを提案する。ここでは、複合的アプローチを用いて分析を行った2つの研究事例について紹介し、我々のアプローチの有効性を検証する。検証を通じて、従来のアプローチと比較して提案する複合的アプローチが有効であるかどうかを明らかにすることを目的とする。

複合的アプローチや複合的分析という言葉は他の研究にも見られるが [7, 8]、その言葉は分野によってバラバラであり定まっていない。本研究で提案する複合的アプローチとは、図 1.1 に示すように人が持つ個人的特徴・要因といった多角的な視点から、統計学的な分析手法や機械学習の分析手法を組み合わせ、複数の手法を用いて分析するアプローチである。複合的アプローチでは、ヘルスデータ分析の対

象は人間であるということを踏まえ、これまでの分析では考慮されてこなかった人間の持つ様々な特徴や要素にまで焦点を当てて分析を行う。さらなる分析を行うためには、統計的手法、機械学習、ディープラーニングなど多くの手法が必要となるため、それらの手法を組み合わせてよりその個人に合わせた結果を得るようにする。複合的に分析することにより、従来研究の結果改善だけでなく、手法を組み合わせることによる相乗効果や、結果を得られるプロセスの見える化、または人が持つ特徴の差を考慮した分析結果も期待することができる。例えば、診断推定のアプローチに生活周期の分析を加えることで、従来の推定精度を改善するだけでなく、その個人の生活周期やその個人差なども明らかにすることができ、さらにはその周期分析手法を他の分析に応用できるなど、多くの成果を見込むことができる。

本研究では、従来の分析だけでなく人間が持つ様々な特徴や要素まで複数の手法を用いて分析して組み合わせる複合的アプローチについて、その有効性を2つの研究事例によって検証する。1つ目の研究では、診断の推定を行う研究に複合的アプローチを適用する。従来の推定手法だけでなく、統計的手法による生活周期の推定やそのスパンを基とした主成分分析（PCA：Principal component analysis）による次元圧縮を組み合わせた診断推定のための複合的分析アプローチについて検証を行う。2つ目の研究では、健康に対する異常検知を行う研究に、複合的アプローチを適用する。従来の検知方法に加え、ヘルスデータに構造方程式モデリングとドメインモデルを用い、病気を引き起こすリスクや疲労度といった潜在的要素を定量化し、さらにその度合を隠れマルコフモデルで状態推定する複合的分析アプローチについて検証を行う。

### 1.3 本論文の構成

本論文の残りの部分は、以下によって構成されている。第2章では、ヘルスデータを用いたヘルスデータ分析について関連研究について述べ、それらと比較した本研究の立ち位置を示す。第3章と第4章では、複合的アプローチに基づいておこなった研究事例を紹介する。第3章では、ヘルスデータからの診断推定のための複合的アプローチを提案し、個人の生活周期を推定し、それをもとに従来の脈診断推定の精度向上が行えるかを検証した研究を示す [9, 10, 11]。第4章では、ヘルスデータを利用した異常検知のための複合的アプローチを提案し、潜在的要因を考慮することで既存の異常検知とほぼ同等の精度を持ち、さらに結果を得るプロセスが見える化した研究を示す [12, 13, 14]。第5章では本研究の総合考察を行い、第6章で本研究のまとめと今後の課題を示す。

## 第2章

# 関連研究

---

この章では、これまでに行われてきた診断推定と異常検知に関する関連研究を示す。ヘルスデータを用いた研究は、病気の兆候検出や診断推定、患者の特徴分析など、ヘルスデータ分析の中でも広く行われている。それらがどのような手法を用いてこれまで研究が行われてきたかを紹介し、人間が持つ特徴や潜在的要素を考慮した研究がほとんど行われてこなかったことを明らかにする。その上で、本研究の立ち位置として、人間科学的視点からそれらに対して分析を行い、従来の研究と組み合わせる複合的アプローチを提案することについて述べる。

---

### 2.1 ヘルスデータ分析による診断推定

ヘルスデータは歩数や、血圧、心拍数といった本人の健康状態が推測できるデータであり、それを基に多くの研究で健康状態の推定や医師の診断の推定が行われている。患者の健康状態を推定した研究 [15] では、ディープラーニングを用いて推定を行った。しかし、人間中心的な視点から複数の手法を組み合わせた分析の例はあまり見当たらない。

Rajkomar et al. [15] は、退院した患者が再度入院するかどうかについてなど、医師のカルテからその後の経過を予測するモデルを提案した。彼らは、複数の医療診断情報のフォーマットを統一的なフォーマットにすることでビッグデータとしての分析を可能にし、患者がどのような経過をたどったかについてや退院したときの状態をディープラーニングで学習させた。その結果、将来的に患者がどのような経過をたどるかというリスク評価の精度を改善させた。Lutze and Waldhor [16] は、スマートウォッチを使用して高齢者のパーソナルヘルスサポートアシスタントのために人間の活動を検出するア

アプリケーションを開発した。そのアプリケーションはニューラルネットワークベースの仕組みを持ち、ウェアラブルデバイスに組み込まれている。動きや睡眠の状態を検知するだけでなく、分析を通じて優れた歯磨きの検知などの応用ができることが述べられている。Liang et al. [17] は、ウェアラブルデバイスによって取得された睡眠追跡データを調査した。異常な睡眠パターンを検知するために順序エントロピーで次元削減を行った後、単変量解析によって異常が検知される。彼らは、異常かどうかの正解の情報がないためさらなる検証が必要であることや、主観的な感覚との関係性を調べる必要があると述べている。Merilahti et al. [18] は、ウェアラブルデバイスを使用して高齢者から取得した概日活動や睡眠パターンなどの長期データを分析した。認知症を含む研究対象者に対して分析を行ったところ、身体機能状態は概日リズムにおける運動強度と関係性があることが明らかとなった。また、身体機能の低下は、活動リズムの強さの低下やおよびリズムの断片化と関係があることが述べられており、活動量を把握することによって身体機能の低下防止につながるということが考察として述べられている。Guo and Ma [19] は、スマートフォンから取得した個人データに基づいて類似の個人を集約し、データベースに格納するアーキテクチャを提案した。ウェアラブルデバイスから送られてくるデータを、スマートフォンを通じてどのように管理するか、プライバシーをどのように保護するか、多くのデータからどのように個人の特徴分析を行うかについての仕組みを提案し、パーソナリティ分析に活用することを提案した。Schuldhaus et al. [20] は人体に4つの慣性センサーを配置し、93.9%の精度で活動を予測できる分類システムを開発した。システムにおいては各センサーが特徴量抽出から動きを判断し、その多数決によってどの活動内容であるかを分類するという仕組みになっている。

ヘルスデータ分析の一分野である中国伝統医学（中医学）に関する研究では、診断結果を推定するための研究が広く行われている。中医師の主観的な感覚によって行われる脈診断を定量的に判定するため、センサーの開発や脈像からの自動診断などが行われている。Yan et al. [21] は、パルス波形速度 (PWV) とパルス反射率 (Rf) が脈診断の推定に重要な指標であることを明らかにした。4種類の脈診断結果に対し、それぞれの脈状にはどのような特徴があるかを分析し、4つの違いを示した。Wang et al. [22] は、静脈のスペクトルグラフを脈特徴量として用い、後方伝播ニューラルネットワークを用いて静脈を6つの脈状に分類した。ニューラルネットワークが広く普及していなかった2007年当時において、平均して88%という高い精度で分類できることを示した。しかし、実際の診断では複数の脈状が診断されることもあるが、彼らの実験では一つの種類だけの分類であるため、脈診断に対する機械学習の展望を示した研究であると言える。Li et al. [23] は、脈像から特徴量を抽出し、勾配ブースティング決定木 (GBDT: Gradient Boosting Decision Tree) を適用することで脈診断を推定した。決定木ベースでの分類では説明変数に重要度のスコアを計算することができるため、Rf や h1, h3 などの脈像

から得られる特徴量が相対的にどの程度重要化を明らかにした。Ma et al. [24] は、脈を推定する際、二段階のステップで推定をすることにより精度が向上することを示した。Yan et al. [25] は、PWV と Rf を特徴量として用いることで、KNN アルゴリズムによって 4 種類の脈を分類できることを示した。Li et al. [26] は畳み込みニューラルネットワークを採用し、おおよそ 95% の精度で脈診断を推定した。Zhang and Sun [27] も、9 層の畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) によって 12 種類の脈を 93.49% の精度で分類できることを明らかにした。これは、脈像を画像として特徴量抽出などをせずそのまま分析対象とし、CNN という画像処理に用いられるディープラーニング手法を適用するものであり、R f や h1 などではなくそのまま分析に用いることができるといった利点がある。Guo and Wang [28] は、確率的ニューラルネットワークを使用して 93.03% の平均精度で 12 の脈を識別できることを報告した。彼らの手法は、後方伝播ニューラルネットワークよりも精度が高く、ノイズにも頑強であることが実験によって検証されている。

また、その他にも中医学に関するヘルスデータ分析として、以下のような研究がおこなわれている。Guo et al. [29] は、高血圧や健康な人の脈をマルチスケールエントロピーを用いて分析し、その特徴を明らかにした。分散分析による結果から、冠状動脈性心臓病、高血圧、健康の 3 グループにおいて、有意差 (サンプル数 = 570,  $p < 0.001$ , F 値: 記述なし) があることが示された。Li et al. [30] は、マルチラベル分類とアンサンブル学習を組み合わせた分類モデルを基に、中医学における高血圧の人を分類する手法を提案した。ここでは、マルチラベル分類の代表的な手法である Multi k-nn など 5 種類の手法を比較し、最大で 68% の精度が得られることが述べられている。Wu and Guo [31] は、これまでの中医学に関する文献のレビューを行い、診断推定のアプローチに関して整理するだけでなく、高齢者の健康管理を地域的に行う際に、中医学の概念を導入した健康管理やアドバイスを提唱した。Layne et al. [32] は、中医学が心血管疾患に有効であるかどうかについてメタアナリシスを行った。その結果、中医学の治療を受けている患者は、西洋的な医薬品の適用に匹敵する血圧の効果が見られ、その有効性が確認された。Guo et al. [33] は、顔の画像を分析し、中医学の医師の診断と組み合わせて健康状態を推定するシステムを提案した。中医学では、人を総合的に捉えて健康かどうかを判断するため、顔色や表情なども診断に考慮されている。そのため、顔の画像を用いることで、診断推定の結果や精度の改善が期待できることが述べられている。さらに、Guo et al. [34] では、主観的な医師のバイアスを除くための一貫性アルゴリズムを提案した。Jin et al. [35] は、症状と漢方薬の共起関係をマルチグラフ畳み込みネットワークを用いて分析することにより、高い精度での推定が可能であることを示した。

中医学における分析だけではなく、遠隔医療や ICT を活用したサービスや健康支援もすでに行われ始めている。例として、手首にセンサーを取り付け、脈の圧力やその速さから診断推定を行う脈診機が

販売されている。それを基に遠隔医療を行う企業も中国で現れており<sup>\*1</sup>、さらに病院との連携も進められている。日本においても、北里大学と富士通が協力した「東西医学を統合したヘルスケアシステムをつくる－健康寿命を延ばす、豊かな社会への挑戦」というプロジェクトが2014年から開始されており、触診などを標準化してICTを活用できるように研究やデバイス開発が進められている。しかし、日本では病院といえば西洋的な治療を受けるイメージが強く浸透しており、定量的な判断が難しい中医学については疑問を持つ声も多い。そのため、中医学に関する自動的な診断推定や遠隔医療についても、いまだ多くの課題が残されている [36].

## 2.2 ヘルスデータ分析による異常検知

ここ10年ほどの期間においては、特にセンサーとスマートウォッチを備えたリストバンドなどのデバイスが幅広く使用されている [4]. また、ヨーロッパにおいても人々の健康増進やQoLの向上を目的としたプロジェクトが進んでおり、研究において大きな関心が寄せられている。そのため、ヘルスデータを用いた研究の中でもバイタルサインや睡眠に関する異常検知が多く行われている。

Liang et al. [17] は、腕時計型のウェアラブルデバイスを用いて一人の女性から睡眠時間を取得し、睡眠データに対する分析手法の提案及び検証を行った。彼らの提案した手法では、順序エントロピーを用いた次元削減をおこない、統計的異常検知をすることで、睡眠の質に対する異常検知を行うことができる可能性が示された。Balasubramaniyam [37] は、Tシャツ型のウェアラブルデバイスを提案し、呼吸数や肺活量をモニタリングすることによる異常検知を提案した。

「異常なデータを検出する多くのアプローチでは、正常なデータを基にして距離を計算したり機械学習を行うことによって判断する事が多い [38].

距離を基にしたアプローチでは、レコードと特定の値の間の距離が計算される。異常かどうかとする基準を閾値として設定し、得られたデータと基準との距離が閾値を超えると異常データとして検出される。Hotellingのコントロールチャート [39] は、距離を基に異常を検出する古典的な手法である。この手法は、変数の平均と分散から異常の度合いを計算する方法であり、正常なレコードが多い場合に使用される。Ma and Lee [40] は、長期的な患者のモニタリングシステムを提案し、Hotellingのコントロールチャートをそのシステムに採用した。他の手法としては、主成分分析で異常度を求める手法もあり、Amor et al. [41] は、PCAを使用して多変量解析を実行する方法を提案した。彼らの方法では、はじ

---

<sup>\*1</sup> 36Kr Japan, 「中医学でもリモート診療 IoT 活用のスマート端末で脈拍を再現 遠隔脈診まで可能に」, <https://36kr.jp/80287/> (最終確認日 2020/12/15)

めに多変量解析から異常度が計算され、異常度が閾値を超えている場合、異常なバイタルサインが何であったのかが調査される。

機械学習や予測によるアプローチでは、k 近傍法 (KNN: k-nearest neighbor) がたびたび使用される。Yan et al. [42] は、健康関係のビッグデータにその手法を採用し、KNN を改善して異常を分散処理においても発見できる手法を提案した。KNN は新しいデータから近い k 個のデータを参考にし、多数決のように一番多いラベルに分類するものであるため、正常なデータがまとまっている事例に対して良い性能を持つと考えられる。その他では、異常を確率的に予測をするため、ベイジアンネットワークなどの確率モデルを使用することもある。Kawanishi et al. [43] は、認知症の特徴を異常と見なし、絵を描く患者の動きから異常なデータを検出した。彼らは変分ベイズ法に基づく手法を採用し、教師なし学習による異常検出を行った。Hela et al. [44] は、日常の環境要因の異常を検出することを提案し、マルコフ連鎖ネットワークを使用して異常を予測した。これは、複数の事象に対する確率を考慮して確率を計算することができるため、因果関係などを組み込んでモデル化するときによく使われるモデルである。

また、近年では異常検知のためにディープラーニングを用いることも増えており、その流れは勢いを増している。Caroprese et al. [45] は、ディープラーニングによる電子医療データ処理のためのツールを調査し整理した。Hsiao et al. [46] は、ディープラーニングを使用して心血管疾患のリスクを分析し、生活環境と疾患リスクの関係を分析した。Li et al. [47] は、ディープラーニングと協調フィルタリングを使用して、糖尿病患者の再入院を予測した。しかし、ディープラーニングでは、結果がどのようにして得られたのかを説明できないという制限がある。ディープラーニングは結果を得るプロセスがブラックボックスであり、学習した中間層の仕組みを明らかにすることは難しいことが知られている。そのため、結果の解釈可能性を向上させるための研究も行われている [48, 49]。

## 2.3 本研究の位置づけ

関連研究では、ヘルスデータを用いた研究が近年広く行われていることを示し、異常検知や診断推定に関わる研究を主に紹介した。しかし、それらは個人の特徴や要素までを考慮することはほとんどされず、画一的な手法によって分析が行われている。また、分析手法や分析における指標を複数組み合わせられている研究 [50] など、ヘルスデータ分析の領域以外では複数指標を用いたものも散見される。例えば、人の目線や注意に基づき、画質の向上のために補完を行う際に人のどこが注視されているかを分析して重みをつけることによって補完品質を向上させる研究 [51] 等がある。このように、これま

でのヘルスデータ分析を発展させて人間中心の視点から複数の指標を取り入れることによって、その個人の特徴を考慮した分析を行うことができるようになり、より個人に適した結果をフィードバックすることが見込める。また、ディープラーニングは高い精度の結果が得られることから近年多用されている手法であるが、ヘルスデータ分析においては信頼性を求められることも多く、特に治療に対しては間違いが許されないためその適用を慎重にならざるを得ない。したがって、ディープラーニングと精度と同等の精度を持ち、結果が得られるプロセスが明確である手法も求められている。

本研究では、上記の問題の解決に貢献できるよう、複合的アプローチを新たに提案する。人間科学的視点から、疾病リスクや生活周期にまで焦点を当て、複数の手法や指標を用いて分析を行う法を示し従来研究の発展を試みる。従って、本研究の立ち位置は、これまでの研究手法と人間が持つ特徴や要素の分析手法を組み合わせることにより、新たにヘルスデータ分析のための複合的手法を提案するものである。

## 第3章

# 生活周期を考慮したヘルスデータ分析による診断推定

---

この章では、従来の診断推定手法に対して、生活周期を分析して組み合わせる複合的アプローチについて述べる。実験では中国伝統医学で行われている脈診断の推定を対象とし、生活周期を分析して組み合わせる複合的アプローチについて提案をおこない検証する。生活周期は自己相関分析によって推定され、得られた1周期の期間分のヘルスデータに対してPCAが行われてヘルス特徴量が抽出される。従来の脈特徴量とヘルス特徴量を組み合わせることにより、マルチラベル分類での推定精度が改善されることを示す。

---

健康に生活を送る上で、自身の健康状態を知ることが非常に重要である。定期検診に行き医師による診察を受ければ健康状態を確認することができるが、その頻度を高くすることは難しく気軽に行うことは出来ない。もしも身近なデバイスで簡易診断を受けることができれば、忙しい人や過疎が進んで医師が不足している地域の人々でも、気軽に健康管理を行うことが可能となる。

多くの人は数日単位での生活周期を持っており、その周期は健康と大きく関係している。例えば、学生や会社に行く人を考えた場合、平日5日間は通学・通勤し、休日2日間では外出や休息などプライベートの時間を過ごすというリズムがある。この場合、そのような人々は7日間という生活周期を持っていると言える。また、高齢者の場合では、1週間ではない場合も考えられる。

そのような生活周期はユーザーが可視化されたグラフなどを見て推測することが主であり、ヘルスデータ分析にまで活用されているとは言い難い。ヘルスデータ分析では、年代や性別などの属性が同じ

人々を対象にして研究が行われることが多いが、それらの人々の中でも生活周期は異なることがある。個人の健康に関わる分析を行う際には、属性に現れない生活周期も関係していると考えられる。したがって、ヘルスデータを詳細に分析することによって従来の研究を改善できる余地があり、人間中心の視点から分析することでさらなる推定の精度向上が期待できる。

本研究では、中国伝統医学の理念である人間を総合的に捉えて健康を考えるという考え方を参考にし、人間が持つ生活周期を考慮することにより脈診断の推定精度を向上させることを目的とする。そのためには、生活周期を分析するための手法が別途必要であり、従来の研究とどのように組み合わせるかについても考える必要がある。ここでは、生活周期を分析する手法やその組み合わせによる診断推定のための複合的アプローチを提案し、その有効性について検証を行う。それに加えて、推定に用いる脈像の特徴量の基準を変化させると精度がどのように変化するのかについても検証を行い、考察する。

## 3.1 脈像からの脈特徴の抽出

### 3.1.1 中国伝統医学の概要

個人の生活周期が大きく関係している診断の一つに、脈診断がある。脈診断とは、中国伝統医学（中医学）で行われている診断方法である。これは健康に対する考え方や治療方法が西洋医学とは大きく異なっており、人体を総合的に捉えることに特色がある。したがって、中医学は医学的な分野で人間中心の視点からの治療を行っていると言える。

このように、西洋医学とは大きく異なる考え方である中医学はその治療方法も異なる。西洋医学では原因となる物質を科学的に取り除くことを第一とするが、中医学では人間が持つバランスを整えて自分の力により病気を治す治療を目指している。そのため、診断方法は原因となる部分だけを見るのではなく、望診、聞診、問診、切診という4つの診断方法を用い、その人の体調全体を医師が判断する。治療においても、もし他と同じ症状で合ったとしても画一的な治療方法はなく、その人の体調や体質に合わせた方法を取り入れている。食事指導、漢方薬の処方、生活アドバイス、整体などが行われ、薬に頼るだけではなく自然治癒力を高める手伝いをするという考え方である。

そのような中医学において、脈の診断は中医学の典型的な診断方法の1つである。中醫師は専門的な訓練により、手首の脈に3本の指を当てて診断を行う。指先の感覚に基づいて、脈の強さと速度から患者の健康状態を診断する [52]。その脈診断は中醫師の指先が持つ感覚による診断であるため、定量的な診断が難しい。自動化を行うためには、未だ多くの課題が残っていると言える。

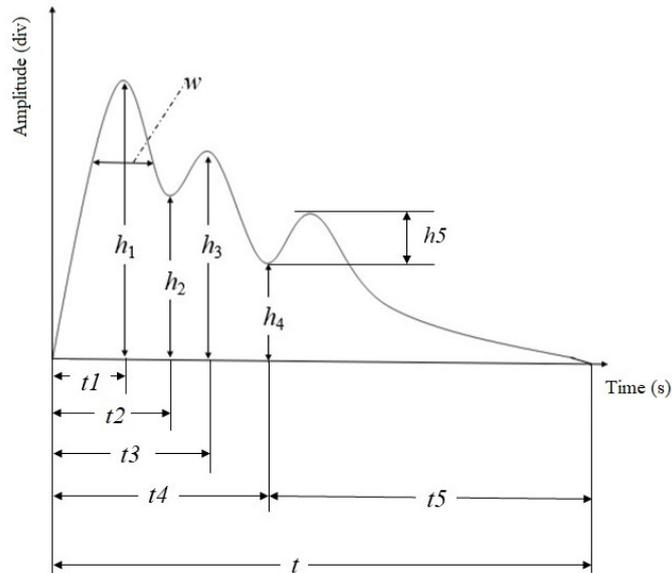


図 3.1 一般的な脈像

### 3.1.2 脈特徴の抽出手法

脈診断の推定に関する研究では、脈の特徴量を記録した脈の波形（脈像）が一般的に用いられている。脈像から脈の特徴量を抽出し、それを基に機械学習や統計的手法によって分析が行われている。

図 3.1 は、代表的な脈像である。この図では、x 軸が時間、y 軸が振幅を表している。h1 や t4 など脈像から得られる特徴量は、多くの研究で脈特徴量として採用されている [23]。

脈像からは多くの種類の特徴量を得ることができるが、機械学習に脈特徴量を用いる場合、どの特徴量を用いるべきかを決定する必要がある。Li et al. [23] の研究では、勾配ブースティング決定木 (GBDT: Gradient Boosting Decision Tree) を用いて、診断推定を行う際の重要度をスコア付けした。GBDT は、決定木ベースのアルゴリズムであり、分類を行った際の各特徴の重要度をスコア化することができる。表 3.1 は、Li et al. [23] が計算した各特徴のスコアを示している。彼らは、 $h_3/h_1$ 、 $h_4/h_1$ 、 $w/t$ 、 $Rf$  の上位 4 つの特徴量を機械学習に使用した。この表 3.1 を見てもわかるように、0.9 以上のスコアを持つ特徴量を採用していることがわかる。

表 3.1 先行研究で示されている脈特徴量の重要度

特徴	スコア
h3/h1	1
h4/h1	0.986
w/t	0.977
Rf	0.953
w	0.839
t3	0.763
h3	0.731
h5/h1	0.653
h1	0.516
t4	0.509
t5	0.431
t	0.396
h5	0.362
PWV	0.326
t1	0.218
h4	0.198

## 3.2 ランダムフォレストを用いた脈診断推定

この実験では、医師の診断を機械学習によって推定した場合と脈診機による診断の精度を検証するため、ランダムフォレストを利用して診断結果の推定を行う [9]。本実験の概要を図 3.2 に示す。はじめに実験協力者のヘルスデータを収集して説明変数として用い、目的変数である医師の診断を学習する。この学習により、ヘルスデータから医師の診断を推定するための分類器を構築することができる。次に、構築した分類器を用いて、ヘルスデータから医師の診断を推定する。その推定された診断と脈診機の診断結果の一致率を検証することで、医師の診断を推定した場合と脈診機の診断がどの程度一致するのかを検証する。

機械学習で構築する分類器としては、ランダムフォレストを採用した。ランダムフォレストを採用し

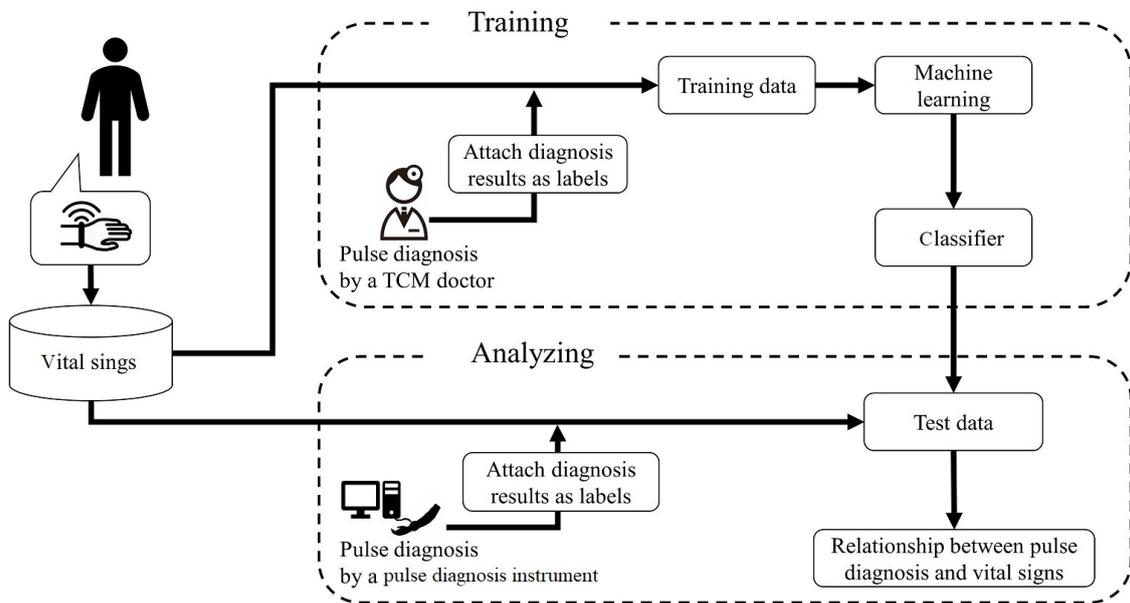


図 3.2 ランダムフォレストを用いた脈診機と医師診断の一致率検証実験の概要

た理由は、マルチラベル分類に対応しているためである。脈診断は、一度の診察で複数の状態を診察されることもあるため、マルチラベル分類に対応した手法が望ましい。ランダムフォレストは、複数の決定木を使用したアンサンブル学習に基づく機械学習手法であり、マルチラベル分類を行うことが可能である。上記の理由により、推定された医師の診断と脈診機の診断についてマルチラベル分類での精度を検証した。

### 3.2.1 ランダムフォレスト分類器の構築と評価指標

ランダムフォレストはマルチラベル分類に対応しているが、複数のラベルを持つレコードをそのまま訓練データとして使用することはできない。この問題を解決するにはいくつかの方法があるが [53]、今回はその中で *dubbed copy-weight* と呼ばれる方法を採用した。これは、マルチラベルとなっている元のデータをコピーして、 $n - 1$  種類のデータを複製して作成する方法であり、ここで、 $n$  はラベルの種類数を示す。もともとのデータとコピーされたデータにはそれぞれ 1 つのラベルが付与され、一つのラベルを持ったレコードが元のデータ数分作成される。例えば、A,B,C の 3 種類のラベルをもつレコードが 10 件ある場合には、30 件に複製され、それぞれ 10 件ずつ A のみ、B のみ、C のみのラベルを持ったレコードがあることとなる。そして、ランダムフォレストで学習を行う際には、これらのレコードの重みは  $1 / n$  に設定される。

表 3.2 ランダムフォレストによる一致率検証結果

ID	性別	年齢	訓練データ内のラベル	推定したラベル	CA (%)
No. 1	男性	69	A,B,C	A	100
No. 2	男性	73	D,E,F,G	D	100
No. 3	男性	66	A,D,H	A	64
No. 4	女性	70	D,I,J	D	91
No. 5	女性	65	D,J,K	J	82
No. 6	女性	68	D,I	D	100
男性全体	-	-	A,B,C,D,E,F,G,H	A または D	64
女性全体	-	-	D,I,J,K	J または D	91

評価においては、医師の診断と脈診機の診断がどれほど一致しているかを示す分類精度指標  $CA$  を式 3.1 により定義する。

$$CA = N_{matched} / N_{total} \quad (3.1)$$

ここで、 $N_{total}$  は、中医学医師の診断に基づいて分類された診断データの総数を表し、 $N_{matched}$  は、脈診機 (pulse diagnosis instrument) による診断が分類結果と一致したテストデータの数を表す。

### 3.2.2 データセット

本実験では、病院に通院している中国の高齢者 6 人を実験協力者として、データを取得した。内訳は、男性 3 人と女性 3 人であり、平均年齢は 68.5 歳であった。彼らは耳鳴りやめまい、胃腸炎などの症状を持っており、その診察のために医師による脈診断を受けていた。その診断の機会に合わせ、同日に脈診機で脈像と脈診機による診断のデータを収集した。また、日常生活におけるヘルスデータもウェアラブルデバイスを用いて取得した。取得したデータは、以下の 3 種類である。

- ヘルスデータ (歩数, 呼吸数, 最低血圧, 最高血圧, 心拍数, 睡眠スコア)
- 脈診機<sup>\*1</sup>が推定した診断結果

<sup>\*1</sup> モデル名: DS01-C, 上海 Daosh Medical Technology Co., Ltd, <http://www.daosh.com/en/product/detail.aspx?id=16>

- 中医師の脈診断結果

ヘルスデータは、37 degree bracele\*<sup>2</sup>を用いて記録され、毎日の夜 8 時前後の最新のデータを 1 日 1 回収集した。病院への通院は週一回であり、そのときに専門医による脈診断を受けた。また、その通院と同じ日に、脈診機による診断推定も行われた。データ取得期間は、2017 年 2 月 8 日から 2017 年 5 月 15 日までであり、早稲田大学倫理委員会の承認を得て取得された (承認番号: 2017-224)。

取得されたヘルスデータと同日に得られた医師の診断を対応させ、ランダムフォレストで学習を行った。訓練データの診断レコードは、1 人あたり 21 件であった。脈診では、1 人の脈が 1 回で複数のカテゴリーに診断される場合がある。したがって、一部の訓練データには複数のラベルがつけられていた。テストデータとして、脈診機から与えられた診断ラベルが 1 つだけのレコード及びその日のヘルスデータを用いた。テストデータの数は 1 人あたり 11 件であった。

### 3.2.3 診断推定の結果と考察

医師の診断と脈診機の診断の一致率の結果を、表 3.2 に示す。この表からわかるように、*CA* はすべてのケースで 60 % を超え、64 % から 100 % までの幅となり、最も低い精度では男性では 64%、女性では 91% であった。これは、一致する確率が 60% 以上であり、ランダムに推定したときよりも高い精度で推定が一致していることがわかる。中医学医師の診断を訓練データとして使用した分類では、*PDI* による診断と近い診断を推定するという傾向があった。また、性別ごとにデータを統合しても、分類精度は男性の最低精度 64% と同じ値であった。

ヘルスデータから推定した医師の診断と脈診機の診断に同じ傾向が見られたということは、ヘルスデータと脈診断に一定の関係があることを示している。これより、脈診機による脈診断推定も、個人のヘルスデータをさらに分析することによって向上できることが期待できる。ただし、この実験においては、訓練データとテストデータの数が少ないという問題があった。そのため、訓練データとテストデータで共通したラベルを十分に用意することが出来ず、複数のラベルを持ったテストデータを分類することは出来なかった。一致率について詳細に検討するためには、十分な数のデータを用意することや、マルチラベル分類を行えるようにするためのアプローチをさらに改善する必要がある。

---

\*<sup>2</sup> 37 degree bracelet, 37 Degree Technology, <http://www.37c.cc/en/index.html> (最終確認日 2018/8/10)

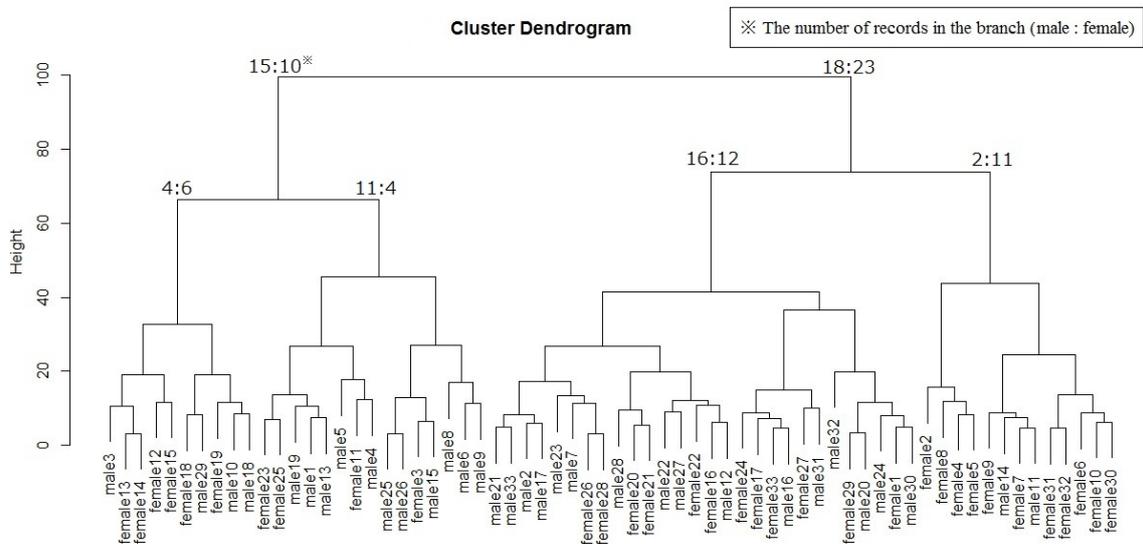


図 3.3 階層クラスター分析によるヘルスデータの分類結果

### 3.2.4 性別に着目したクラスター分析の結果と考察

診断推定では、性別で分けて分析することにより、男女に精度の差が見られた。これは、女性の方が男性よりも脈診断とデータとの関係性が強いいため、良い精度となった可能性がある。もし、男女で異なる傾向があるならば、分類器を性別で分けて作成するなど、性差を考慮した分析が必要となる。

そこで、クラスター分析によって男性と女性のレコードがどのようにクラスターを作成するかについての分析を行った。クラスター分析を実行するため、上述のランダムフォレストによる脈診断推定の実験で用いたデータセットを用いて分析を行った。ヘルスデータに「男性/女性 + データ番号」としてラベルを付け、4つのパラメーター（最高血圧、最低血圧、呼吸数、心拍数）を使用して階層クラスター分析を行った。その分析では各データ間の距離を計算する方法が複数存在するが、この分析では典型的な方法であるワード法 [54] を採用した。

図 3.3 は、階層的クラスター分析の結果を示している。ここにおいて、各階層にある割合の数字は、そのクラスター内に男性と女性のレコードがそれぞれいくつ含まれているかを示しており、左側の数字が男性、右側の数字が女性のレコード数を示している。この図を見ると、上から 2 番目の階層では、左から 2 番目のブランチ（枝）の方が女性よりも男性のレコードが多く、4 番目のブランチの方が男性よりも女性のレコードが多くなっていることがわかる。

性別に着目したクラスター分析では、女性の方が男性よりも分類の精度が高かった。この結果は、こ

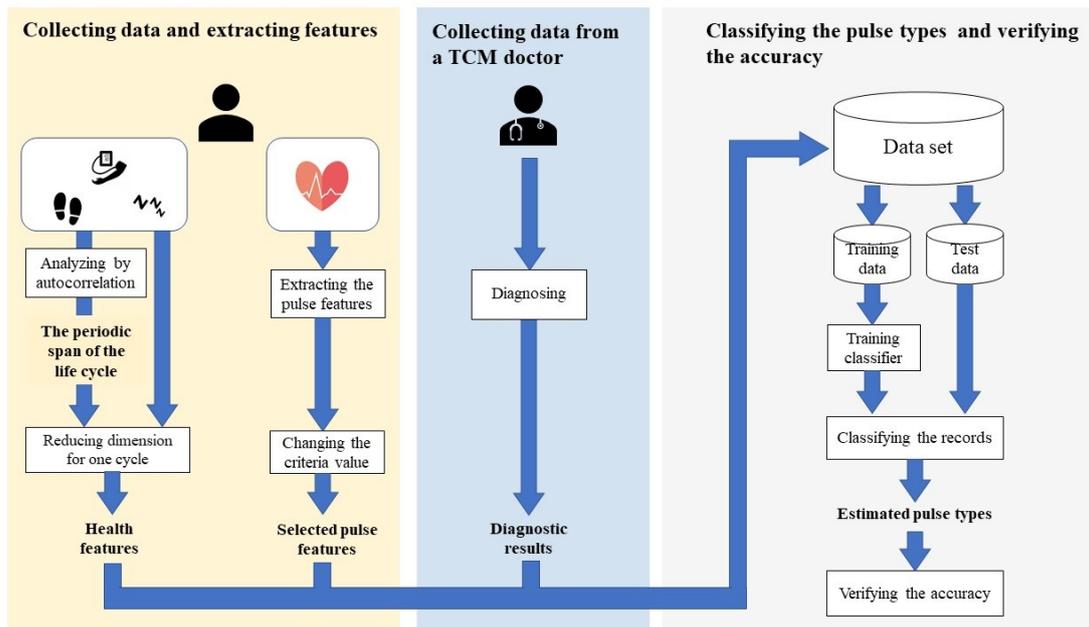


図 3.4 診断推定のための複合的アプローチ

これらのデータに性別ごとに類似した特徴がある可能性を示している。これは、女性のヘルスデータのほうが脈診断に大きく関わっていることや、診断ラベルごとにヘルスデータの傾向が異なっている可能性が考えられる。ただし、一番上の階層では、その差はわずかであった。この分析では明確な結果は得られなかったため、性別による分割が有効であるかについてもデータ数を増やし、実験デザインを改良するなどして今後さらなる検証が必要である。

### 3.3 生活周期を考慮した脈診断推定

上で紹介した実験では、ヘルスデータは脈診断と関係しており、それにより推定精度が向上する可能性があることが示唆された。また、分析の際には、診断日のデータだけでなく日々の活動や生活周期も脈診断推定のために考慮することが望ましい例えば、日々散歩をしている人や決まった曜日に体操がある人などは、脈診断当日以外にも健康に関わる活動を行っており、それが診断結果に影響を与えている可能性がある。これは中医学における人間を総合的に考えるという理念による考え方であり、実際に生活指導や運動指導は中医師によって行われることが一般的である。

診断日だけでなく、近日のヘルスデータまで取り入れようとした場合、どの範囲まで診断日と関係するデータであるかの範囲を決めた上で分析する必要がある。そこで、ここでは生活周期に着目した。生

生活周期は人の生活のリズムであり、学生や企業に勤務する社会人は平日 5 日間、休日 2 日間といった生活周期を持っていることが多い。そこで、ヘルス関連データから生活周期まで分析し、その周期を診断日と関係するデータとして、脈診断推定に取り入れるアプローチを検証した [10]。

生活周期を考慮した脈診断推定のための複合的アプローチを図 3.4 に示す。ここでは、生活周期のスペンを自己相関分析により推定し、診断日からその期間分前までのデータを PCA によって次元圧縮して機械学習に用いる。得られた特徴量は、ヘルス特徴量として、分析に使用される。これにより、脈特徴量とヘルス特徴量を組み合わせることができ、個人の性質まで機械学習に取り入れられることが可能となる。

ヘルスデータでは血圧、心拍数、呼吸数など多くの種類が考えられる。これらをリアルタイムや狭い感覚でデータを取得した場合には、種類や量ともに膨大な量のデータになる可能性がある。そのため、ヘルスデータをそのまま利用して機械学習に適用した場合には、次元の呪いと呼ばれる問題によって推定精度が悪くなる可能性がある。次元の呪いは Bellman [55] によって提唱された用語であり、変数の次元が増えるに従って計算量やパラメータは指数関数的に増大してしまうという問題を指す。それにより、計算を収束させることが不可能になってしまったり精度が大きく低下するなどの問題が発生する。

本研究では、この問題を解決してヘルスデータを従来の脈診断と組み合わせるため、推定した生活周期スペンを用いて次元削減を行う。次元削減とは、変数の数を減らし、データの特徴量を抽出する方法である。それにより、ヘルス特徴量として使用することで、脈診断推定の精度を改善できるかどうかについて検証を行う。

### 3.3.1 自己相関分析に基づく生活周期の推定手法

次元削減を行うには、PCA が代表的な手法として存在する。PCA は、多くの変数を要約して合成変数を作成することができるため、これを用いて診断日から一定期間遡ったデータの範囲で次元削減を行い、ヘルス特徴量を作成する。しかし、PCA を行うためには、どの範囲の過去データを使用するか決めなければならない。そこで、統計的手法である自己相関分析に基づいて生活周期を推定し、それを基に分析する手法を提案し検証を行う。

生活周期のスペンを分析するために、時系列データであるヘルス関連データに対して自己相関分析を行う。この自己相関分析では、データが一定期間前のデータと相関しているかどうかを調べることができる。これにより、歩数の周期が一週間ごとに繰り返されているなど、ヘルス関連データがどのような周期で繰り返されているのかを調査する。

自己相関分析においては、算出された自己相関が有意かどうかを調べるため、Ljung-box 検定 [56] という統計的手法を用いる。このテストでは、統計検定量  $Q$  が式 3.2 によって定義される [56]。

$$Q = n(n+2) \sum_{j=1}^h \frac{\hat{p}_j^2}{n-j} \quad (3.2)$$

ここで、 $n$  はレコード数を示しており、 $\hat{p}_j$  は、期間の長さが  $j$  の場合の自己相関を示し、 $h$  は検証する最大期間の長さを表している。自己相関パラメータ  $\hat{p}_j$  は過去のデータとの相関係数であり、式 3.3 で定義される [57]。

$$p_j = \frac{Cov(y_t, y_{t-j})}{\sqrt{Var(y_t)}\sqrt{Var(y_{t-j})}} \quad (3.3)$$

ここで、 $y_t$  は時間  $t$  での変数  $y$  の値を表し、 $y_{t-j}$  は時間  $t-j$  での変数  $y$  の値を表す。  $Var$  は分散を求める関数であり、 $y_t$  と  $y_{t-j}$  の 2 つの変数に対して使用される。  $Cov$  は、2 つの変数間の共分散を計算するものである。統計検定量  $Q$  に基づいて、 $p$  値が計算され、自己相関が有意かどうかを検証するために使用される。  $p$  値とは、帰無仮説の下で  $Q$  の値が得られる確率を表す。この  $p$  値が特定のしきい値よりも小さい場合、帰無仮説を棄却して有意であると判断することができる。なお、多くの研究では、このしきい値は 5% に設定されている [58]。Ljung-box 検定の帰無仮説は、データに自己相関がないことであり、対立仮説はデータに自己相関があることである。もし  $p$  値が定められた閾値よりも小さければ、帰無仮説を棄却して、検定した期間において自己相関があるとみなすことができ、有意であるということはデータが  $j$  のスパンで自己相関を持っていると見なすことができることを意味する。本研究では、歩数を分析し、 $p$  値が 5% 未満の場合、そのスパンを被験者の生活における生活周期とみなした。歩数を採用した理由は、血圧や心拍数よりもデータ取得の瞬間の状況に左右されにくく、1 日の活動を反映していると考えたためである。得られた生活周期を用いて、診断日からそのスパン 1 周期分前までのデータを診断と関係する日であるとみなし、PCA を実行して次元圧縮によるヘルス特徴量を作成した。

### 3.3.2 データセット

生活周期を考慮した脈診断推定の実験では、健康な中国の高齢者 8 名（平均年齢 70 歳）からデータを収集した。実験協力者の詳細な年齢や性別を表 3.3 に示す。本実験では実験協力者のインフォームドコンセントも得られており、実験計画は早稲田大学倫理委員会に承認されている（承認番号：2017-224）。

実験協力者からは、医師の診断データ、脈診機の診断・脈像データ、ヘルスデータを収集した。得ら

表 3.3 実験協力者の概要

ID	性別	年齢	身長 (cm)	体重 (kg)
A	男性	65	165	60
B	男性	68	168	75
C	男性	68	167	68
D	男性	75	165	65
E	女性	72	160	62
F	女性	76	154	58
G	女性	67	158	60
H	女性	69	160	55

れたデータの詳細は以下のとおりである。

- 医師の診断データ：これは医師による診断データであり，毎週 1 回中医師によって行われた。データ取得期間は，2018 年 3 月 18 日から 2018 年 7 月 1 日までである。この期間中に，すべての被験者から合計で 9 種類の脈が取得された。この診断結果には，1 度の診断で一つの脈状が診断される場合と，複数種類の脈状が記録される場合があった。
- 脈診機診断・脈像データ：上記と同じ期間に，脈診機による診断も行われた。脈診機による診断は医師の診断と同じ日に行われ，それと同時に，脈像から得られる脈特徴量も脈診機によって取得された。この脈診機による診断においても，一度の診断で一つの脈状が診断されるケースと複数の脈状が診断されるケースが存在した。
- ヘルスデータ：これは個人の日々のヘルスデータであり，ブレスレット型のウェアラブルデバイス<sup>\*3</sup>を手首に装着することで取得された。取得されたデータは，歩数 (Step)，睡眠スコア (Sleep score)，最低血圧 (DBP)，最高血圧 (SBP)，心拍数 (Heart rate)，経皮的動脈血酸素飽和度 (SpO2) である。このデータは，2018 年 3 月 15 日から 2018 年 8 月 15 日までであり，一日一回，午後八時付近の最新データを記録した。

以上のデータの件数については，8 人全て合わせて医師の診断及び脈診機診断データが 123 件，ヘルスデータのデータ数が 1,232 件である。

<sup>\*3</sup> モデル名：EHP-A86, Shenzhen OUDMON Technology Co., Ltd, <http://en.oudmon.com/>

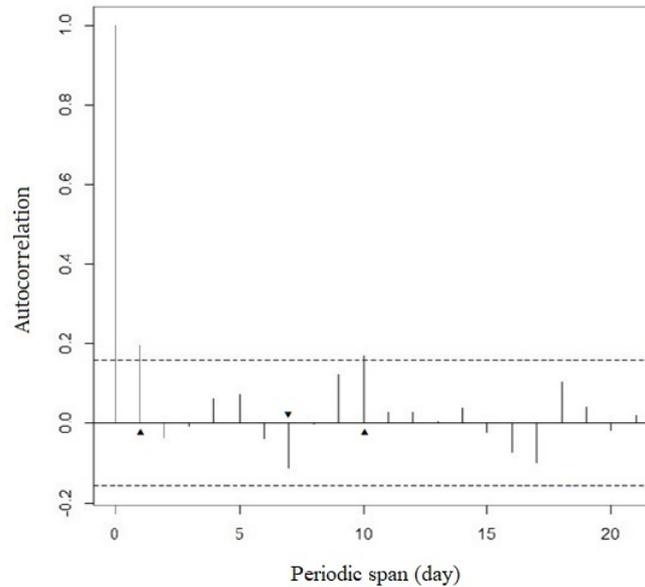


図 3.5 自己相関分析結果の例

### 3.3.3 脈特徴量の抽出と生活周期推定の結果

本実験では脈特徴量として、Li et al. [23] の調査に基づき、 $h3/h1$ ,  $h4/h1$ ,  $w/t$ ,  $w$  の 4 つの指標を用いた。その調査は GBDT による脈診断推定への変数重要度であるため、今回の機械学習にも有効であると考えてこれらの指標を採用した。今回採用した特徴は、上位から順に 1,2,3,5 番目である。4 番目の Rf (Reflection rate) は今回の脈診機では取得できなかったため、その次に高い重要度を持つ  $w$  を採用した。

生活周期は、ヘルス関連データが持つ自己相関を分析することによって推定される。この実験では、通勤や通学、散歩や外出など、歩数に周期が現れやすく周期を明確にできると考え、歩数に対して分析を行った。ヘルス関連データの歩数を用いて、一人の実験協力者 (Subject ID: A) に対して生活周期の推定を行った結果を、図 3.5 に示す。この図では、上点線と下点線を超えたとき、p 値が 5% 以下となったことを示しており、ここではスパンが 1 と 10 のときに点線を超えている。期間の長さが 1 のときは一日のズレ、すなわち前日であることを意味しており、前の日との間に相関があるといえる。期間の長さが 10 の場合は、ある日と 10 日前とのデータに相関があるといえる。このように、各個人の歩数に対して分析を行い、どの期間に周期性があるかについて分析を行った。

図 3.5 の結果では、この実験協力者においては 10 日間の生活周期があると見なすことができる。し

かし、他の7人の実験協力者に対して分析を行ったところ、全てのスパンにおいて有意な自己相関が得られなかった。図 3.5 では、スパンが7の場合、自己相関は2番目に高い値を示している。労働者や学生など多くの人は生活周期として一週間のリズムで活動しており、7日周期が採用されている。そのため、有意な自己相関を得られなかった7人の対象者に対しては、7日の生活周期を設定した。

生活周期を推定した後、得られたスパンを基にヘルスデータを用いて PCA を行った。診断日を起点とし、推定された1周期のスパンまで遡ったヘルスデータをその診断と関係するデータであるとみなし、診断日ごとに次元削減を行った。PCA では、元のデータをどれだけ説明できるかという割合を寄与率として得ることができるため、80%を設定したところ、2つの特徴量がヘルス特徴量として得られた。この80%という数字は、一般的にPCAでの累計寄与率の閾値に用いられている基準である [59]。そして、次元削減によって得られたヘルス特徴量を脈特徴量と組み合わせ、機械学習での説明変数として利用した。

### 3.3.4 診断推定の結果と考察

#### (A) シングルラベル分類の結果と考察

機械学習によって、生活周期を考慮したヘルス特徴量と脈特徴量を学習させることにより、医師の脈診断結果を推定する精度が向上するかどうかについて検証した。ベースラインは、従来行われている脈特徴量のみによる診断推定であり、検証ではこれと本アプローチとの比較を行った。

はじめに、推定する脈状が一種類の場合に、シングルラベル分類としての精度がどのように変化するかを検証した。シングルラベルの分類には、結果を比較するために Li et al. [23] が採用した GBDT を使用した。

Li et al. は4つの脈状を脈特徴量から分類しており、それを参考にしてこの実験でも4種類の脈状を分類した。はじめに、脈状ラベルが1つしかない35レコードがデータセットから抽出された。その後、分類器を構築するため、訓練データとして20レコード、残りの15レコードがテストデータとして使用された。

GBDTによって分類した結果を、表 3.4 に示す。本実験では、精度はベースラインと同じであり、Li et al. [23] による実験の結果よりも低い精度であった。この結果は、データセットのサイズが彼らのものよりも小さいためであると考えられ、生活周期を考慮したアプローチとベースライン間での改善を示すことが出来なかった。

表 3.4 シングルラベル分類での GBDT による分類精度

手法	Accuracy
ベースライン	0.80
本研究	0.80
Li et al. [23]	0.905

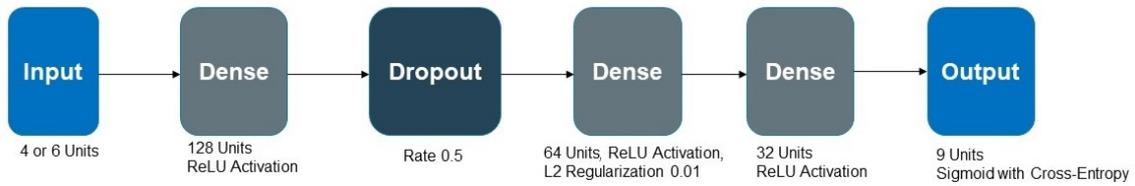


図 3.6 構築したディープラーニングモデル

表 3.5 マルチラベル分類の分類精度

	ベースライン	本研究
	Accuracy (F-measure)	Accuracy (F-measure)
ランダムフォレスト	0.41 (0.50)	0.48 (0.57)
ディープラーニング	0.87 (0.53)	0.88 (0.56)

### (B) マルチラベル分類による分類の結果と考察

次に、複数のラベルを一度に推定するマルチラベル分類を行った。中医師は複数の種類の脈を同時に診断することがあるため、マルチラベル分類はシングルラベルよりも実際の診断に近い分類方法であるといえる。GBDT はマルチラベル分類に対応していないため、ランダムフォレストとディープラーニングの 2 つの方法を使用して、マルチラベル分類の精度を検証した。

ランダムフォレストは決定木に基づくアンサンブル学習法であり、特徴量に基づく分類や回帰に有用であるとして、多くの研究で用いられている [60, 61]。ディープラーニングで構築したモデルを図 3.6 に示す。このモデルにおいて、入力数が 4 のときはベースラインのアプローチによるものであることを意味しており、入力が脈特徴量のみであることを示す。入力数が 6 のときには、脈特徴量に加えて生活周期を考慮したヘルス特徴量を加えられていることを示す。学習におけるエポックの数は 500 を設定し、分類を行った。

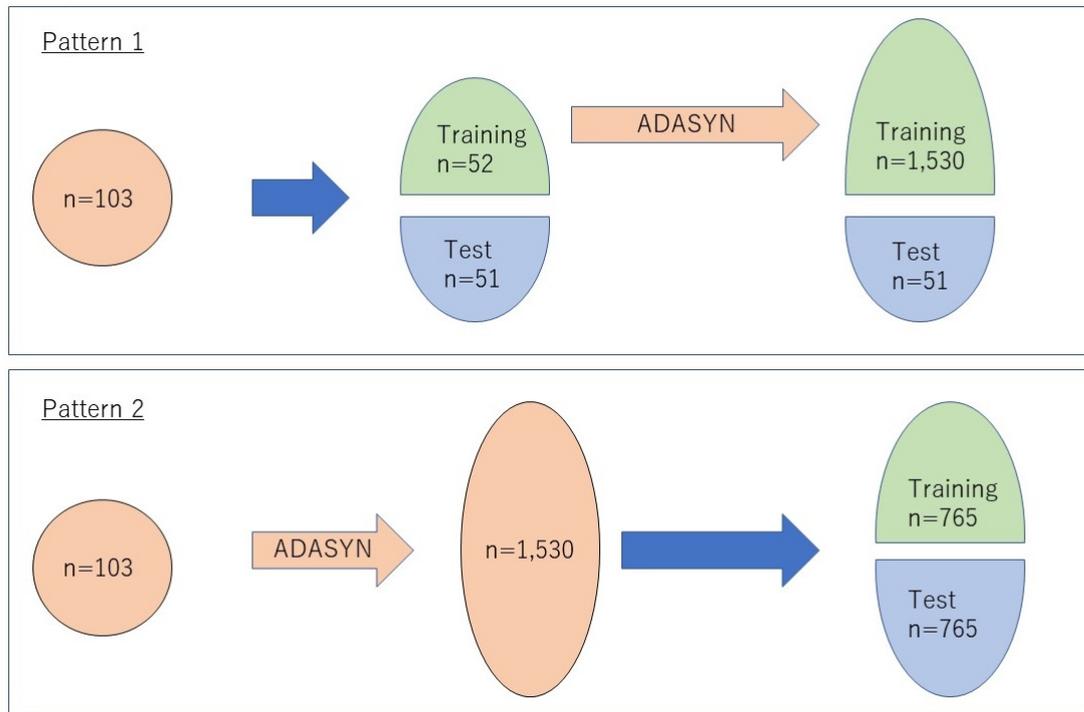


図 3.7 データの分割方法

分類結果を表 3.5 に示す。ここにおける精度 (Accuracy) は、全ての脈状ラベルの個数に対する正解した脈状ラベルの個数を示している。診断の完全一致ではなく個別の脈状に対する一致率を精度としており、一つの脈状が診断されたかされなかったかについてカウントしたものを意味する。

ここで提案したアプローチの精度は、どちらの手法を用いた場合でもベースラインよりも高い精度を得ることが出来た。これは、生活周期を考慮することによって従来の脈診断推定精度を高められることを示している。

しかし、まだデータセットが小さいという問題が残っている。データセットが小さい場合は、作成した分類器が訓練データの偏りに大きく影響を受けることが考えられ、分類器の精度を十分に検証できていないとは言えない。したがって、この結果だけで終わること無く、データセットを大きくしてさらなる検証を行う必要がある。そこで、この問題を解決するために、人工的なデータを作成することによってデータ数を増やす方法を適用した。そのようなデータを人工的に増やす方法はオーバーサンプリングと呼ばれており、複数の方法が存在する。本実験では、ADASYN [62] と呼ばれる手法を適用した。この方法は、畳み込みニューラルネットワークを用いて脈診断を推定した研究 [26] でも使われている手法である。この ADASYN を用いて、さらなる実験を行った。

### (C) ADASYN を適用したマルチラベル分類の結果と考察

実験 1, 2 では, 実験に使用したレコード数は 103 であった. ここでは, ADASYN [62] による人工的データの作成と追加によって, データサイズを大きくしてさらなる実験を行った.

ADASYN は, データ数が少ないために分類器での学習重みが小さくなっているカテゴリのデータ量を増やすことにより, 分類の精度を向上させる. この方法では, 一番データ数が少ないカテゴリの密度分布を使用して, 合成サンプルの数を決定する. 一番データ数が少ないカテゴリの各データについて, KNN によって近傍データを検出し, 少数派カテゴリのデータの周りに多数派カテゴリが多いほど合成サンプルを動的に作成するという方法である. これにより, まわりに多数派カテゴリのデータが多いほど合成サンプルも多く作られ, 周りに多数はカテゴリのデータがない場合にはあまり合成サンプルが作られない. すなわち, カテゴリの境界線近くに多く作成されることとなる.

Peimankar et al. [63] は, ADASYN を用いてオーバーサンプリング後に交差検証を行った. 本実験においても, 彼らの研究を参考にし, データセットの大きさの問題を解決するため, ADASYN を適用した. ADASYN によるオーバーサンプリング後, データセットを訓練データとテストデータに分割した. オーバーサンプリング後のデータセットの大きさは, は 3,060 レコードである. 新しいデータセットでは, 70% を訓練データ, 30% をテストデータとして使用した. このデータセットを用いて, 以下の 6 つの機械学習手法を用いて推定精度を検証した. 具体的な訓練データとテストデータに分割して増やす方法については, 図 3.7 にも示すように, 2 つのパターンをそれぞれ検証した. 以下に, その 2 パターンについての詳細を示す.

- パターン 1: 元のデータセットを, 訓練データとテストデータに分割する. ADASYN によってデータ数が増加するのは訓練データのみである. なお, すべてのラベルが訓練データに少なくとも 1 回含まれるようにレコードが選択され, 残りはランダムに抽出される. 訓練データを抽出して ADASYN を適用した後, それによって分類器を構築してテストデータを分類する.
- パターン 2: ADASYN を, 分割されていないデータセットに適用する. データサイズを増やした後, ランダムに選択されたデータが訓練データとテストデータに分割される.

ADASYN を適用し, マルチラベル分類を行った結果が表 3.6 である. この実験では, ランダムフォレストによる精度は低かったが, ディープラーニングではどのパターンにおいても高い精度で分類を行うことが出来た. パターン 1 でランダムフォレストの精度が低くなっている原因としては, パターン 1 での訓練データの元となるデータ数が少なく, それを基にデータを増加させたためにバイアスが大きく

表 3.6 ADASYN を適用した後のマルチラベル分類精度

		ベースライン	本研究
		Accuracy (F-measure)	Accuracy (F-measure)
ランダムフォレスト	(Pattern 1)	0.32 (0.42)	0.24 (0.32)
ディープラーニング	(Pattern 1)	0.92 (0.70)	0.99 (0.93)
ランダムフォレスト	(Pattern 2)	0.83 (0.87)	0.89 (0.92)
ディープラーニング	(Pattern 2)	0.92 (0.79)	0.99 (0.98)
Li et al. [26]	-	0.95	-

なってしまったことが考えられる。元々の 52 個のデータの偏りが増幅される形でデータ数が増加してしまい、ディープラーニングに比べて単純な仕組みであるランダムフォレストは、上手く分類を行うことが出来なかったと思われる。パターン 2 ではどちらの手法による分類においても精度が向上したが、より高い精度となったのはディープラーニングであった。こちら、ランダムフォレストよりも複雑な仕組みであるディープラーニングのほうが、上手く学習できたことを示唆しており、その有効性を示していると言える。

#### (D) 複数の機械学習手法によるマルチラベル分類の結果と考察

実験 3 にて、ランダムフォレストとディープラーニングについて分類精度を検証し、その比較を行った。ここでは、この結果を基に、さらに複数の手法と比較し、生活周期を考慮した本アプローチの有効性を検証した。用いた手法は、以下の通りである。

- ロジスティック回帰 (Binary relevance)
- サポートベクターマシン (SVM) (Binary relevance)
- ナイーブベイズ (Binary relevance)
- マルチ k-NN
- ランダムフォレスト
- ディープラーニング

このうち、ロジスティック回帰、SVM、ナイーブベイズはシングルラベル分類の手法であり、マルチラベル分類をサポートしていない。そのため、Binary relevance と呼ばれる手法を用いて、これらの

表 3.7 6つの機械学習手法による分類精度比較

Method	ベースライン	本研究
	Accuracy (F-measure)	Accuracy (F-measure)
ロジスティック回帰 (Binary relevance)	13.0 (19.1)	18.0 (27.5)
SVM (Binary relevance)	15.0 (18.0)	47.2 (61.1)
ナイーブベイズ (Binary relevance)	24.3 (61.0)	31.8 (74.5)
マルチ k-NN	98.0 (94.6)	95.1 (91.7)
ランダムフォレスト	95.6 (93.0)	97.8 (93.0)
ディープラーニング	95.5 (89.5)	98.5 (96.2)

手法をマルチラベル分類に適用した。Binary relevance では、シングルラベル分類器であるクラスに分類されるかどうかの二値分類を、クラスの数分を行うことにより、擬似的にマルチラベル分類を行うという手法である [64]。マルチ k-nn は、k 近傍法をマルチラベル分類用に改良したものである [65]。また、ここで用いたディープラーニングのモデルは図 3.6 のモデルである。訓練データとテストデータは、これまでの結果に基づき、実験 3 のパターン 2 でデータ数を増加させたものを使用した。

表 3.7 は、複数の機械学習手法による推定結果を示している。Binary relevance での精度は低かったが、マルチラベル分類法を用いた場合では、どれも 95% 以上の精度が得られた。中でも、ランダムフォレストとディープラーニングで精度が向上した。これは、ヘルス関連データから生活周期を考慮することで精度が向上することを示している。一方、Multi k-NN では精度が改善できなかった。このデータセットでは、近傍法による分類が不適切である可能性や、データセットの増やし方についても関係がある可能性があり、さらなる検証が必要である。

このように、生活周期を自己相関分析によって分析し、更に PCA も適用してヘルス特徴量を抽出することによって、脈診断の推定精度が向上する結果となった。中医学では体全体の状態が脈に現れてきているという考え方も合致するものであり、体質や生活周期を考えることは合理的であると言える。ヘルスデータは体全体の状態と密接に関連するデータであるため、個人が持つ生活周期という観点からそれに対して分析を行う複合的ヘルスデータ分析が有効であると言える。この実験の改善点としては、さらなるデータセットの拡張、舌診や触診といった他の診断結果の考慮、歩数以外の周期性の分析などが考えられる。それらについても、実験デザインの改良やデータの収集など、今後さらに検討する予定である。

## 第4章

# 潜在的要因に着目したヘルスデータ分析 による異常検知

---

この章では、従来の異常検知手法に対して、個人が持つ潜在的要素を分析して組み合わせる複合的アプローチについて述べる。実験では血圧や心拍数が記録されたバイタルサインのデータを用い、従来の異常検知手法に加えて疾病リスクや疲労など人の内部に潜んでおりバイタルサインに影響を与える要因の分析手法を示し、組み合わせる複合的アプローチの有効性を検証する。潜在的要因は統計的手法である構造方程式モデリングによって因子得点として定量化され、得られた因子得点を隠れマルコフモデルの入力とすることによって度合いの推定を行う。推定された度合いごとに異常検知を行うことにより、これまでの手法よりも細かくパラメータを調整でき、異常検知の精度が改善されることを示す。また、ディープラーニングと同様の精度を持ち、さらに結果が得られるプロセスを見える化できる手法であることについても述べる。

---

健康管理は生活において重要であるが、そのうちの一つに病気の早期発見や予防が大きな課題としてあげられる。早めに治療を行うことにより、体の負担軽減、回復の早期化だけでなく医療費の削減や医療資源の節約を行うことができるため、本人だけでなく社会的にも良い効果をもたらす。

本研究では、人が持つ疲労度や内部に潜む病気の進行度に焦点を当て、これまでの異常検知手法と組み合わせる複合的アプローチを提案し検証する。人から得られるバイタルサインは、疲労や病気などの隠れた影響を受けて、得られるものであると考える。また、異常検知においてはディープラーニングが用いられることが多いが、結果が得られる過程がブラックボックスであるという性質上、その

信頼性には限界があると言われている [48]. 医療現場では人の命に関わることもあるため、どのような過程で得られたものかを明確にできないと、責任の問題や医療過誤の問題があるため、信頼性の確保は重要な課題である. したがって、ディープラーニングとほぼ同等の精度を持ち、結果が得られるプロセスが明確で信頼性の高い手法が求められていると言える. そのため、結果が得られるプロセスを明確にでき、潜在的要因を考慮することによって人の状態に合わせたより繊細で精度の高い異常検知手法を行うことを目的として研究を行う.

## 4.1 潜在的要因の定義と定量化

健康異常の検出に関する研究は数多く行われてきている. 例えば、脳波の異常信号は異常検出の主要なテーマの一つである [66, 67]. しかし、これらの研究では、脳波信号を記録するために特殊な装置を用いており、普通の人々が日常的に使用できるものではない. そのため、気軽に異常検知のためのモニタリングをするという事は難しく、機材などを別途用意しなくてはならない. また、これらの研究では、バイタルサインのわずかな変化も重要であるにもかかわらず、病気が発生した時の異常が現れる値そのものに焦点を当てており、他の要因は考慮されていない. 例えば、病気を引き起こす要因であるとか、基礎疾患、他のバイタルサインなどの指標との関係性などである. 健康に関する異常というのは他の指標や要因と関係して引き起こされることが多いため、他の要素を考慮することにより更に精度の高い異常検知が期待できる.

人が病気になるときは、病気を誘発する潜在的要因が存在し、その要因によって人の健康が影響を受けたり病気が発病すると考えられる. 例えば、健康な状態から突然病気になるわけではなく、その人に潜む病気を引き起こす要因が進行することで病気になる. 風邪では抵抗力が徐々に低下したり、他の病気でも血栓の肥大化やがん細胞の増殖など本人には気づかないところで病気は進行している. このような病気の進行は自覚的な症状もなく、さらに主観的な感覚としても現れてこないため、直接分析することが難しい. しかし、病気を引き起こす潜在的要因の進行が健康に影響し、自覚的でなくとも病気の兆候としてヘルスデータに現れている場合がある. このような兆候を素早く検知することができれば、早期の治療や予防が期待できる. 他にも、仕事の負担が多くなったことをきっかけとして、睡眠不足や運動不足などに悩まされるようになり、免疫力が低下して体温の上昇や咳の症状など風邪の症状が現れることが考えられる. この例では、仕事量の負担が睡眠の量に影響を与え、それがさらに免疫力に影響を与えることにより、病気の発生を導いている. これらの要因の影響は、直接観察することが困難であるため、定量化することも難しい. また、間接的かつ暗示的に病気を誘発する要因もあれば、感染した

表 4.1 要因の種類

	直接的	間接的
明示的	(1) 直接的, 明示的	(2) 間接的, 明示的
暗示的	(3) 直接的, 暗示的	(4) 間接的, 暗示的

人と接触してしまったという直接的な要因も考えられる。したがって、そのような潜在的要因から影響を受ける事柄から、その背後にある要因の進行度合いを推定することができれば、その人の環境や状況に合わせて異常検知を行うことが可能となる。

本研究では潜在的要因を考慮した複合的アプローチを提案するが、はじめに潜在的要因について定義する。多くの研究では「要因」という用語を使用しているが、その意味は一貫していない。本研究では、表 4.1 に示すように、直接的、間接的、明示的、および暗示的の 4 つのカテゴリを使用して要因を定義する。この表では、背景が灰色のセル (2)、(3)、および (4) が潜在的要因を示している。各カテゴリは、次のように定義される。

- (1) 直接的とは、結果とのつながりが唯一な要因を表す。
- (2) 間接的とは、結果と直結しないが、2 つ以上の関係性から結果を明らかにできる要因を表す。
- (3) 明示的とは、データから結果を明確に特定できる要因を表す。
- (4) 暗示的とは、データから結果を明確に特定できないが、何らかの暗黙的知識を適用することで結果を明らかにできる要因を表す。

この定義における結果とは、データを分析して得られた現象、事実、結果を指している。上記の定義を適用することにより、それぞれの組み合わせは表 4.1 と対応して以下のように説明することができる。

- (1) 結果との関係性が唯一であり、明確に特定できる要因。病気になった後、それは 1 つの直接的かつ明示的な要因によって引き起こされたことがわかる。例えば、腐った食品を食べたことが原因で食中毒になる。
- (2) 結果と直結しておらず 2 つ以上の関係性があり、データから明確に結果を特定できる要因。病気になった後、2 つ以上の連鎖反応の間接的および明示的な要因によって引き起こされたことがわかる。例えば、インフルエンザの予防接種をしなかった人が感染した人と接触し、インフルエンザにかかる。

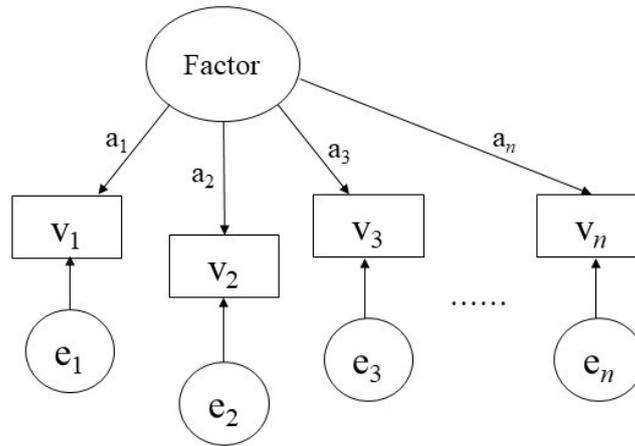


図 4.1 ドメインモデル

- (3) 結果との関係性が唯一であり，何らかの暗黙的知識の適用によって結果を明らかにできる要因．例えば，喫煙者などの特定のタイプの人々が肺がんになる傾向がある場合，その要因は，その人が肺がんになる前の喫煙頻度と喫煙期間から推定される．
- (4) 結果と直結しないが 2 つ以上の関係性があり，暗黙的知識が結果を明らかにするために必要となる要因．例えば，特定のタイプの人々がライフスタイル関連の病気にかかる傾向があり，その人が仕事で忙しく，睡眠不足の場合，要因は就業時間と睡眠時間から推定される．

本研究のアプローチでは，(2)，(3)，及び (4) を潜在的要因として扱い，分析を行う．

潜在的要因を分析する際，要因の度合いを直接計測することは難しい．潜在的要因が互いに関連しているという複雑なモデルを検証する場合，構造方程式モデリング (SEM : Structural Equation Modeling) がよく使用される．May et al. [68] は，SEM を運転体験とユーザー情報に適用し，自動運転への態度に影響を与える潜在的要因を含むモデルを検証した．Khakurel et al. [69] は，ウェアラビリティ，ウェアラブルデバイスのデザイン，およびその他のいくつかの要因が使用の意図に影響しているかどうかを検証した．Tong et al. [70] は，SEM を用いてインターネット中毒者の心理状態と生活の質との関係を調査した．Demerouti et al. [71] は，潜在的な要因として看護師の仕事の需要と仕事のリソースを想定し，燃え尽き症候群と人生の満足度についてのモデルを検証した．これらを参考にし，本研究では統計的手法である SEM を採用し，変数の組み合わせから潜在的要因の度合いを定量化する．

SEM では，図 4.1 に示すようなドメインモデルが必要である．ドメインモデルは，潜在変数から影響を受ける顕在変数の組み合わせが表されている．潜在的要因から影響を受ける変数をモデル内で表現

することにより、すでに観測されている変数の増減から潜在的要因の度合いを推定する。例えば、心拍や呼吸数が高くなっているときは、病気の危険度が増して健康状態が悪くなっている可能性がある。これは、危険度が増加したことにより、心拍や呼吸数が高くなったと考えられる。

SEM はパス解析や因子分析を含んだ分析手法であり、潜在的要因の度合いや変数への影響の重みは因子分析の手法を基にして推定される。SEM を適用すると潜在的要因の度合いは潜在因子の因子得点として計算され、これは共分散を基に計算が行われる。例えば、2つのデータセットが  $X, Y$  が  $(X, Y) = (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  として得られたとき、 $X$  と  $Y$  の共分散は式 (4.1) によって求めることができる [72].

$$\begin{aligned} Cov(X, Y) &= E[x_i y_i] \\ &= E[(X - \bar{X})(Y - \bar{Y})] \end{aligned} \quad (4.1)$$

ここで、 $E(z)$  は  $z$  の期待値を示しており、 $\bar{X}$  と  $\bar{Y}$  はそれぞれ  $X$  と  $Y$  の平均値を表す。

たとえば、潜在的要因が3つの明示的な変数  $(v_1, v_2, v_3)$  に影響する場合、変数  $v_1, v_2, v_3$  と因子得点  $F = \{f_1, f_2, \dots, f_n\}$  は以下の式で表すことができる [72].

$$v_1 = a_1 f_i + e_1 \quad (4.2)$$

$$v_2 = a_2 f_i + e_2 \quad (4.3)$$

$$v_3 = a_3 f_i + e_3 \quad (4.4)$$

ここで、 $f_i$  は潜在因子得点を示す。豊田 [73] に示されている定義により、 $f_i$  の分散は1と見なされ、 $f_i$  の平均は0と見なされる。 $a_i$  は、相関の程度を示す回帰係数である。 $e_i$  は独立した定数値であり、潜在的要因以外の影響を示す値（誤差）である。SEM の前提条件として、 $e_i$  は独立という仮定であるため、他の変数の影響を受けない。そのため、 $e_i$  には他の変数との共分散はなく、 $e_i$  の分散は1と見なされ、平均は0と見なされる。

2つの変数  $(v_1, v_2)$  を考えたとき、 $v_1$  と  $v_2$  の共分散は次のように計算できる [73].

$$\begin{aligned} Cov(v_1, v_2) &= E[(a_1 f_i + e_1)(a_2 f_i + e_2)] \\ &= a_1 a_2 E[f_i^2] + a_1 E[f_i e_2] + a_2 E[f_i e_1] + E[e_1 e_2] \\ &= a_1 a_2 \end{aligned} \quad (4.5)$$

ここで、 $E[f_i^2]$  は1に等しい。なぜならば、 $E[f_i^2]$  は式 (4.1) によって  $E[(f_i - E[f_i])^2]$  と表すことができ、 $f_i$  の分散を示しているためである [73]。また、 $f_i$  と  $e_1$ 、 $f_i$  と  $e_2$ 、 $e_1$  と  $e_2$  の共分散は独立しており0と定義されるため、 $E[f_i e_1]$ 、 $E[f_i e_2]$ 、 $E[e_1 e_2]$  は0となる。

$Cov(v_2, v_3)$  と  $Cov(v_3, v_1)$  は,  $Cov(v_1, v_2)$  と同じ方法で計算できる. それにより,  $a_1a_2$ ,  $a_2a_3$ , および  $a_3a_1$  の値を計算することができる.  $v_1$ ,  $v_2$ ,  $v_3$  は明示的な変数 (歩数, 心拍数など) であるため,  $Cov(v_1, v_2)$ ,  $Cov(v_2, v_3)$ , および  $Cov(v_3, v_1)$  は式 (4.5) によって計算することができる. その結果, 次の連立方程式を得ることができる [73].

$$a_1a_2 = C_1 \quad (4.6)$$

$$a_2a_3 = C_2 \quad (4.7)$$

$$a_3a_1 = C_3 \quad (4.8)$$

ここで,  $C_1$ ,  $C_2$ , および  $C_3$  は, それぞれ  $Cov(v_1, v_2)$ ,  $Cov(v_2, v_3)$ , および  $Cov(v_3, v_1)$  の値を表している.

式 (4.6) を解くことにより,  $a_1$ ,  $a_2$ , および  $a_3$  の値を得ることができるため, これらの値を使用して因子分析を行うことができる. 因子分析では,  $f_i$  の値を推定するために重回帰分析が実行される [74].  $F$  は式 4.9 によって推定できる [74].

$$\hat{F} = XR^{-1}Af \quad (4.9)$$

ここで  $V$  は説明変数の平均偏差行列,  $Af$  は因子負荷行列であり影響の重みを示す. 上記の例では,  $Af$  は以下のように表すことができる [74].

$$Af = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

ここで,  $R$  は説明変数間の相関行列を表す.  $v_1$ ,  $v_2$ , および  $v_3$  の相関行列は以下のように表すことができる [74].

$$R = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{pmatrix}$$

ここで,  $r_{12}$ ,  $r_{23}$ ,  $r_{13}$  はそれぞれ  $(v_1, v_2)$ ,  $(v_2, v_3)$ ,  $(v_3, v_1)$  の相関変数を表している. 以上より,  $F$  を潜在的要因の度合い,  $Af$  を変数に対する影響の重みとして推定することができる.

SEM では, ドメインモデルがデータにどの程度適応しているかを適合度指標によって評価でき, モデルの当てはまりの良さを示す指標が複数の提案されている. それらを使用することにより, そのモデルが上手く事象をモデル化できているかどうかについて判定できる. 3つの指標, AIC (赤池情報量基準) [75], RMSEA (二乗平均平方根誤差) [76], CFI (比較適合指数) [77] は, モデルのデータに対す

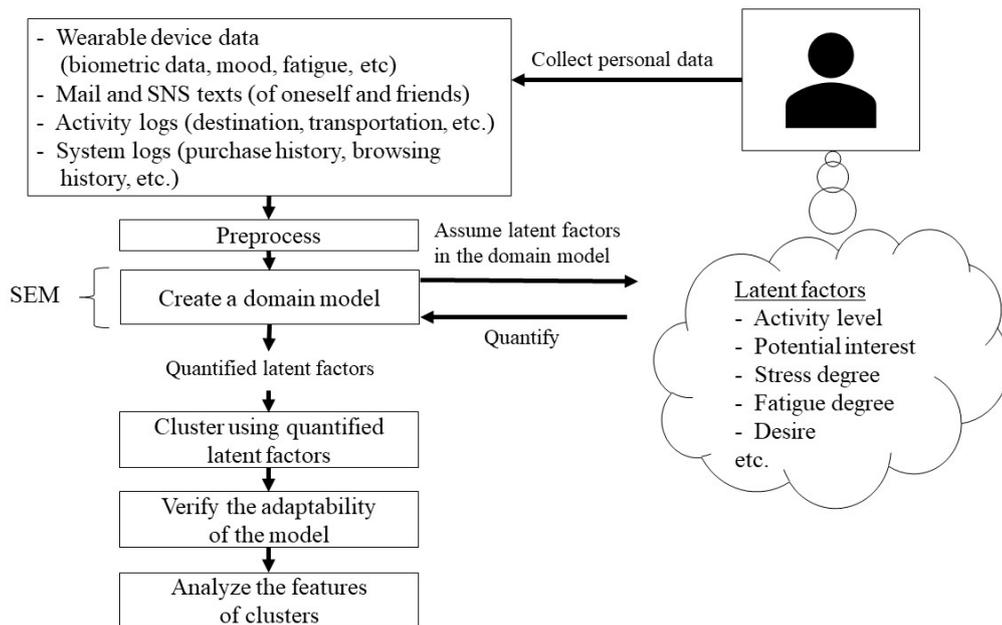


図 4.2 クラスター分析を用いた潜在的要因の分析概要

る適応性を示すものである。AIC は、小さいほどモデルがデータによりよく適合していることを示す。RMSEA が 0.05 未満の場合では、モデルは適切なモデルであると見なすことができる。CFI が 0.9 より大きい場合では、モデルはデータに適合していると判断できる。

## 4.2 構造方程式モデリングとクラスター分析を用いた潜在的要因定量化

はじめに、上記で紹介した SEM の有効性について検証するため、人が持つ活動を引き起こす活動因子を潜在的要因とみなし、その度合い定量化によってどのように人の行動が変化するのかを検証した [12]。本実験のアプローチを図 4.2 に示す。ここでは、歩数や心拍が活動因子によってどのような傾向を示すのか、定量化した活動因子からヘルスデータをクラスタリングすることによってそのクラスタ別の特徴を分析した。また、その有効性を検証するため、クラスタリングされたデータ別にモデルの適合度指標を検証した。適合度指標が高い値であれば、従来のラベルよりも効果的に日々のデータをクラスタリングをおこなうことができると考えられる。検証では、クラスタリングした本アプローチとは別に、クラスタリングを行わずに分析したものと外出ラベルによってクラスタリングしたものを用意して比較した。

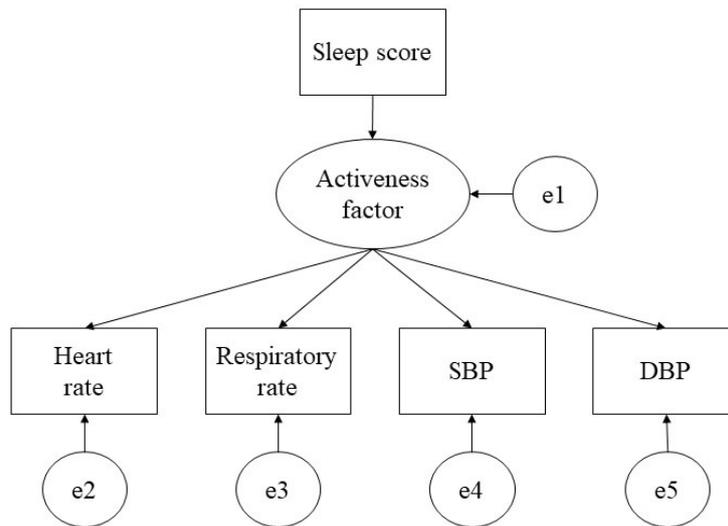


図 4.3 活動因子を分析するためのドメインモデル

#### 4.2.1 ドメインモデルの作成と構造方程式モデリングの適用

次に、前処理を行ったデータを用いて活動度モデルを個人データ分析に適用した。このプロセスでは、毎日の活動因子の度合いが SEM によって定量化される。

潜在的要因を定量化するためには、潜在的要因と顕在変数との関係性を表すモデルが必要である。そこで、個人から得られたデータを用いてドメインモデルとして活動モデルを構築した。図 4.3 は、構築された活動モデルを示している。活動因子は毎日のデータに対して計算され、活動による影響を受けるヘルスデータ：心拍数 [78]，呼吸数 [79]，収縮期血圧（SBP），拡張期血圧（DBP） [80]，との関係性を示した。睡眠スコアは、ウェアラブルデバイスによって計算される夜間の睡眠の質を表すスコアであり、活動因子に影響を与えるものである。

#### 4.2.2 活動因子に基づいたクラスター分析

SEM を用いて定量化された活動因子をもとに、クラスター分析を行う。クラスター分析には、代表的な手法である k-means 法を採用する。SEM および k-means 法の適用においては、データ分析用の

オープンソースソフトウェアである R<sup>\*1</sup>を使用した。

k-means 法は、Guo の研究 [19] でも個人データをクラスター化するために使用される手法であり、教師なし学習でのクラスターリングを行うことができる。k-means 法では最初にクラスター数  $k$  を決める必要があるため、3つのクラスター（活動因子の度合い：高，中，低）を想定して、パラメーター  $k$  の値を 3 に設定した。クラスターリングの結果を図 4.4 に示す。この図より、一番左のクラスターには高い活動因子スコア、3 番目のクラスターには低い活動因子スコアが分布している事がわかる。

検証では、3つの方法（クラスターリングなし、ラベルに基づくクラスターリング、活動因子に基づくクラスターリング）を比較した。それにより、モデルに対する適合度を検証することで、潜在的要因の定量化によってクラスター別に傾向が違っているかを検証した。クラスターリングなしでは、すべてのデータを一つのクラスターと考え、すべてのデータを使用して適合度を算出する。ラベルに基づく分類では、ある 1 日にどここの場所で活動していたかを示すラベルを付与し、そのラベルごとに適合度を算出するものである。ラベルは、Google map からの行動データを基に、自宅、大学、外出の 3つのラベルのうちどれかを判定し、1 日毎に付与した。

### 4.2.3 データセット

表 4.2 は、取得したデータの種類と頻度を示す。ヘルス関連データを取得するため、手首に装着するウェアラブルデバイスである 37 degree ブレスレットを使用した。さらに、スマートフォンに内蔵されている Google マップのタイムラインと呼ばれる行動記録を参照し、1 日の目的地と交通手段を取得した。実験協力者は男性であり、健康な大学院生、27 歳であった彼はほぼ毎日大学に行き、研究活動を行っていた。データ取得期間は、2017 年 8 月 30 日から 2018 年 2 月 13 日であった。なお、この実験協力者は研究者本人であるため、倫理委員会での倫理審査は行っていない。

次に、データの前処理を行った。得られたヘルスデータを、10時から18時の時間までに区切り、その時間帯に対して平均値を計算した。このように時間を分割する理由は、実験協力者が通常 10 時に家を出て 19 時頃に家に帰ることがヒアリングによって判明しており、主に活動する時間帯とみなせるためである。ヘルスデータの 10時から18時の時間帯の平均を計算した後、各項目の分散が 1 で平均が 0 になるように、データの正規化を行った。最後に、欠損値を含む 1 日のデータを削除し、最終的なデータ数として 132 レコードを得た。

---

\*1 R Development Core Team, <http://cran.r-project.org/>

表 4.2 ウェアラブルデバイスとスマートフォンによって取得されたデータ

Device or App	Attributes	Interval
37 degree bracelet	Step	Once an hour
	Moving distance	Once an hour
	Consumed calories	Once an hour
	Systolic blood pressure	Once an hour
	Diastolic blood pressure	Once an hour
	Heart rate	Once an hour
	Respiratory rate	Once an hour
	Sleep score	Once a day
	Mood	Once a day
	Fatigue	Once a day
Google map	Destination	Real time
	Transportation	Real time

#### 4.2.4 実験の結果と考察

表 4.3 に、モデル適合度を検証した結果を示す。ヘルスデータをラベルで分類した場合、自宅ラベルでは、データ数が少なすぎて NaN で示されているようにモデルが収束しなかった。また、AIC を比較したときも、大学ラベルでの値はクラスタリングなしのものに近く、大きな値となった。これは、行動履歴だけでクラスタリングした場合にはヘルスデータの特徴を上手く分析することができないことを示している。一方、活動因子に基づいて分類した場合は、全てのケースでモデルとして有効であると判断された。これにより、潜在的要因を使用する場合のモデル適合性は、分類されない場合やラベルで分類された場合よりも良い値であり、上手くモデルに当てはまっているといえる。

また、クラスターごとにヘルスデータがどのように分布しているかについても調査を行った。図 4.5 は、クラスターごとのヘルスデータの分布を示している。睡眠スコア、血圧、呼吸数はクラスターごとに大きな違いは見当たらないが、心拍数は活動因子と強く相関していることがわかる。このことより、活動度スコアは心拍数に影響を与えることがわかるが、活動することにより一日の平均心拍数が変動することは合理的であると考えられる。

活動因子を用いた実験では、3 つの改善点があることが明らかとなった。1 つ目は、ドメインモデル

表 4.3 モデル適合度の比較

	分類なし	ラベルによる分類			活動因子に基づく分類		
		大学	自宅	外出	High	Middle	Low
データ数	132	116	2	14	26	69	37
AIC	1498.2	1290.4	N/A	179.0	290.0	609.4	367.8
RMSEA	0.000	0.000	N/A	0.000	0.000	0.000	0.000
CFI	1.000	1.000	N/A	1.000	1.000	1.000	1.000

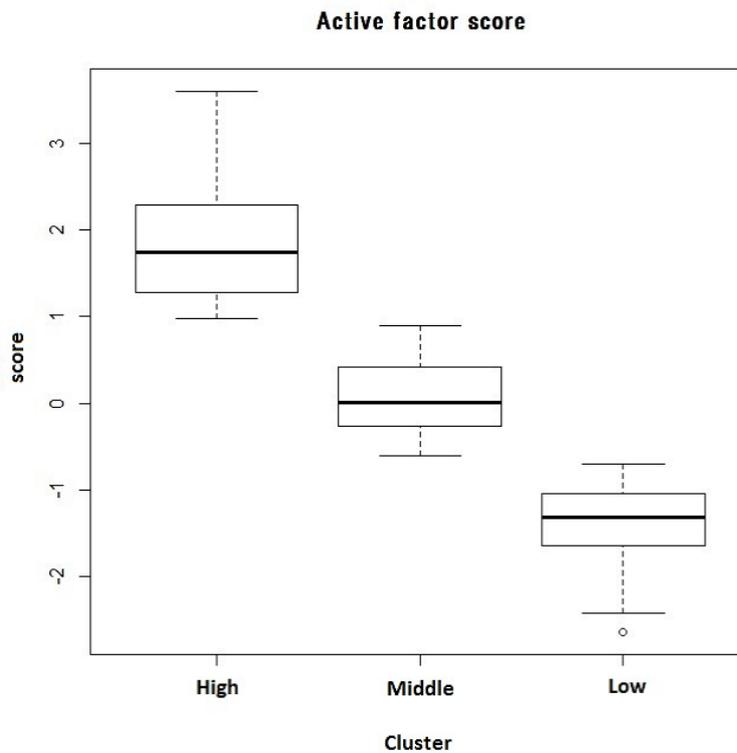


図 4.4 クラスターごとの活動因子の度合いの分布

についてである。この実験では、活動モデルを構築し、ヘルスデータに対するモデルの適合度を検証した。ただし、実験結果は使用するモデルに大きく依存する。複雑なモデルを使用する場合や他の潜在的要因に対してモデルを構築する場合、その構築方法やモデルの有効性を慎重に検証する必要がある。

2つ目は、ヘルスデータの前処理方法である。この分析では、1日の時間のうち、10時から18時のヘルスデータを平均して使用した。しかし、分単位や一日ごとに違う時間で区切るなど、より細かい分

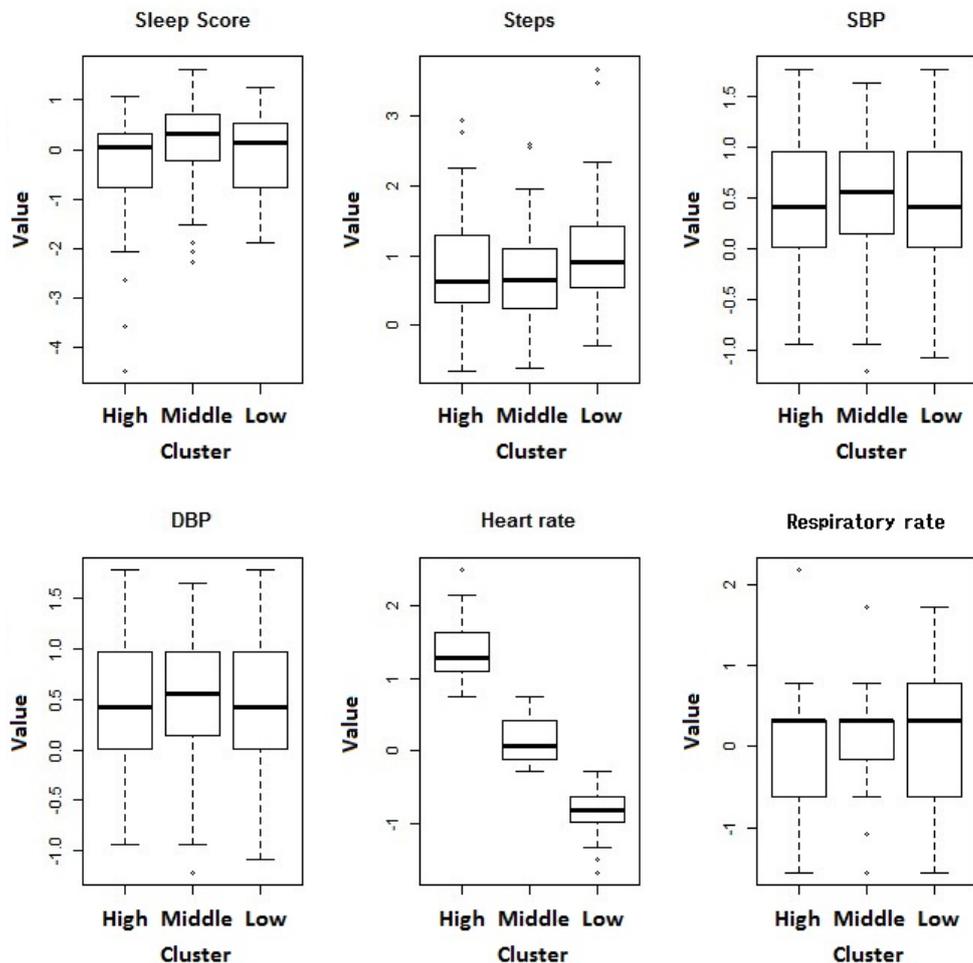


図 4.5 クラスタごとのヘルスデータの分布

割方法も可能である。一律に時間を分割するだけでなく、その個人の活動時間に合わせて区切ることに  
より、さらにその個人の活動についての分析が可能となる。どのように区切れば一日の活動が反映され  
ているかについては、さらなる検証が必要である。

3つ目はデータセットのサイズについてである。この実験では、一人の大学院生からのデータを用い  
ており、その個人に合わせてラベルを設定した。他の人の場合ではより細かく分けることも可能である  
と考えられ、モデルの適合度も変化すると考えられる。より多くの人からデータを収集し、ラベルもそ  
の個人に合わせてなどさらなる検証が必要である。

### 4.3 疲労度合い推定による異常検知

潜在的要因を活動因子として定量化する実験においては、従来のラベルよりも効果的にクラスタリングすることができ、モデルに対する適合度が高いことが明らかとなった。しかし、その定量化においては時系列の影響は考慮されていなかった。より個人のことを詳細に分析するためには、時系列の影響を考慮することが望ましい。例えば、前日に激しく運動した日は次の日に少し休むことが多いなど、我々の生活は独立しておらず過去の影響も受けている。そのため、定量化した潜在的要因を更に分析し、時系列の影響を考慮した度合いの推定を行う。それにより、よりきめ細やかな異常検知を目的として実験を行う。

ここでは、SEM で得られた潜在的要因の度合いを時系列データとして分析し、異常検知を行う手法を検証する。本実験の概要を図 4.6 に示す。ここでは時系列の影響を考慮するため、計算された潜在的要因の度合いに対し、隠れマルコフモデル (HMM: Hidden Markov Model) を適用して状態推定を行う。先述したように、SEM で計算された度合いは時系列データではないため、過去のデータによる影響が考慮されていない。HMM は、一つ前のデータの状態から現在の状態を推定することが可能であるため、その影響を考慮して現在の状態を推定することができる。これにより、HMM の入力として定量化された潜在的要因を使用し、推定で得られた状態を度合いとして用いる。

この研究では、人が持つ疲労度合いを潜在的要因と考え、その度合い別にヘルスデータの異常検知を行うことにより、より詳細に異常検知を行うことを目的として実験を行った。異常検知で用いる複合的アプローチを図 4.6 に示す。ここでははじめにユーザーからヘルスデータを取得し、SEM や HMM によって潜在的要因の度合いを推定した後、その度合い別に異常検知が実行される。それにより、検出された異常があれば本人や病院、家族に通知を送ることが期待できる。

#### 4.3.1 隠れマルコフモデルと HotellingT2 法による異常検知手法

HMM は確率モデルの 1 つである。HMM では、データは現在の状態に基づいて取得され、次の状態は確率的に現在の状態によって決定されると想定されている [81]。このつながりはマルコフ連鎖と呼ばれており、現在の状態から次の状態を時系列データとして推定する。HMM では、データの確率分布は隠れたモデルの状態に依存しており、すでに得られている変数はその隠れモデル状態に従って出現したものであると考える。

本研究での SEM と HMM の関係を図 4.7 に示す。ここでは、SEM によって定量化された潜在的要

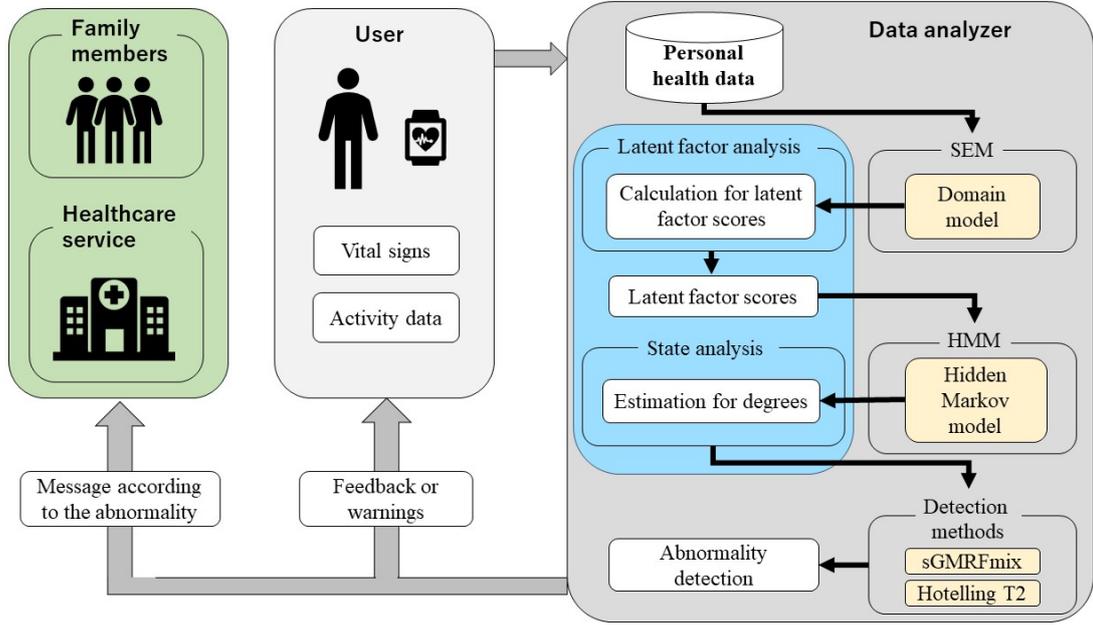


図 4.6 異常検知のための複合的アプローチ

因の度合いを、HMM の入力として用い、そこから隠れた状態の推定を行う。これにより、SEM のみではできなかった時系列を考慮した度合い推定が可能となる。

HMM では、状態推定は以下の手順で行われる。はじめに、状態の数  $S = \{S_1, S_2, \dots, S_M\}$  が決定される。状態の数  $M$  はパラメータとして事前に設定されて与えられる。次に、SEM によって定量化された潜在因子  $F = \{f_1, f_2, \dots, f_n\}$  について、状態が推定される。パラメータ  $\lambda$  は、式 (4.10) で表される [81].

$$\lambda = (Atp, B, \Pi) \quad (4.10)$$

ここで、 $Atp$  は遷移確率分布であり、状態  $S_i$  から状態  $S_j$  に遷移する確率を表す。 $Atp$  は、式 (4.11) で表すことができる [81].

$$Atp = \{a_{ij} = Pr(q_{t+1} = S_j | q_t = S_i)\}, 1 \leq i, j \leq n \quad (4.11)$$

ここで、 $B$  は観測シンボル確率分布である、これは、観測データ  $V_k$  が状態  $S_i$  から取得される確率を表す。 $B$  は、式 (4.12) で求めることができる [81].

$$B = \{b_j(V_k) = Pr(F_{t,k} | q_t = S_j)\} \quad (4.12)$$

また、 $\Pi$  は初期状態分布である。これは、 $t$  が 1 に等しいときに状態が  $S_i$  になる確率を表す。 $\Pi$  は、

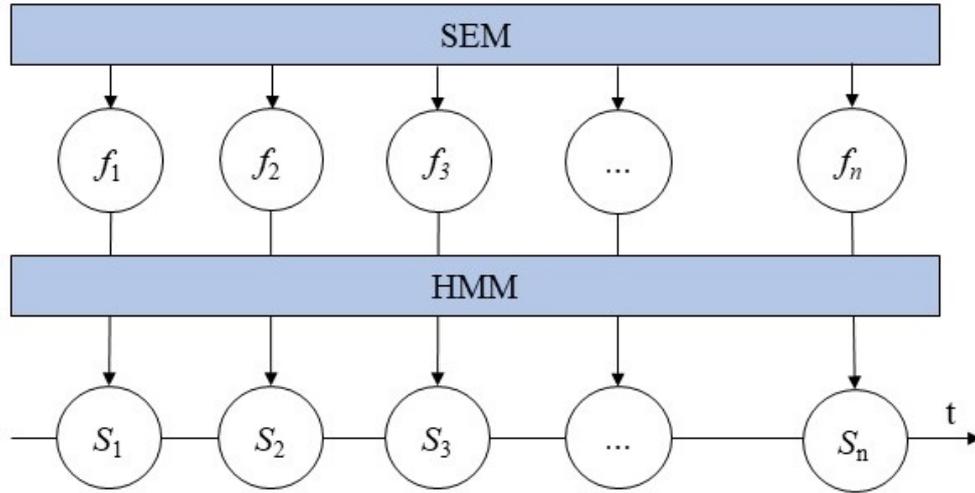


図 4.7 SEM と HMM の関係性

式 (4.13) として求めることができる [81].

$$\Pi = \{\pi_i = Pr(q_1 = S_i)\}, 1 \leq i \leq n \quad (4.13)$$

これら隠れモデルのパラメータを求めるため、Baum-Welch アルゴリズム [82] が使用される。このアルゴリズムでは、最尤推定のための手法の一つである期待値最大化法 (expectation-maximization algorithm) [83] をもとに推定を行っている。

パラメータを推定した後、ビタビアルゴリズム [84] によって状態の遷移が推定される。ビタビアルゴリズムは、動的計画法と分割統治法によってビタビパスを計算するアルゴリズムであり、それにより状態遷移のシーケンスを推定することが可能である。ビタビパスとは一連の状態遷移であり、 $Q = \{q_1, q_2, \dots, q_n\}$  として表すことができる。ここで、 $q_i$  は各レコード  $F_i$  の状態を示す。このアルゴリズムを適用することにより、隠れた状態とその遷移を推定できる。

HMM で状態推定を行った後、状態別に異常検知を行う。異常検出を実行するには、代表的な手法として HotellingT2 法 [39] がある。この手法は統計的手法の 1 つであり、Ma and Lee の研究 [40] で採用されている手法である。

この理論では、異常度  $a$  はデータの分散と平均をもとに計算され、各データに対して求められる。例えば、1 つの変数としてデータ  $x_1, x_2, \dots, x_n$  が得られた場合、 $x_i$  の  $a$  は式 (4.14) に定義される。

$$a(x_i) = \left( \frac{x_i - \mu}{\sigma} \right)^2 \quad (4.14)$$

ここで、 $\mu$  は  $x_i$  の平均であり、 $\sigma$  は  $x_i$  の標本分散を表す。

Hotelling の理論では、 $a$  の分布はカイ 2 乗分布に従うことが証明されている [39]。カイ 2 乗分布は確率密度分布であるため、 $a(x_i)$  を求めれば、その値以上が得られる確率を計算することができる。したがって、任意の確率を設定し、得られた値が異常かどうかを確率によって判定することで、異常検出を行うことができる。

### 4.3.2 データセット

本実験では、ウェアラブルデバイスから得られる疲労度と、本研究で提案する疲労度合いの推定手法を比較し、それぞれ HotellingT2 法で異常検知をした場合に結果がどのように異なるかを検証した。実験協力者は高齢者 18 名（男性 8 名、女性 10 名）であり、彼らのヘルスデータをブレスレット型のウェアラブルデバイスを用いて取得した。なお、この実験は早稲田大学の倫理審査委員会の承認を受けている（承認番号：2017-224）。平均年齢は 63.4 歳であり、取得したデータの種類は、歩数、収縮期血圧（SBP）、拡張期血圧（DBP）、心拍数、呼吸数、気分、疲労度であった。ウェアラブルデバイスでは、20 時時点での最新のヘルスデータと、17 時時点でのデバイスが推定した疲労度を収集した。収集期間は 2017 年 1 月 29 日から 2017 年 3 月 15 日までである。

### 4.3.3 疲労度合い推定の結果

SEM では、各データを正規化する必要がある。そこで、各バイタル信号とステップ数を 0~1 の間で正規化した。図 4.8 に示す疲労度を推定するためのドメインモデルでは、前日の歩数と睡眠の質が疲労因子に影響を与え、SBP と心拍数が疲労因子に影響することを示している。このモデルは、病院のガイドラインや医学的知識をもとに作成された。

高齢者ごとに毎日の疲労因子を定量化した後、HMM を用いて疲労度合いを推定する。疲労度合いはウェアラブルデバイスの分類では 4 つの状態に分かれているため、HMM でも 4 つの状態としてパラメータを与えて疲労状態を推定した。デバイスのラベルに対応するように、疲労因子の平均値が最も大きいクラスターを重度疲労状態、2 番目に大きいクラスターを中等度疲労状態、3 番目に大きいクラスターを軽度疲労状態、4 番目に大きいクラスターを非疲労状態とした。異常検知では、各状態において、Hotelling の理論に基づいて異常値を検出し、異常検知を行った。

Hotelling の異常検知手法で異常値を検出するためには、閾値を設定する必要がある。本実験では、HotellingT2 法で異常検知を行っている関連研究 [40] を参考に、閾値を 0.95 と 0.99 に設定し、それぞれの場合での異常検知について検証した。

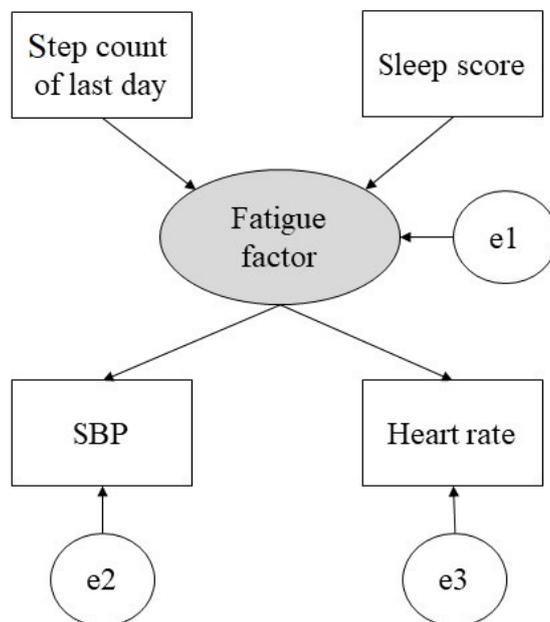


図 4.8 疲労度の要因モデル

表 4.4 異常検知の結果 (男性 A)

	クラスタリングなし						ラベルによる分類						HMMでの疲労度合いによる分類					
	Threshold = 0.95			Threshold = 0.99			Threshold = 0.95			Threshold = 0.99			Threshold = 0.95			Threshold = 0.99		
	SBP	DBP	Heart rate															
疲労度なし	異常データ件数	2	6	7	0	5	5	0	0	0	0	0	0	0	0	0	0	0
	異常データの平均値	159	108.66	66.42	N/A	107.6	51.4	N/A	N/A	N/A	N/A	N/A	N/A	N/A	139	N/A	68	N/A
	状態内の平均値	131.87	89.95	81.46	131.87	89.95	81.46	N/A	N/A	N/A	N/A	N/A	N/A	N/A	121.93	86.9	84.74	121.93
軽度の疲労度	異常データ件数	N/A	N/A	N/A	N/A	N/A	N/A	1	1	1	0	1	1	0	1	1	0	0
	異常データの平均値	N/A	N/A	N/A	N/A	N/A	N/A	161	114	51	N/A	114	41	N/A	71	N/A	N/A	N/A
	状態内の平均値	N/A	N/A	N/A	N/A	N/A	N/A	134.25	90.25	79.79	137.5	90.25	79.79	137.5	87.64	79.79	140.15	88.94
中程度の疲労度	異常データ件数	N/A	N/A	N/A	N/A	N/A	N/A	0	5	6	0	5	3	1	1	1	0	1
	異常データの平均値	N/A	N/A	N/A	N/A	N/A	N/A	N/A	107.6	72.2	N/A	107.6	51	N/A	101	51	N/A	N/A
	状態内の平均値	N/A	N/A	N/A	N/A	N/A	N/A	131.41	90	82.3	131.41	90	82.30	140.15	88.94	78.52	140.15	
重度の疲労度	異常データ件数	N/A	N/A	N/A	N/A	N/A	N/A	0	0	0	0	0	0	1	1	0	0	0
	異常データの平均値	N/A	N/A	N/A	N/A	64	N/A	N/A	N/A									
	状態内の平均値	N/A	N/A	N/A	138.06	95.83	77.77	138.06										

表 4.5 異常検知の結果 (女性 E)

	クラスタリングなし						ラベルによる分類						HMM での疲労度合いによる分類					
	Threshold = 0.95			Threshold = 0.99			Threshold = 0.95			Threshold = 0.99			Threshold = 0.95			Threshold = 0.99		
	SBP	DBP	Heart rate	SBP	DBP	Heart rate	SBP	DBP	Heart rate	SBP	DBP	Heart rate	SBP	DBP	Heart rate	SBP	DBP	Heart rate
疲労度なし	異常データ件数	3	5	9	2	0	2	0	0	0	0	0	0	4	1	0	0	0
	異常データの平均値	152.66	89.8	79.22	154	N/A	51	N/A	N/A	N/A	N/A	N/A	N/A	105.25	101	N/A	N/A	N/A
	状態内の平均値	121.3	86.11	77.31	121.3	86.11	77.31	117.14	84.14	73.57	117.14	84.14	73.57	118.23	81	118.23	81	70.52
軽度の疲労度	異常データ件数	N/A	N/A	N/A	N/A	N/A	N/A	3	2	4	0	0	2	1	2	0	0	0
	異常データの平均値	N/A	N/A	N/A	N/A	N/A	N/A	100	67	68.75	N/A	N/A	76	131.5	101	N/A	N/A	N/A
	状態内の平均値	N/A	N/A	N/A	N/A	N/A	N/A	122.3	88.15	77.71	122.3	88.15	77.71	114.56	86.86	114.56	86.86	76.65
中程度の疲労度	異常データ件数	N/A	N/A	N/A	N/A	N/A	N/A	3	3	3	0	0	0	4	1	0	0	0
	異常データの平均値	N/A	N/A	N/A	N/A	N/A	N/A	152.66	106.33	69	N/A	N/A	N/A	117	N/A	N/A	N/A	N/A
	状態内の平均値	N/A	N/A	N/A	N/A	N/A	N/A	121.46	84.19	78.09	121.46	84.19	78.09	124.61	82.04	124.61	82.04	81.28
重度の疲労度	異常データ件数	N/A	N/A	N/A	N/A	N/A	N/A	0	0	0	0	0	0	0	1	0	0	0
	異常データの平均値	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	状態内の平均値	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	119.86	84.4	119.86	84.4	94.28

表 4.6 男性 A のドメインモデル適合度

指標	結果	適合基準
GFI	0.93	0.9 より大きい
CFI	0.90	0.9 より大きい
SRMR	0.08	0.1 より小さい
RMSEA	0.06	0.1 より小さい

表 4.7 女性 B のドメインモデル適合度

指標	結果	適合基準
GFI	0.97	0.9 より大きい
CFI	0.95	0.9 より大きい
SRMR	0.01	0.1 より小さい
RMSEA	0.04	0.1 より小さい

#### 4.3.4 実験の結果と考察

18 人の高齢者のデータに疲労モデルを適用した結果、モデルが収束せず、8 人はエラーとなり計算不能であった。これは、各データ間の相関関係が矛盾していたためであると考えられる。疲労モデルが一致できた 10 人の高齢者のうち、男性は 4 人 (ID: A, B, C, D)、女性は 6 人 (ID: E, F, G, H, I, J) であった。

そのうち、67 歳の男性 A と 67 歳の女性 E の結果を表 4.4 と表 4.5 に示す。2 つの表において、疲労度の推定を行わなかった場合、その結果は非疲労状態の欄に示した。また、N/A は、該当するデータがなかったために計算されなかったことを示している。また、男性 A と女性 E データに対するモデルの適合度を、表 4.6 と表 4.7 に示す。

一般に、疲労の度合いが高くなると収縮期血圧や心拍数が上昇することが知られている。表 4.4 にもあるように、HMM で推定した疲労度が高くなると、収縮期血圧の平均値が上昇する傾向がみられ、一致していると言える。デバイスのラベルで集約した場合には、疲労度が中等度・軽度に集中しているため、このような傾向は見られなかった。また、表 4.5 では、HMM で推定した疲労度が高くなるにつれ

表 4.8 ウェアラブルデバイスと HMM による疲労度推定数の比較 (男性 A)

		HMM によって推定された疲労度				合計
		疲労度なし	軽度の疲労度	中程度の疲労度	重度の疲労度	
デバイスにより 推定された 疲労度	疲労度なし	0	0	0	0	0
	軽度の疲労度	8	2	7	7	24
	中程度の疲労度	35	12	12	23	82
	重度の疲労度	0	0	0	1	1
合計		43	14	19	31	107

て心拍数が増加する傾向が見られる。この傾向はラベルによる集約よりも明らかであり、潜在的要因としての疲労度を推定することで、上手くデータをまとめることができていると言える。

閾値についてであるが、Hotelling の理論では、結果を確認しながら閾値を調整することが望ましいとされている。結果からも見られるように、HMM で疲労度を推定したところ、0.99 の閾値ではほとんど異常値が検出されなかった。したがって、最初の設定値は 0.95 にした方が良いと思われる。

この実験では、改善または解決すべき問題点が 3 つ見つかった。1 つ目はドメインモデルである。この実験では、8 人分のデータにおいて、変数間の相関が一貫していなかったため、収束しなかった。そのため、疲労度を推定するドメインモデルを改善し、データについてもさらなる収集をして検証する必要がある。

2 つ目は、異常の検出方法についてである。このアプローチで用いた異常検知手法では、平均値とデータとの距離に基づいて異常を検出した。しかし、異常値が平均値より大きいのか小さいかということや、他の変数との関係性は考慮されていない。例えば、最低血圧だけが低下したが最高血圧はそのままである場合や、どちらとも低下した場合や上昇した場合など、異常は単独の変数で見ただけでなく、複数変数を考慮して異常の種類まで検出することが望ましい。それにより、さらに細かく異常を検知することが期待できるため、さらなる検討が求められる。

3 つ目は、状態推定の精度である。この実験では、表 4.8 と表 4.9 にあるように、デバイスによって疲労度が推定された場合に比べ、HMM による疲労度合いの推定を行ったときはその件数がより均等になる傾向があった。本手法による推定はバランスが取れているように見えるが、その精度については今後の検討が必要である。状態推定の精度を検証するために、すでに異常が記録されているデータセットや他の異常検知研究で使用されたオープンデータを用い、さらに検証することが望ましい。

表 4.9 ウェアラブルデバイスと HMM による疲労度推定数の比較 (女性 E)

		HMM によって推定された疲労度				合計
		疲労度なし	軽度の疲労度	中程度の疲労度	重度の疲労度	
デバイスにより 推定された 疲労度	疲労度なし	6	5	3	0	14
	軽度の疲労度	12	14	26	0	52
	中程度の疲労度	20	6	8	7	41
	重度の疲労度	0	0	0	0	0
合計		38	25	37	7	107

#### 4.4 疾病リスク度合い推定に基づく異常検知アプローチの検証

上記の実験では、潜在的要因に基づく異常検知の複合的アプローチを提案した。しかし、使用したアルゴリズムは単純なものであり、異常を正しく検知できているかも検証できなかった。そのため、この研究では上記の研究を発展させ、異常検知の際に他の変数との関係性も考慮できる手法を採用し、さらに他の研究でも用いられているデータセットを使ってさらに検証した。

実験の流れは図 4.6 と同じであり、SEM と HMM によって潜在的要因の度合いを推定し、その度合別に異常検知を行う。潜在的要因としては疾病リスクを想定し、HMM を使用してリスクの程度を推定した後、その度合いごとに異常が検出される。異常を検出するため、ここでは Ide et al. [85] が提案した sGMRFmix アルゴリズムによる検出手法を採用した。この手法は、教師あり機械学習の一種であり、他の変数との関係性を踏まえた異常スコアを計算することで異常検知を行うことができる。この方法では、変分ベイズを使用し、変数に対する重みとガウス・マルコフ確率場のパラメーターが推定される。これらのパラメーターを推定した後、異常スコアがその負の条件付き対数尤度 [85] に基づいて計算される。異常スコアがしきい値より高い場合、異常と見なす事ができる。この手法では、変数ごとの異常スコアを計算することができるため、どの変数に異常が発生したかを識別も可能である。

##### 4.4.1 疾病リスク度合い別の sGMRFmix アルゴリズムによる異常検知

はじめに、統計処理ソフトウェア R の lavaan ライブラリを用いて、SEM を適用するためのドメインモデルを作成した。病気を引き起こすリスクを潜在的要因として推定し、図 4.9 に示すモデルを構築した。脳損傷の場合、病状が変化すると血圧と心拍数が変化することが知られている [86]。また、脳損傷のガイドラインにおいても、血圧と心拍数のモニタリングが必要であることが示されている [87]。さ

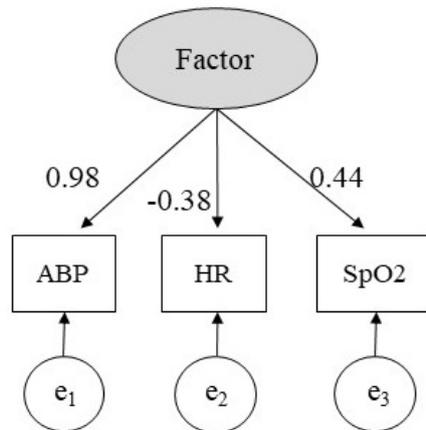


図 4.9 疾病リスクを潜在的要因としたドメインモデル

表 4.10 疾病リスク定量化のためのドメインモデル適合度

指標	結果	適合基準
GFI	0.99	0.9 より大きい
CFI	0.96	0.9 より大きい
SRMR	0.03	0.1 より小さい
RMSEA	0.11	0.1 より小さい

らに、医療現場では、SpO2のモニタリングも行われている [88]。これらの調査をもとに、潜在因子が ABP, HR, SpO2 に影響を与えるドメインモデルを作成した、

図 4.9 に示されている Factor から各バイタルサインに伸びている矢印上の数字は、SEM によって得られた疾病リスクが顕在変数に与える影響の重みを示す。潜在因子スコアと血圧の間には正の相関があると考えられる。さらに、HR は潜在因子スコアとの間に弱い負の相関があると考えられる。この図より、リスクが高くなると血圧が高くなること、また、リスクが低いほど心拍数は高くなるが血圧ほどの変動は少ないことがわかる。そのため、この患者の場合、疾病リスクは主に血圧に関係していることが考えられる。

図 4.10 は、潜在因子を定量化したスコアを 1 分ごとに平均化し、折れ線グラフで示したものである。ここで、図に示された潜在因子スコアは、疾病リスクとみなすことができるため、スコアが変化すると、バイタルサインに影響が出て、疾病リスクが変化することを意味する。リスクが 0 から大きく離れているときは、バイタルサインが大きく変化していることを示している。

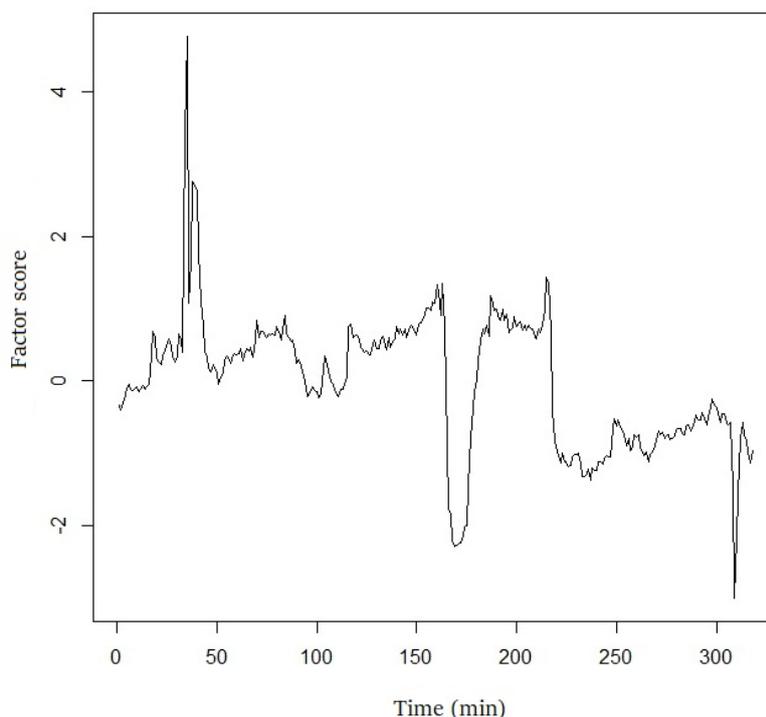


図 4.10 因子得点の推移

表 4.11 実験で利用した異常データ

No.	レコード番号	対象のバイタルサイン	挿入された異常
1	1000-1099	HR	その度合の中の最小値から 20 低い値
2	2000-2099	ABP, SBP, DBP	各バイタルサインのその度合の中での最大値
3	3000-3099	SpO2	その度合の中の最小値から 10 低い値
4	4000-4099	SBP, DBP, HR, PULSE	各バイタルサインのその度合の中での最大値

データに対するドメインモデルの適合度の指標として、ここでは GFI(Goodness of Fit Index ) [89], CFI (Comparative Fit Index)[90], SRMR (Standardized Root Mean square Residual) [77], RMSEA(Root Mean Square Error of Approximation) [91]) を採用した。その結果を表 4.10 に示す。そのうち RMSEA は 0.11 であり、ここでは基準を満たしていない。しかし、RMSEA は許容できる閾値に近く、他の複数の指標も適合度を満たしている。したがって、この指標についても問題はないと判断し、本研究では図 4.9 に示すモデルを採用した。

sGMRFmix アルゴリズムによる異常検知では、レコード  $Rec_i$  の異常度  $a_i(Rec)$  は式 (4.15) で計算される [85].

$$a_i(Rec) = -\ln \sum_{k=1}^K g_k^i(Rec) \mathcal{N}(Rec_i | u_i^k, w_i^k) \quad (4.15)$$

ここで、 $K$  は状態の数であり、アルゴリズムによって決定される。 $g_k^i(Rec)$  は変数の重みである。 $k$  はモードの数、 $i$  は変数の種類である。異常検出は、計算された各レコードの  $a_i(Rec)$  にしきい値を設定することで実行される。

$g_k^i(Rec)$  は、式 (4.16) で表すことができる [85].

$$g_k^i(Rec) \approx \frac{\bar{\theta}_k^i \mathcal{N}(Rec_i | u_i^k, w_i^k)}{\sum_{l=1}^K \bar{\theta}_l \mathcal{N}(Rec_i | u_i^l, w_i^l)} \quad (4.16)$$

ここにおいて、 $\bar{\theta}_k^i$  が計算できれば、 $g_k^i(Rec)$  を得ることができる。 $\bar{\theta}_k^i$  を計算するため、Ide et al. によって提案された sGMRFmix アルゴリズムが使用される。このアルゴリズムでは、sGMRFmix アルゴリズムはパラメータとして  $\mathcal{D}$ ,  $\rho$ , および  $\alpha$  を用いて実行される。 $\mathcal{D}$  はデータセット、 $\rho$  はペナルティの重み、 $\alpha$  はモード  $k$  の混合重みである。モデルの出現確率がゼロでない場合、 $\alpha$  は 1 に設定され、それ以外の場合、 $\alpha$  はゼロに設定される。以上のパラメータを推定することにより、異常度  $a_i(Rec)$  を求めることが出来、それを用いて異常検知が行われる。

#### 4.4.2 データセット

実験では、MIMIC データベース\*2 として公開されている患者のモニタリングデータセットの一つを使用した。このデータベースでは複数のデータセットが公開されているが、関連研究と結果の比較を行うため、Amor et al. [41] の研究で使用されている No. 221 のデータセットを使用した。No. 221 のデータセットでは、約 24 時間の脳外傷女性患者から得られたバイタルサインが記録されている。記録されているバイタルサインは、平均血圧 (ABP)、収縮期血圧 (SBP)、拡張期血圧 (DBP)、酸素飽和度 (SpO2)、呼吸数 (RR)、心拍数 (HR)、および脈拍数 (Pulse) であり、サンプリングレートは、約 1 秒であった。レコード数は 83,363 で、[41] は最初の 20,000 レコードを使用した。彼らの研究と結果を比較するため、この研究でも最初の 20,000 レコードを使用した。その中にはいくつかの欠損値があり、欠損値があるレコードを削除したところ、最終的に 19,951 レコードとなった。この内、機械学習の訓練データとして、19,951 レコードのはじめの 900 レコードを訓練データとして用い、19,051 レコード

\*2 <https://archive.physionet.org/mimic2/> (最終確認日 2020/12/16)

表 4.12 ベースラインとなる異常検知手法と本アプローチとの比較

手法	リスク度合い推定	異常検知手法
PCA ベースの手法	-	Squared Prediction Error に基づく [41]
ディープラーニング	-	オートエンコーダー [92]
度合い推定なし	-	sGMRFmix アルゴリズム [85]
本研究	○	sGMRFmix アルゴリズム [85]

をテストデータとして使用した。

#### 4.4.3 実験データの準備と評価指標

異常検知の精度を評価するための手法として、人工的にデータを作成し、挿入する方法がある [41]。ここでは、Amor et al. [41] と同じデータセットを利用し、その結果を比較するため、同じ方法によって実験用のデータセットを作成した。具体的には、健康の異常を示すデータとして、表 4.11 にあるように異常データを挿入した。これらのデータについて、リスク度合い推定アプローチでは推定された度合い別での挿入し、度合いの推定がない場合ではそのまま挿入した。そして、これらを検出できるかを検証した。

ベースラインとしては、3 種類のアプローチによる異常検知を行い、精度を比較した。PCA を用いた研究 [41]、疾病リスクを推定しない場合、ディープラーニングを用いた場合である。それぞれの違いをまとめたものを、表 4.12 に示す。

これらについての詳細は、以下の通りである。

1. PCA ベースのアプローチ [41]: このアプローチは、先行研究によるものであり、同じデータセットを用いて異常検出を行っている。この先行研究では、異常なヘルスデータが挿入され、PCA を用いて異常度を算出し、異常検知を行った。なお、異常検知のアルゴリズムは、この実験による手法とは異なったものである。
2. ディープラーニング: このアプローチでは、すべてのレコードを一つの度合いとみなし、表 4.11 に示すように 400 件の異常レコードを挿入した。ディープラーニングモデルには、オートエンコーダーを用いた。このモデルでは、学習データに異常が含まれていなくても、異常の程度を計算することが可能である。訓練データを用いてモデルを学習し、テストデータの異常度を算出した。オートエンコーダーのハイパーパラメータについては、損失関数をバイナリクロスエントロ

ピーとし、エポック数を 100 とした。また、異常検知の閾値については F1-score が最も高くなるよう調整し、閾値を設定した。

3. 度合い推定なし: このアプローチでは、全てのレコードを一つの度合いとみなし、表 4.11 に示すように、400 件の異常レコードをテストデータに挿入した。InjectedAnomaly に示すように、試験データに 400 件の異常レコードを挿入した。異常検出の方法は、Ide et al. [85] による手法である。閾値については、検知の F1-score が最も高くなる閾値を設定した。

評価では、異常検出の性能を代表的な指標である Accuracy, Recall, FPR (False positive rate) で評価した。関連研究では Recall と FPR を使用して評価を行っていたため、3 つのベースラインとの比較ではその 2 つを採用した。また、5 交差検証においてはその 2 つに加え Accuracy を採用し、精度を検証した。

Accuracy は、式 (4.17) で定義される。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (4.17)$$

ここで、TP は True Positive, TN は True Negative, FP は False Positive, FN は False Negative を表す。本研究では、True は正しい検出を意味し、False は誤った検出を意味する。Positive は異常であることを示し、Negative は非異常であることを意味する。したがって、TP は正しく検出された異常の数を示し、TN は正しく分類された非異常の数を示す。FP は誤って判定された異常（実際には正常）の数、FN は誤って分類された非異常（検出できなかった異常）の数である。したがって Accuracy とは、本研究でどれだけ異常が正しく検出されたかを示す指標である。

Recall は、式 (4.18) で求められる。

$$Recall = \frac{TP}{TP + FN} \quad (4.18)$$

Recall は、どれだけ真の異常が検出されたかを示す指標である。

FPR は、式 (4.19) で定義される。

$$FPR = \frac{FP}{FP + TN} \quad (4.19)$$

FPR は、異常ではないものをどの程度誤って異常と分類したかを示す指標である。

#### 4.4.4 実験の結果と考察

3 つのベースラインと比較した結果を表 4.13 に示す。また、上記の 3 つのベースラインに対する検証とは別に、度合い推定なしについては 5 交差検証も行った。その結果を表 4.14 に示す。

表 4.13 異常検知の結果

手法	番号	Recall (%)	FPR (%)	異常検知のための閾値
PCA ベース [41]	Average	98.02	0.7	Algorithm based on Jackson and Mudholkar approximation
	No. 1	100	18.0	Anomaly >0.3
ディープラーニング	No. 2	100	0.001	Anomaly >0.8
	No. 3	100	0	Anomaly >0.3
	No. 4	100	0	Anomaly >1.1
	Average	100	4.5	-
	No. 1	100	0.8	HR = NA
度合い推定なし	No. 2	100	0.3	DBP >200
	No. 3	100	0	SpO2 >NA
	No. 4	100	0.3	SBP >20
	Average	100	1.25	-
	No. 1	100	0	HR = NA
本研究 (高リスク)	No. 2	100	0.04	DBP >100
	No. 3	100	0	SpO2 = NA
	No. 4	100	0	SBP = NA
	Average	100	0.01	-
	No. 1	100	0	HR = NA
本研究 (中リスク)	No. 2	90	0.1	DBP >20
	No. 3	100	0	SpO2 = NA
	No. 4	100	0.1	DBP >20
	Average	97.5	0.05	-
	No. 1	100	0	HR >100
本研究 (低リスク)	No. 2	95	0.1	DBP >20
	No. 3	100	0	SpO2 = NA
	No. 4	100	0.1	DBP >20
	Average	98.75	0.05	-
	各度合いの平均	Average	98.75	0.186

リスク度合いを推定したアプローチでは、Recall が 98.75%、FPR が 0.186% と、関連研究よりも良い結果を得た。ディープラーニングと度合い推定なしは全て 100% の Recall であったが、FPR が高い数値であった。FPR が高いということは偽陽性が高く、正常を間違えて異常として検出してしまうことを示している。複合的アプローチによる異常検知では、Accuracy は 100% に届かなかったがそれに

表 4.14 5 交差検証による検証結果

手法	番号	Accuracy (%)	Recall (%)	FPR (%)	異常検知のための閾値
度合い推定なし	No. 1	98.3	96.5	0	HR: NA
	No. 2	75.2	65.0	0.3	DBP > 200
	No. 3	100	100	0	SpO2: NA
	No. 4	68.7	61.0	0.3	SBP > 20
	Average	85.6	80.6	0.1	-
本研究（高リスク）	No. 1	100	100	0	HR: NA
	No. 2	99.5	96.0	0.4	DBP > 100
	No. 3	100	100	0	SpO2: NA
	No. 4	100	100	0	SBP: NA
	Average	99.9	99.0	0.1	-
本研究（中リスク）	No. 1	100	100	0	HR: NA
	No. 2	99.7	99.0	0.02	DBP > 20
	No. 3	100	100	0	SpO2: NA
	No. 4	100	100	0	DBP > 20
	Average	99.9	99.7	0.005	-
本研究（低リスク）	No. 1	100	100	0	HR > 100
	No. 2	98.0	89.3	0.1	DBP > 20
	No. 3	100	100	0	SpO2: NA
	No. 4	99.1	95.0	0	DBP > 20
	Average	99.3	96.1	0.03	-
各度合いの平均	99.7	98.3	0.045	-	

近く、また、FPR は大きく低くすることができた。

さらに、5 交差検証においてリスク度合い推定なしと比較した場合においても、Accuracy, Recall, FPR は、疾病リスクの度合い別に異常検知をしたほうが良い結果となった。

本研究では、バイタルサインと疾病リスクという潜在的要因の関係性からリスク度を推定し、度合い別に異常度を計算して閾値を設定することで異常検出性能を改善した。リスク度が高いほど閾値を大きく設定することで、より高い精度と低い FPR を得ることができた。また、5 交差検証においても性能が向上した。これは、疾病リスク度合いを推定したことによってより細かく異常検知の閾値を設定することができたため、細かい異常まで検知することができたことを示している。従来の異常検知では閾値を一律にする必要があったが、度合い推定をすることにより、リスクが高いときと低いときで分けるこ

とが必要であることが有効であると言える。

ドメインモデルを作成する際、疾病リスクから影響を受ける変数を特定する必要がある。脳外傷の場合、測定・監視すべき指標は広く検討されており、医療ガイドラインにも記載されている。ベースラインとして用いた関連研究においては、血圧や脳波などの指標が異常であるかどうかを直接判断しているが、人の健康状態や疾病リスクの度合いまでは考慮されていない。そのため、異常に対する閾値は一律となり、個人に合わせた異常検知までは行えていない。この実験では、病気を誘発する潜在因子を考慮して度合いを推定することにより、人から得られるバイタルサインに対して感度の高い異常検出が可能であることを示した。

この手法の大きな利点は、目に見える形で異常を検出できることに加え、データセットに依存しないことである。この手法はディープラーニングとほぼ同等の精度を持っており、さらに偽陽性が低い。従来のディープラーニングの欠点であった結果が得られるプロセスを可視化でき、疾病リスクの度合いや推移をグラフとして表すことができた。これにより、どのように結果が得られたかというプロセスの信頼性が向上され、より医療現場での使用に適した手法であると言える。

また、本研究では、個人のヘルスデータを用いた異常検出に着目し、1人の被験者のデータセットを用いて提案手法を評価し、同じデータセットを用いたベースラインとの比較を行った。それにより、本手法はデータセットが特定のものでなくとも健康異常の検出に有効であることが示された。定量化された因子得点から疾病リスクの度合いを推定することで、度合いに応じた異常検知閾値を設定することが可能であるといえる。将来的には、システムの構築を行い、サーバやクラウド上でデータ解析を実行することで、リアルタイムの異常検知が期待できる。それにより、バイタルサインを常にチェックすることで異常を早期に発見することができ、病気の早期発見だけでなく、予防にも貢献することができる。

本アプローチの制限と改善点をまとめると、以下が挙げられる。まず、ドメインモデルは1つの疾患のみを対象に設計されているため、複数の疾患を同時に検出することは考えられていない。ドメインモデルはその名の通りドメイン依存型であり、疾患とそのバイタルサインとの関係についての調査や医学的知見に基づいている。したがって、このようなドメインモデルを作成するには、その疾患に適した調査や専門知識が必要となる。ドメインモデルを作成する際には、モデルの適応性や合理性に加えて、モデルに含まれる明示的な変数（潜伏因子の影響を受けるバイタルサイン）をどのように決定するかを慎重に検討する必要がある。探索的にモデルを作成し、適合度が良いモデルについて合理性を確認するなど、モデルの作り方については課題が残されていると言える。

Pavel et al. [93]によると、人間の生活をモデリングする際には、バイタルサインなどの生のデータを扱うだけでなく、センサーなどからのアクティビティの変容の認識、健康行動の変容モデルが必要

である。本研究ではバイタルサインとその症状に関わる知識を基にドメインモデルを作成したが、生活に関わるイベントや体への負荷がかかるイベントもモデルに入れることにより、異常検知精度を更に改善できる可能性がある。その個人がどのように生活しているのかをセンサーで取得し、スマートフォン等の通話記録や行動履歴からパターンを抽出することにより、個人の生活モデルを構築できることが考えられる。センサーが見えないところに設置されたスマートホームの普及によって、より多くの要素を取り入れたモデル構築が見込める。さらに、Ismail et al. [94] で提案されている健康知識モデルの構築方法も参考になると考えている。これらの研究や構想を基に、ドメインモデルの構築方法をさらに改善・検証して行く予定である。また、比較実験では、複数のバイタルサインを組み合わせると1つの異常を検出する閾値を設定することも可能であるが、簡単のために、1つの異常検出につき1つのバイタルサインに閾値を設定した。例えば、ハイリスク度の異常スコアの閾値設定については、DBPの異常スコアの閾値をFPRを低下させるように設定した。しかし、表 4.13 に示すように、Abnormality No.2 では、リスク度が高くなるにつれて DBP の閾値が大きくなった。迅速な検出のためには、リスク度が高いほどより厳しい閾値を設定することが望ましい。この課題を解決するために、より感度の高い検出方法をさらに提案し、検証する必要がある。

## 第5章

# 総合考察

---

この章では、診断推定と異常検知のための複合的アプローチで得られた実験結果に対し、総合的な考察を行う。関連研究との定量的、定性的な比較や、本研究が持つ特色、将来的な展望について示す。その上で、今回得られた課題についても考察を行い、今後の方針や改善点について述べる。

---

### 5.1 診断推定研究と異常検知研究からの総合考察

本論文では、人間を中心とした観点から従来の分析を発展させ、個人が持つ特徴や潜在的要因を考慮して従来の手法と組み合わせる複合的アプローチを提案した。

診断推定と異常検知における複合的アプローチの2つを研究事例として紹介し、精度が改善できるだけでなく、さらには結果が得られるプロセスも可視化することができることを示した。どちらにおいても、人のもつ特徴や潜在的要因を複数の手法を用いて分析したことにより、従来の手法では分析できなかった健康状態や体調を取り入れた高い精度での分析を行うことができた。これにより、人間の個人差や特徴を含めてどのように分析するかについて考える必要があり、複数の手法を統合させて人間科学的視点からも分析することが重要であることが明らかとなった。

診断推定における複合的アプローチは、ヘルスデータが脈診断と関係していることを機械学習を用いた実験で明らかにした。そこで得られた知見をもとに、ヘルスデータからの生活周期を用いて従来の脈診断推定と組み合わせる手法を提案し検証した。これらの研究により、東洋医学の理念だけでなくデータ分析の面から見てもヘルスデータが脈診断と関係しており、それを考慮してさらなる分析を行う必要があることが示された。

表 5.1 診断推定における関連研究との比較

研究	分類手法	分類の種類	用いた特徴量	精度 (%)
Li et al. [23]	GBDT	シングルラベル分類	脈像	90.5
Yan et al. [25]	KNN	シングルラベル分類	脈像	97.8
Li et al. [26]	ニューラルネットワーク	シングルラベル分類	脈像	95.0
Zhang et al. [27]	CNN	シングルラベル分類	脈像の画像	93.5
Guo et al. [28]	確率的ニューラルネットワーク	シングルラベル分類	パワースペクトル等	93.0
本研究 [10]	ニューラルネットワーク	マルチラベル分類	脈像 & ヘルスデータ	98.5

脈診断を推定する関連研究においては、脈特徴量のみを用いて分析したものが多い。表 5.1 に示すように、本研究では、他の研究と比較しても高い精度を持つ推定を行うことが出来たが、これは他の研究と比較してヘルスデータの特徴量を取りれたためであると考えられる。さらに、シングルラベル分類ではなくマルチラベル分類を行って得られた精度であるため、より実際の診察に近い分類で高い精度での推定を行うことが出来た。また、ヘルスデータを用いて医師の診断を予測した場合でも、脈診機の結果と一致することが多く、ヘルスデータは脈診機の推定結果を向上させる事がわかった。このことは、中医学の基本的な理念である人体を総合的に見るという考え方と一致しており、人間が生活の中で持つ様々な特徴を加えることにより、高い精度での診断推定が行えたといえる。

ヘルスデータをクラスター分析した実験では、男女によって明確に分かれることはなかったが、その差を示唆する結果が得られた。これにより、脈診断推定においても男女を分けることで結果が改善する可能性があるが、この点についてはさらなる検証が必要である。多くの関連研究 [23, 25] では、患者を男女や年代で分けることはせず、実験協力者の個人差までは注目していない。そのため、機械学習の特徴量として取り入れたり、データセットをその特徴で分割するなど、学習に反映させることにより、さらなる改善が行える可能性がある。

中医学は主観的な診断が主であるため定量的研究が難しく、自動化のための研究においてもセンサーの開発、体調の見立て、漢方薬の処方方法など取り組む点は多く細分化している。人体を総合的に見るという観点から考えると、現実の診断方法と同じように、陰陽、気血水、五行などの観点から体全体を踏まえた上で脈診断を考える必要がある。医師の診断を完全に再現するためには陰陽、気血水などを判定するためには体温や顔色、生気、声色などその人全体を踏まえなければならない。従来の関連研究がおこなっているような脈像という単一の視点から診断推定を行うのではなく、生活周期も体調の一部として取り入れることで、より中医学の理念に近づけることができた。これにより、体を総合的に捉えて

分析を行うことを一歩前進させることが出来たと言える。

診断推定における今後の課題としては、どのような要素が診断を関係あるかを考えて、さらに多くの要素を取り入れて総合的に人間を分析していくことである。今回の脈診断ではヘルスデータ分析によって生活周期を推定したが、生活周期の他、中医学の観点からは他の特徴（体温、顔色、疲労感）なども取り入れて分析することが望ましい。また、実現に向けたシステム構築や、デバイスとの連携も必要であるため、実装に向けてさらなる検討が必要である。生活周期をどのように推定するか、また、歩数だけでなく位置情報や変数の組み合わせなどもさらに手法を追加して考慮する必要がある。また、脈診機の診断結果もあるため、この診断を医師の診断に近づける手法についても研究する予定である。これにより、ヘルスデータ分析によって従来の研究の推定精度を高めるだけでなく、ウェアラブルデバイスと脈診機のみでも精度の高い結果が得られることが期待できる。

異常検知における複合的アプローチでは、その個人にある潜在的要因をどのように定量化するか、また、それを用いた従来の分析手法との組み合わせについて提案して検証を行った。潜在的要因として活動因子を用い、ヘルスデータから定量化した場合においては、その活動因子でクラスタリングしたときにクラスタ毎の特徴があり、行動や体の状態が異なっていることが明らかとなった。さらに、疲労を潜在的要因とした実験においては、ウェアラブルデバイスによる疲労度の推定では発見できないような細かな外れ値も発見できることを示した。この研究においても、SEM と HMM を用いて複合的に推定した疲労度は、疲労が大きくなるにつれて心拍数が増加するなど疲労が溜まっている状態と似た特徴が現れていた。活動因子でクラスタリングすることによる特徴と同様に、この研究においてもクラスターに特徴が現れていると言え、疲労度に従って心拍が高くなるなどの傾向が見られた。これは、SEM と HMM による潜在的要因の定量化がうまく行えていることを示すものであると考えることができる。オープンデータを用いた異常検知実験においても、疾病リスクの度合いを推定して度合い別に異常検知することにより、精度が高く偽陽性が低い異常検知を行うことが出来た。さらに、異常検知を行う際には、命に関わることもあるため、その結果に対する信頼性や責任が求められる。ディープラーニングでは結果を得られるプロセスがブラックボックスであるという性質上、信頼性には限界がある。本アプローチのように病気の度合いを推定してそれぞれに細かい検知を行うことにより、ディープラーニングとほぼ同等の精度が得られ、ホワイトボックスな手段を提案することが出来たということは、信頼性が向上したと言える。そのため、潜在的要因を考慮することにより、ただ精度が向上しただけではなく、結果が持つ信頼性も向上させる事ができた。

異常検知のための複合的アプローチと関連研究について、その比較を表 5.2 に示す。ここからもわかるように、異常検知は直接バイタルサインなどを見て判断することがほとんどであり、個人に潜む疲労

表 5.2 異常検知における関連研究との比較

研究	アプローチの特徴	使用したデータ	個人が持つ特徴の分析
Liang et al. [17]	睡眠パターンの異常検知	睡眠データ	-
Balasubramaniyam et al. [37]	T シャツによる呼吸異常の検知	呼吸データ	-
Ma and Lee [40]	異常の重要度別の可視化	バイタルサイン	-
Amor et al. [41]	総合的指標による異常検知と異常変数の特定	バイタルサイン	-
本研究	SEM と HMM による疾病リスク推定	バイタルサイン	○

度や疾病リスク，肉体的負担，精神的負担，ストレス度などを考慮することは少ない。様々なデータに対して異常検知は行われているものの，個人の特徴や潜在的要素に対してまで分析を行い，組み込んでいる研究は多くない。疾病リスクを推定する本アプローチでは，従来の研究で得られている精度を向上させることができることを示したが，潜在的要因の対象を変えることで，他のデータに対しても適用でき，その精度を改善させることができると思われる。現在のウェアラブルデバイスにも異常検知や転倒検知，緊急通報を行う機能がついているものもあるが<sup>\*1</sup>，その個人に潜む要因までは捉えられていない。このようなウェアラブルデバイスに実装することができれば，異常の早期発見や予防が期待できる。今後は，本研究で得られたアプローチをどのように実装するか，また，地域や医療機関と連携させるかについても考えていく必要がある。

異常検知における今後の課題としては，推定された状態が本当はどのような状態であったかについてのさらなる検証や，コンテキストまで含めた分析手法の提案が必要である。ウェアラブルデバイスとスマートフォンとの連携についても，実装に向けて検討していきたい。異常検知における実験では，被験者の数の不足，ドメインモデルの作成方法，推定された状態と実際の状態の比較についても課題が残っている。自動的なドメインモデルに向けて，変数間の関係性から合理性が失われないように辞書を作成して組み込んでおき，得られたデータから探索的にモデルを探っていくその辞書と照らし合わせるなど，別途改良が必要である。これらについてさらなる実験デザインを考え，複合的アプローチのさらなる改善について探っていく予定である。

## 5.2 本研究の特色と今後の展開

従来の研究と比較し，本研究では人間中心の観点から，生活周期や疾病リスクと言った人間が持つ要素や特徴に対して複合的に分析を行うアプローチを示し，画一的な分析手法と比較してその精度を向上

<sup>\*1</sup> Apple Watch Series 5, <https://www.apple.com/jp/apple-watch-series-5/health/> (最終確認日 2020/12/16)

できることを示した。その有効性を示すことにより、人間がもつ様々な要素を考慮することがより個人に適した結果のために必要であるという人間科学的視点に立つことの重要性を示す事ができた。また、健康管理や福祉に関わる研究に対し、どのように情報科学が関係し、その分野に対して貢献できるのかを示すことができた。

さらに、精度が向上できるだけでなく、ディープラーニングと同等の精度を持ち、結果を得られるプロセスを見える化できるホワイトボックスでのアプローチの有効性も示した。これにより、従来の研究結果を更に発展させるだけでなく、実際に利活用する際の結果に対する信頼性も向上した。人間がもつ特徴や要素の分析にて結果を得られる過程を見える化することによって、信頼性が向上するだけでなく生活周期や疾病リスク度をユーザーが視覚的に理解しやすくなる。これは自分の健康状態へのより直感的な理解に繋がり、健康管理システムの普及や誰でも使えるサービスへ応用することが期待できる。人間科学的視点からの分析を行う上で様々な手法を組み合わせることにより、従来では分析できなかった事柄に対しても定量的に可視化することができることが明らかとなった。疾病リスクや生活周期と言った人間が持つ特徴を定量的に分析する手法を改善していくことにより、従来との研究と組み合わせた発展的な研究も期待できる。さらには、専門医がいない地域においても公民館などで脈診機とウェアラブルデバイスを組み合わせたより高精度な健康管理が行えれば、健康診断など健康増進事業の一環としても活用することが可能である。このように、個人の特徴や潜在的要素を考慮した複合的アプローチが将来的に実装されることにより、より個人に適した異常検知や診断推定が行えるため、社会が抱える医療負担、健康寿命の増進に貢献できる。

これまでのヘルスデータ分析においては、人が持つ特徴や要因は考慮されることが少なく、変わった点を持つ人物は外れ値として除外されることも多かった。そのような人に対してもその個人に適した結果をフィードバックすることが重要であり、誰でも自身の健康を理解し管理することが必要である。この研究では個人が持つ様々な要素や特徴を取り入れたことで、ICTを活用するでも人間性を尊重して分析を行うべきであるということを示した。これにより、人と技術のあり方についてヘルスデータ分析という観点を通じて明らかにすることにより、人間科学へ貢献することができたと言える。

今後の発展としては、健康管理の精度向上による生活の質向上や健康寿命の増進に寄与するべく社会実装を行っていきたいと考えている。複合的アプローチを用いて、ウェアラブルデバイスなどでより個人に適した診断推定や早期の異常検知が行えるようになれば、健康寿命を伸ばすことができ公費医療負担の削減も行うことが出来る。現代社会が抱える大きな問題である高齢化に伴う後期高齢者医療制度による医療費抑制や、介護保険制度の負担軽減などの効果も期待できる。そのため実現化に向けて、システムの試作や複数の病気への対応、さらなるデータセットでの検証などを行っていきたいと考えている。

## 第6章

### 結論

---

この章では、まとめと今後の課題について述べる。本研究で提案した複合的アプローチを紹介し、実験で得られた結果について整理する。また、学術的寄与や社会的貢献について記述する。最後に、考察で得られた今後の課題や取り組むべき点を示す。

---

#### 6.1 本研究のまとめ

ここでは、人間中心という視点から人が持つ特徴や潜在的要因に着目し、従来の単一な手法ではなく複数の手法を組み合わせて用いるヘルスデータ分析のための複合的アプローチを提案した。診断推定における複合的アプローチでは、従来の脈像からの特徴量を用いるだけでなく、人が持つ生活周期も自己相関分析と PCA によってヘルス特徴量として用いることで、脈診断の推定精度を最大で 98.5% まで向上させられることを明らかにした。異常検知における複合的アプローチでは、潜在的要因を構造方程式モデリングや隠れマルコフモデルによって疲労や疾病リスクの度合いを推定し、その度合い別に異常検知を行うことで、精度が 98.75%、偽陽性が 0.185% と、性能が改善されてより人の状態に合わせた異常検知が行えることを示した。さらには、この異常検知手法はディープラーニングと違って統計的手法を用いているため結果を得られる過程や式を明示的に示すことができるため、見える化によって結果の信頼性も向上させることができた。これらの研究事例から、ヘルスデータを人間中心の観点からも分析し、そのために分析手法を組み合わせることにより従来の研究が改善されることを示した。

本研究では、人間中心の視点から考えて個人の持つ特徴や性質まで分析する複合的アプローチによって従来研究の精度が改善でき、さらに結果を得るプロセスの見える化や個人に合わせたデータ分析が可

能となることを示した。また、人間中心の観点から多角的に複数の分析手法を組み合わせる複合的なアプローチを新たに提案し、計算論的アルゴリズムによる実装の可能性と有効性を示唆した。人間中心の視点から個人の持つ特徴や潜在的要素まで分析することで、その結果を改善できるだけでなく従来の研究に新たな視点を示すことにより、学術的発展に寄与する。この複合的アプローチが実用化されることで、医師による診断の補助や個人に合わせた高い精度での健康管理サービスを提供することが可能となり、社会的貢献が期待できる。

## 6.2 今後の課題

複合的アプローチによって従来の研究を発展させることができるが、異常検知や診断推定など個別の研究分野において、何をどのように組み合わせるか、分析するかについては方法論がまだ確立できていない。生活周期の他に何をとり入れるべきか、疾病リスクの他に関係する潜在的要因は無いか、またそれらをどのように分析すればよいかについてなど、一つ一つ検証していく必要がある。それに加え、複合的アプローチによって得られた人間が持つ特徴や潜在的要素の分析方法についても、さらに改良して精度高く分析できるようにする必要がある。実際の状況と比較して潜在的要因の度合いはどの程度合致しているか、また、それをどのように検証するのかについても今後考察していきたい。さらなるデータの収集やオープンデータによる実験を通じ、今回提案したアプローチがどのような病気や診断に適用できるかについて更に検証していく予定である。さらに、ウェアラブルデバイスだけではなく他のデバイスとの組み合わせも想定した総合的な異常検知や生活支援につながるよう、研究を進めていきたいと考えている。例えば、既存のデバイスに人間科学的な視点を取り入れて分析できるようにするだけではなく、家の中のカメラから動作を検出して異常検知と組み合わせたり、スマートフォンの内部センサーやカメラとも組み合わせ分析を行うなど、さらなる分析手法の拡張に取り組んでいきたいと思っている。複合的アプローチは、異常検知や診断推定以外にも、他のヘルスデータ分析にも応用可能な手法であるため、どのような分析を発展させられるか、どのように人間科学的視点を組み込んでいくか、今後の軸となる研究テーマとしてさらに検討していきたいと考えている。

# 謝辞

本論文を執筆するにあたり、学生、教員、職員の皆様など関係したすべての皆様に深謝いたします。社会人経験者かつ文学部出身という異例な経歴にも関わらず、大きな心で研究室に受け入れて頂いただけでなく、修士課程早期修了や助手としての推薦、懇意かつ丁寧に最後までご指導いただいた金先生には、心より感謝申し上げます。また、副査である西村先生は、実験のデザインや統計的手法での分析方法など、自分が気づかないような多くの点から忌憚なくご指導いただきました。同じく副査の扇原先生にも助けて頂き、情報科学とは別角度である社会福祉や健康管理といった観点から助言を頂いただけでなく、介護施設見学や病院との連携プロジェクトに参加させてもらい、自分のテーマを実践でどのように活かせるかを考える上で大変参考になる経験をさせていただきました。

博士論文のプレッシャーを共有して共に励まし合いながら進んでいった博士後期課程の粕谷清治さんと藤谷猛さんには、大学院において苦楽を共にし大変お世話になりました。共同研究者として論文執筆の際やディスカッションに積極的に関わっていただいたサミー株式会社の高木皓介さんには、技術面や執筆に関して大きく助けていただきました。ブロックチェーンの有用性を教えてくれた日本 IBM 株式会社の伊藤賢一さんにおきましては、その知識や考え方が大変参考となりました。貴重な脈診断のデータを共有していただいた周思宇さんは、同じテーマで切磋琢磨しつつ助け合うメンバーとして、論文を執筆する上で欠かせない存在でした。王海東さんには、ディープラーニングのプログラミングやデータサイエンス勉強会での運営など、研究以外の面でも大変お世話になりました。加えて、事務職員の皆様、情報システム運営室や技術管理室の皆様、安全管理していただいている警備の皆様、快く研究ができるように掃除をしていただいている清掃スタッフの皆様など、普段は意識しないようなところでも直接・間接を問わず多大なるご支援をいただきました。この場を借りて、改めまして関係した皆様に御礼申し上げます。

最後に、大学院での研究を支え、いつでも応援してくれた家族に感謝の言葉を添えて、本論文の謝辞とさせていただきます。

## 参考文献

- [1] 厚生労働省 高齢者医療制度等改革推進本部事務局: 医療制度改革の課題と視点, <http://www.globus.org/toolkit/>. (Last Accessed on 2020/7/3).
- [2] Li, H., Wu, J., Gao, Y., Shi, Y.: Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective, *International journal of medical informatics*, Vol. 88, pp. 8–17 (2016).
- [3] Mahata, J., Phadikar, A.: Recent advances in human behaviour understanding: A survey, *Proceedings of the 2017 Devices for Integrated Circuit (DevIC)*, pp. 751–755 (2017).
- [4] Belimpasakis, P., Roimela, K., You, Y.: Experience Explorer: A Life-Logging Platform Based on Mobile Context Collection, *Proceedings of the 2009 3rd International Conference on Next Generation Mobile Applications, Services and Technologies*, pp. 77–82 (2009).
- [5] Deng, Z., Yang, P., Zhao, Y., Zhao, X., Dong, F.: Life-Logging Data Aggregation Solution for Interdisciplinary Healthcare Research and Collaboration, *Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, pp. 2315–2320 (2015).
- [6] Zheng, X., Mukkamala, R. R., Vatrappu, R., Ordieres-Mere, J.: Blockchain-based Personal Health Data Sharing System Using Cloud Storage, *Proceedings of the 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom2018)*, pp. 1–6 (2018).
- [7] Zhang, H., Jiang, M.: An integrated approach for implantable medical devices fatigue reliability prediction, *Proceedings of the 2017 Annual Reliability and Maintainability Symposium (RAMS2017)*, pp. 1–6 (2017).
- [8] Kachroo, P., Eraso, J. M., Beres, S. B., Olsen, R. J., Zhu, L., Nasser, W., Bernard, P. E.,

- Cantu, C. C., Saavedra, M. O., Arredondo, M. J. et al.: Integrated analysis of population genomics, transcriptomics and virulence provides novel insights into *Streptococcus pyogenes* pathogenesis, *Nature genetics*, Vol. 51, No. 3, pp. 548–559 (2019).
- [9] Tago, K., Ogihara, A., Nishimura, S., Jin, Q.: Analysis of Pulse Diagnosis Data from a TCM Doctor and a Device by Random Forest, *New Frontiers in Artificial Intelligence – JSAI-isAI 2018. Lecture Notes in Computer Science* (Kojima, K., Sakamoto, M., Mineshima, K., Satoh, K., eds.), Vol. 11717, Cham, Springer International Publishing, pp. 74–80 (2019).
- [10] Tago, K., Wang, H., Jin, Q.: Classification of TCM pulse diagnoses based on pulse and periodic features from personal health data, *Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM2019)* (2019).
- [11] Tago, K., Nishimura, S., Ogihara, A., Jin, Q.: Improving diagnosis estimation by considering the periodic span of life cycle based on personal health data, *Big Data Research*. (submitting).
- [12] Tago, K., Takagi, K., Ito, K., Jin, Q.: Specifying Latent Factors with a Domain Model for Personal Data Analysis, *Proceedings of the 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*, pp. 286–291 (2018).
- [13] Tago, K., Jin, Q.: Detection of Anomaly Health Data by Specifying Latent Factors with SEM and Estimating Hidden States with HMM, *Proceedings of the 9th IEEE International Conference on Information Technology in Medicine and Education*, Hangzhou, China, pp. 137–141 (2018).
- [14] Tago, K., Takagi, K., Jin, Q.: Detection of Health Abnormality Considering Latent Factors Inducing a Disease, *IEEE Access*, Vol. 8, pp. 139433–139443 (2020).
- [15] Rajkomar, A. et al.: Scalable and accurate deep learning with electronic health records, *NPJ Digital Medicine*, Vol. 1, No. 18 (2018).
- [16] Lutze, R., Waldhör, K.: Personal Health Assistance for Elderly People via Smartwatch Based Motion Analysis, *Proceedings of the 2017 IEEE International Conference on Healthcare Informatics (ICHI2017)*, pp. 124–133 (2017).
- [17] Liang, Z., Martell, M. A. C., Nishimura, T.: A Personalized Approach for Detecting Unusual Sleep from Time Series Sleep-Tracking Data, *Proceedings of the 2016 IEEE International*

*Conference on Healthcare Informatics (ICHI2016)*, pp. 18–23 (2016).

- [18] Merilahti, J., Viramo, P., Korhonen, I.: Wearable Monitoring of Physical Functioning and Disability Changes, Circadian Rhythms and Sleep Patterns in Nursing Home Residents, *IEEE Journal of Biomedical and Health Informatics*, Vol. 20, pp. 856–864 (2016).
- [19] Guo, A., Ma, J.: A Smartphone-Based System for Personal Data Management and Personality Analysis, *Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, pp. 2114–2122 (2015).
- [20] Schuldhaus, D., Leutheuser, H., Eskofier, B. M.: Classification of Daily Life Activities by Decision Level Fusion of Inertial Sensor Data, *Proceedings of the 8th International Conference on Body Area Networks*, Brussels, BEL, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), p. 77–82 (2013).
- [21] Yan, H., Wang, Y., Liu, Z., Guo, R., Li, F., Run, F., Hong, Y.: Feature Extraction for Pulse Waveform in Traditional Chinese Medicine by Hemodynamic Analysis, *Proceedings of the 2009 IEEE International Conference on Bioinformatics and Biomedicine*, pp. 234–238 (2009).
- [22] Wang, B., Guo, H., Zheng, S., Yang, X.: TCM Pulse-Condition Classification Method Based on BP Neural Network, *Proceedings of the 1st International Conference on Bioinformatics and Biomedical Engineering*, pp. 629–632 (2007).
- [23] Li, X., Fu, Z., Qian, P., Wang, L., Zhang, H., Zhou, X., Zhang, W., Li, F.: Computerized Wrist pulse signal Diagnosis using Gradient Boosting Decision Tree, *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM2018)*, pp. 1941–1947 (2018).
- [24] Ma, C., Xia, C., Wang, Y., Yan, H., Li, F.: An improved approach to the classification of seven common TCM pulse conditions, *Proceedings of the 4th International Conference on Biomedical Engineering and Informatics (BMEI2011)*, Vol. 2, pp. 621–624 (2011).
- [25] Yan, H., Wang, Y. Q., Guo, R., Z.Liu, Li, F., Run, F., Hong, Y., Yan, J.: Feature extraction and recognition for pulse waveform in Traditional Chinese Medicine based on hemodynamics principle, *Proceedings of the 8th IEEE International Conference on Computer Application (ICCA2010)*, pp. 972–976 (2010).
- [26] Li, H., Xu, B., Wang, N., Liu, J.: Deep Convolutional Neural Networks for Classifying Body

- Constitution, *Proceedings of the 25th International Conference on Artificial Neural Networks (ICANN2016)*, Springer International Publishing, pp. 128–135 (2016).
- [27] Zhang, S., Sun, Q.: Human Pulse Recognition Based on Convolutional Neural Networks, *Proceedings of the 3rd International Symposium on Computer, Consumer and Control (IS3C2016)*, pp. 366–369 (2016).
- [28] Guo, H., Wang, B.: Research on TCM pulse condition identification using probabilistic neural networks, *Proceedings of the 3rd International Conference on Biomedical Engineering and Informatics*, Vol. 6, pp. 2352–2355 (2010).
- [29] Guo, R., Yan, H., Wang, Y., Yan, J.: TCM pulse analysis of the patients with coronary heart disease based on Multiscale Entropy, *Proceedings of the 11th IEEE International Conference on Control Automation (ICCA2014)*, pp. 1135–1138 (2014).
- [30] Li, J., Li, J., Yi, D.: Research on the Hypertension Syndrome Elements Differentiation of TCM Based on Multi-label Learning and Ensemble Learning, *Proceedings of the 2018 International Conference on Big Data and Computing*, pp. 99–103 (2018).
- [31] Wu, H., Gao, H.: The Application of TCM Five-Pattern Personality and Constitution Identification System in Population Health Management of Elderly People in Community, *Proceedings of the 9th International Conference on Information Technology in Medicine and Education (ITME2018)*, pp. 300–303 (2018).
- [32] Layne, K., Ferro, A.: Traditional Chinese medicines in the management of cardiovascular diseases: a comprehensive systematic review, *British journal of clinical pharmacology*, Vol. 83, No. 1, pp. 20–32 (2017).
- [33] Guo, F., Dai, Y., Li, S., Ito, K.: Inferring individuals' sub-health and their TCM syndrome based on the diagnosis of TCM doctors, *Proceedings of the 2010 IEEE International Conference on Systems, Man and Cybernetics*, pp. 3717–3724 (2010).
- [34] Guo, F., Dai, Y., Lin, Y., Li, S., Ito, K.: Consistency based rules mining on sparse and diverse TCM sub-health diagnosis data, *Proceedings of the 2010 IEEE International Conference on Intelligent Computing and Intelligent Systems*, pp. 896–901 (2010).
- [35] Jin, Y., Zhang, W., He, X., Wang, X., Wang, X.: Syndrome-aware Herb Recommendation with Multi-Graph Convolution Network, *Proceedings of the 2020 IEEE 36th International Conference on Data Engineering (ICDE2020)*, pp. 145–156 (2020).

- [36] Liu, Y., Sui, Z., Zhou, Y., Wang, Z.: Research on the Key Technology Based-NLP Of Chinese Medicine Pulse's Mathematical Quantifying, *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, Vol. 3, pp. 307–310 (2008).
- [37] Balasubramaniyam, H., Vignesh, M. S., Florence S, E., A, A., Abanah, A.: Design and Development of a IoT based Flexible and Wearable T-Shirt for Monitoring Breathing Rate, *Proceedings of the 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC2019)*, pp. 376–379 (2019).
- [38] Ren, H., Ye, Z., Li, Z.: Anomaly detection based on a dynamic Markov model, *Information Sciences*, Vol. 411, pp. 52–65 (2017).
- [39] Hotelling, H.: The Generalization of Student's Ratio, *Annals of Mathematical Statistics*, Vol. 2, No. 3, pp. 360–378 (1931).
- [40] Ma, F., Lee, C. E.: Integrated Control Chart for Vital Signs Early Warning of Long-term Care Patients, *Proceedings of the 7th International Conference on Ubi-Media Computing and Workshops*, Ulaanbaatar, Mongolia, pp. 313–318 (2014).
- [41] Amor, L. B., Lahyani, I., Jmaiel, M.: Data accuracy aware mobile healthcare applications, *Computers in Industry*, Vol. 97, pp. 54–66 (2018).
- [42] Yan, K., You, X., Ji, X., Yin, G., Yang, F.: A Hybrid Outlier Detection Method for Health Care Big Data, *Proceedings of the 2016 IEEE International Conference on Big Data Cloud Computing, Social Computing Networking, Sustainable Computing Communications*, Atlanta, GA, USA, pp. 157–162 (2016).
- [43] Kawanishi, K., Kawanaka, H., Takase, H., Tsuruoka, S.: A study on dementia detection method with stroke data using anomaly detection, *Proceedings of the 2017 6th International Conference on Informatics, Electronics and Vision 2017 7th International Symposium in Computational Medical and Health Technology*, pp. 1–4 (2017).
- [44] Hela, S., Amel, B., Badran, R.: Early anomaly detection in smart home: A causal association rule-based approach, *Artificial Intelligence in Medicine*, Vol. 91, pp. 57–71 (2018).
- [45] Caroprese, L., Veltri, P., Vocaturo, E., Zumpano, E.: Deep Learning Techniques for Electronic Health Record Analysis, *Proceedings of the 2018 9th International Conference on Information, Intelligence, Systems and Applications (IISA2018)*, pp. 1–4 (2018).

- [46] Hsiao, H. C. W., Chen, S. H. F., Tsai, J. J. P.: Deep Learning for Risk Analysis of Specific Cardiovascular Diseases Using Environmental Data and Outpatient Records, *Proceedings of the 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE2016)*, pp. 369–372 (2016).
- [47] Li, X., Li, J.: Health Risk Prediction Using Big Medical Data - a Collaborative Filtering-Enhanced Deep Learning Approach, *Proceedings of the 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom2018)*, pp. 1–7 (2018).
- [48] Vasquez-Morales, G. R., Martinez-Monterrubio, S. M., Moreno-Ger, P., Recio-Garcia, J. A.: Explainable Prediction of Chronic Renal Disease in the Colombian Population Using Neural Networks and Case-Based Reasoning, *IEEE Access*, Vol. 7, pp. 152900–152910 (2019).
- [49] Amarasinghe, K., Kenney, K., Manic, M.: Toward Explainable Deep Neural Network Based Anomaly Detection, *Proceedings of the 2018 11th International Conference on Human System Interaction (HSI2018)*, pp. 311–317 (2018).
- [50] Bluysen, P. M.: Towards an integrated analysis of the indoor environmental factors and its effects on occupants, *Intelligent Buildings International* (2019). <https://doi.org/10.1080/17508975.2019.1599318>.
- [51] Chen, J., Xu, Y., Wu, L., Liu, Y., Zhao, T.: Eye-tracking-Based Quality Assessment for Image Interpolation, *2018 IEEE Visual Communications and Image Processing (VCIP2018)*, pp. 1–4 (2018).
- [52] Yang, C., Chang, T., Chen, Y.: Microwave sensors applying for Traditional Chinese Medicine pulse diagnosis, *Proceedings of the IEEE 8th International Workshop on Electromagnetics: Applications and Student Innovation Competition (iWEM2017)*, pp. 113–115 (2017).
- [53] Tsoumakas, G., Katakis, I., Vlahavas, I.: Mining multi-label data, *Data mining and knowledge discovery handbook*, Springer, pp. 667–685 (2009).
- [54] Ward Jr, J. H.: Hierarchical grouping to optimize an objective function, *Journal of the American statistical association*, Vol. 58, No. 301, pp. 236–244 (1963).
- [55] Bellman, R. E., Dreyfus, S. E.: *Applied dynamic programming*, Princeton university press (2015).
- [56] Ljung, G. M., Box, G. E. P.: On a measure of lack of fit in time series models, *Biometrika*, Vol. 65, No. 2, pp. 297–303 (1978).

- [57] Schelter, B., Winterhalder, M., Timmer, J.: *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, Wiley - VCH Verlag GmbH & Co. KGaA (2006).
- [58] Davier, A. V.: *Statistical Models for Test Equating, Scaling, and Linking*, Springer, New York (2011).
- [59] Higashi, K., Koyama, K., Ozawa, R., Nagata, K., Wan, W., Harada, K.: Functionally Divided Manipulation Synergy for Controlling Multi-fingered Hands (2020). <https://arxiv.org/abs/2003.11699>.
- [60] Xia, J., Liao, W., Chanussot, J., Du, P., Song, G., Philips, W.: Improving Random Forest With Ensemble of Features and Semisupervised Feature Extraction, *IEEE Geoscience and Remote Sensing Letters*, Vol. 12, No. 7, pp. 1471–1475 (2015).
- [61] Rallapalli, S., Suryakanthi, T.: Predicting the risk of diabetes in big data electronic health Records by using scalable random forest classification algorithm, *Proceedings of the 2016 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, pp. 281–284 (2016).
- [62] He, H., Bai, Y., Garcia, E. A., Li, S.: ADASYN: Adaptive synthetic sampling approach for imbalanced learning, *Proceedings of the IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pp. 1322–1328 (2008).
- [63] Peimankar, A., Jajroodi, M. J., Puthusserypady, S.: Automatic Detection of Cardiac Arrhythmias Using Ensemble Learning, *Proceedings of the IEEE Region 10 Conference (TENCON2019)*, pp. 383–388 (2019).
- [64] Zhang, M.-L., Li, Y.-K., Liu, X.-Y., Geng, X.: Binary relevance for multi-label learning: an overview, *Frontiers of Computer Science*, Vol. 12, No. 2, pp. 191–202 (2018).
- [65] Zhang, M.-L., Zhou, Z.-H.: ML-KNN: A lazy learning approach to multi-label learning, *Pattern Recognition*, Vol. 40, No. 7, pp. 2038 – 2048 (2007).
- [66] Hammad, M., Maher, A., Wang, K., Jiang, F., Amrani, M.: Detection of abnormal heart conditions based on characteristics of ECG signals, *Measurement*, Vol. 125, pp. 634–644 (2018).
- [67] Sharma, M., Patel, S., Acharya, U. R.: Automated detection of abnormal EEG signals using localized wavelet filter banks, *Pattern Recognition Letters*, Vol. 133, pp. 188–194 (2020).
- [68] May, K. R., Noah, B. E., Walker, B. N.: Driving Acceptance: Applying Structural Equation Modeling to In-Vehicle Automation Acceptance, *Proceedings of the 9th International Con-*

- ference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct*, pp. 190–194 (2017).
- [69] Khakurel, J., Knutas, A., Immonen, M., Porras, J.: Intended Use of Smartwatches and Pedometers in the University Environment: An Empirical Analysis, *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, p. 97–100 (2017).
- [70] Tong, Q., Zou, X., Gong, Y., Tong, H.: Analysis of psychological health and life qualities of internet addicts using structural equation model, *Psychology*, Vol. 1, No. 1, pp. 22–26 (2010).
- [71] Demerouti, E., Bakker, A. B., Nachreiner, F., Schaufeli, W. B.: A model of burnout and life satisfaction amongst nurses, *Journal of advanced nursing*, Vol. 32, No. 2, pp. 454–464 (2000).
- [72] R.O., M.: *Basic Principles of Structural Equation Modeling*, Springer, Piscataway, NJ, USA (1996).
- [73] Toyoda, H.: *Covariance Structure Analysis*, Asakura Shoten, Tokyo (1998).
- [74] Asano, H., Suzuki, T., Kojima, T.: *Getting started in Covariance Structure Analysis*, Kodansha, Tokyo (2005).
- [75] Akaike, H.: Information theory and an extension of the maximum likelihood principle, *Proceedings of the 2nd international symposium on information theory*, pp. 267–281 (1973).
- [76] Steiger, J. H.: Statistically based tests for the number of common factors, *Paper presented at the annual meeting of the Psychometric Society, Iowa City* (1980). <https://doi.org/10.1080/10705511.2016.1217487>.
- [77] Bentler, P. M., Bonett, D. G.: Significance tests and goodness of fit in the analysis of covariance structures., *Psychological bulletin*, Vol. 88, No. 3, p. 588–606 (1980).
- [78] De Meersman, R. E.: Heart rate variability and aerobic fitness, *American heart journal*, Vol. 125, No. 3, pp. 726–731 (1993).
- [79] Riediker, M., Danuser, B.: Exhaled breath condensate pH is increased after moderate exercise, *Journal of aerosol medicine*, Vol. 20, No. 1, pp. 13–18 (2007).
- [80] Alpert, B. S., Dover, E. V., Booker, D. L., Martin, A. M., Strong, W. B.: Blood pressure response to dynamic exercise in healthy children - black vs white, *The Journal of Pediatrics*, Vol. 99, No. 4, pp. 556 – 560 (1981).
- [81] Bilmes, J. A.: A gentle tutorial of the EM algorithm and its application to parameter esti-

- mation for Gaussian mixture and hidden Markov models, Technical Report Technical Report ICSI-TR-97-021, International Computer Science Institute, Berkeley (1998).
- [82] Baum, L. E., Petrie, T.: Statistical Inference for Probabilistic Functions of Finite State Markov Chains, *The Annals of Mathematical Statistics*, Vol. 37, No. 6, pp. 1554–1563 (1966).
- [83] Dempster, A. P., Laird, N. M., Rubin, D. B.: Maximum likelihood from incomplete data via the EM algorithm, *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38 (1977).
- [84] Viterbi, A.: Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, *IEEE Transactions on Information Theory*, Vol. 13, No. 2, pp. 260–269 (1967).
- [85] Ide, T., Khandelwal, A., Kalagnanam, J.: Sparse Gaussian Markov Random Field Mixtures for Anomaly Detection, *2016 IEEE 16th International Conference on Data Mining*, pp. 955–960 (2016).
- [86] Kuroda, M., Kohno, M., Go, K., Sugino, T., Ukai, T.: Hormonal Study of Blood Pressure, Heart Rate, and Survival in Head-injured Patients, *Japanese Association for Acute Medicine*, Vol. 13, No. 7, pp. 361–368 (2002).
- [87] Carney, N., Totten, A. M., O’Reilly, C., Ullman, J. S., Hawryluk, G. W. J., Bell, M. J., Bratton, S. L., Chesnut, R., Harris, O. A., Kissoon, N., Rubiano, A. M., Shutter, L., Tasker, R. C., Vavilala, M. S. and Wilberger, J., Wright, D. W., Ghajar, J.: Guidelines for the Management of Severe Traumatic Brain Injury, Fourth Edition, *Neurosurgery*, Vol. 80, No. 1, pp. 6–15 (2016).
- [88] Miyagi, T., Maeda, M., Inoue, Y., Kondou, D., Yoshimura, F., Ookura, A., Morioka, M.: Acute Therapy for Traumatic Head Injury, *The Japanese Journal of Rehabilitation Medicine*, Vol. 50, No. 7, pp. 557–569 (2013).
- [89] Jöreskog, K. G., Sörbom, D.: Recent developments in structural equation modeling, *Journal of marketing research*, Vol. 19, No. 4, pp. 404–416 (1982).
- [90] Bentler, P. M.: Comparative fit indexes in structural models., *Psychological bulletin*, Vol. 107, No. 2, pp. 238–246 (1990).
- [91] Browne, M., Cudeck, R.: Alternative ways of assessing model fit, *Testing Structural Equation Models*, Vol. 21, No. 2, pp. 230–258 (1992).
- [92] Hinton, G. E., Salakhutdinov, R. R.: Reducing the Dimensionality of Data with Neural Net-

works, *Science*, Vol. 313, No. 5786, pp. 504–507 (2006).

- [93] Pavel, M., Jimison, H. B., Korhonen, I., Gordon, C. M., Saranummi, N.: Behavioral Informatics and Computational Modeling in Support of Proactive Health Management and Care, *IEEE Transactions on Biomedical Engineering*, Vol. 62, No. 12, pp. 2763–2775 (2015).
- [94] Ismail, W. N., Hassan, M. M., Alsalamah, H. A., Fortino, G.: CNN-Based Health Model for Regular Health Factors Analysis in Internet-of-Medical Things Environment, *IEEE Access*, Vol. 8, pp. 52541–52549 (2020).