

早稲田大学審査学位論文  
博士（人間科学）  
概要書

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

センサーやデバイスの発展により、自分のバイタルサインなどの健康に関わるヘルスデータをウェアラブルデバイスなどを通じて手軽に取得できるようになった。心拍数の急激な上昇などを警告するデバイスも市販されており、自分のデータをわかりやすく可視化できたり、手軽に健康管理に役立てられるとして人気が高い。しかし、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するだけでなく、人間が持つ特徴や潜在的要素を考慮し、多くの角度からその個人を分析していくことにより、さらなる知見を得ることや研究成果を改善できる可能性がある。そのため本論文では、従来の研究や分析手法を発展させ、人間が持つ様々な要素を分析して組み合わせることによ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、その有効性について検証を行った。本学位論文は全6章から構成される。

第1章では研究背景と目的について述べた。現在、健康意識の増加に伴ってどのようなウェアラブルデバイスやサービスが提供されているかを紹介し、それぞれの特色についてまとめた。また、そのようなデバイスを用いてヘルスデータ分析を行う際の現状の課題について整理し、人間科学的視点に立ったヘルスデータ分析が必要であることを指摘した。加えて、複合的アプローチとは従来の研究手法に対して人間がもつ特徴や潜在的要素の分析を組み合わせたアプローチであることを示し、それによって従来の研究が改善される可能性があることを示した。

第2章では先行研究レビュー及び本研究の位置づけについて述べた。ヘルス分析に関わる先行研究をテーマ別に整理し、異常検知や診断推定におけるアプローチについてまとめた。その上で、個人のもつ特徴や病気リスクと言った潜在的要素がほとんど考慮されていないことを明らかにした。最後に、関連研究と比較した際の本研究の立ち位置として、これまでの研究を土台として人間科学的視点からの分析を別途提案し、組み合わせることによる複合的アプローチを提案することについて言及した。

第3章では生活周期を考慮した診断推定の研究について述べた。この研究では、従来のバイタルサインから診断を推定する研究を発展させ、生活周期の分析を組み合わせた複合的アプローチを提案し検証を行った。はじめに、人間が持っている体のバランスの崩れによって病気になるという伝統中国医学の考え方にに基づき、脈診断の結果とヘルスデータに関係性があるかどうかについて検証を行った。その結果、ヘルスデータに機械学習を適用して脈診断を推定した際、最低でも60%を超える精度が得られ、両者における関係性が示唆された。次に、得られた知見をもとにして、ヘルス関連データから周期的特徴を考慮した脈診断結果の推定を行った。生活周期の周期の長さを、歩数に対する自己相関分析から推定し、得られた1周期分の長さの期間にあるヘルスデータを主成分分析によって次元圧縮した。次元圧縮によって得られた健康データの特徴量を、脈像から得た特徴量と組み合わせて機械学習で推定した。その結果、深層学習やランダムフォレストによるマルチラベル分類で約98%の精度で推定できることを示し、従来の脈像からの特徴量のみを用いた推定よりも精度が改善できることを明らかにした。

第4章では人が持つ潜在的要因に着目した異常検知の研究について述べた。この研究では、潜在的要因を顕在的や暗示的などの4つのカテゴリから定義した。例えば、疲労度や疾病リスクなど、その度合が高まることによってバイタルサインや行動に影響を及ぼしうる要因である。この潜在的要因の度合いを推定し、従来の異常検知に組み合わせた複合的アプローチの検証を行った。はじめに、構造方程式モデリング (SEM) とドメインモデルを用いてバイタルサインから活動因子を定量化し、クラスタリングした際にヘルスデータの傾向が類似するかどうかについて検証を行った。結果として、従来の一日の活動概要カテゴリによる集約よりも、活動因子で集約したほうがクラスタごとにヘルスデータの傾向が現れることが明らかとなった。ここで得られた知見を基に、バイタルサインに影響を得与える潜在的要因として疲労因子を想定し、SEM と隠れマルコフモデル (HMM) を用いて疲労度合を推定した。その度合い別に異常検知を行った結果、ウェアラブルデバイスが行う疲労度度合い別に異常検知を行った際と比較して、本手法ではより細かく外れ値を検知できることを示した。さらに、異常検知の精度を比較するため、関連研究とデータセットを揃えて異常検知性能の比較を行った。SEM と HMM により、血圧や心拍数に影響を与える疾病リスクの度合いを求め、推定された度合い別に異常検知を行うことにより、再現率が 98.75%、偽陽性率が 0.186% と、関連研究で示されている性能よりも改善できることを示した。その上、本手法は深層学習ともほぼ同等の精度を持っており、結果を得るプロセスが可視化されているため、結果に対する信頼性も向上させることができた。

第5章では総合考察を行い本論文の特色と今後の展望を述べた。2つの研究事例から、従来の研究に人間が持つ特徴や潜在的要因を組み合わせて分析を行う複合的アプローチは、その精度や性能を向上させることができることを示した。複合的アプローチと従来の研究との違いについて定量的、定性的な比較を行い、人間科学的視点から複数の手法を組み合わせることによってなぜ改善できたか、また、結果を向上できる程度について考察した。人間科学に対する貢献として、ヘルスデータ分析においても人間中心からの分析が重要であることを示し、検証によってその有効性を示すことができた。今後の発展としては、見守りシステムや IoT と組み合わせた社会実装を今後行うことが求められ、それによって医療費負担の軽減や健康寿命の増進が期待できることについて言及した。

第6章では本研究のまとめと今後の課題について述べた。人間科学的視点から人が持つ特徴や潜在的要因を考慮して分析する手法を提案したこと、及びこれまでの研究と組み合わせて分析することによって結果が改善されることを示した。提案した複合的アプローチによってさらなる研究結果の改善が行えるだけでなく、人間中心の視点から分析することの重要性を示すことができ、人間科学の発展に寄与することができた。最後に今後の課題として、人が持つ特徴や要素に合わせた分析手法の選択方法の検討や、オープンデータを用いたさらなる関連研究との比較を行っていく予定であることを示した。