5H-03

介護現場感染症対策支援に向けた体調変調検知の機械学習およびデータエンジニアリングに関する考察

長坂　知美[†]　　千葉　祥樹[†]　　蔦谷　雄一[†]　　石黒　由紀[†]
杉本　千佳[‡]　　小林　匠[‡]　　河野　隆二[*]
富士通株式会社[†]　　横浜国立大学 大学院工学研究院[‡]　　一般社団法人ＹＲＰ国際連携研究所[*]

1. はじめに

COVID-19 感染拡大により，特に感染リスクの高い高齢者が集まる介護施設においては継続的な感染症予防対策が求められている．予防対策には COVID-19 感染判断の指標となる血中酸素飽和濃度(SpO_2)等のバイタルデータを個人毎にモニタすることが有効と考えられるが，人手不足が課題となっている介護施設では頻繁に検査を行う事は困難である．このような課題に対して，スマートウォッチで取得可能なデータや，バイタルセンサを無線ボディエリアネットワーク[1] (BAN)で繋ぎ収集した SpO_2，心拍，呼吸数等のバイタルデータを機械学習して，体調の異常を自動的に検知する取り組みが行われている[2]．しかし，バイタルセンサで取得するデータは被験者の状況や装着するデバイスによって様々であり，検知精度など実用化に向けて課題がある．

本論文では，COVID-19 等の感染による体調不良を検知するため，COVID-19 の特徴を踏まえた特徴量[3] を生成することで体調不良の検知精度の向上に取り組んだ結果について報告する．

2. データエンジニアリングについて

(1) 検証に利用するデータ

2 つの介護施設の協力を得て，介護者 9 名，被介護者 21 名，計 30 名の Fitbit Sense[a][4](Fitbit)のデータおよび日々の健康状況を記録した問診票を 10 カ月分収集した．

(2) 精度向上施策の検討

先行研究[2]の課題の 1 つとして検知精度の向上があげられている．バイタルデータの変化から風邪や COVID-19 などの病気罹患時の体調変化をより高い精度で異常として検知するためには，多くのデータを集める事が有効ではあるが，罹患前，罹患時，罹患後のデータを大量に収集する事は困難である．本論文では十分なデータの準備が難しい本課題に対して，バイタルデータや COVID-19 の特性を考慮し，体調の変化を捉えやすいデータとして加工，効率よく症状の特徴を学習する特徴量設計について下記 2 種の精度向上を行った．

・バイタルデータの特性を考慮した特徴量設計
・症状の特性を考慮した特徴量設計

(3) バイタルデータの特徴量作成の方針

収集したバイタルデータの特性を確認した所，図 1 が示すように運動時の心拍が上昇している等，睡眠や運動などの行動と共にデータが変化していることが分かった．

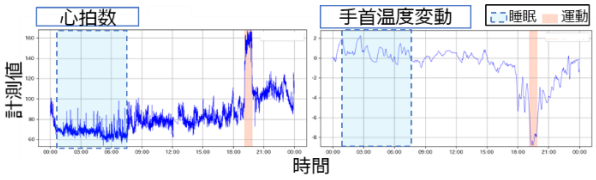

図 1　取得したバイタルデータ

また，全体のデータを確認した結果，バイタルデータの特性として以下の傾向があることが分かった．

- 起床や運動など行動に伴う突発的な変化がある
- 多くの被験者には運動等の時間に規則性がない
- 運動や睡眠などの各状態で傾向が異なり，同じ状態ではバイタルデータの推移が類似している

以上から，運動や睡眠といった行動状態ごとに特徴量を作成するアプローチが適切であると判断し，以下の特徴量を作成した．

表 1　バイタルデータの特性を考慮した特徴量

項番	Fitbit データ	状態
1	SpO_2（睡眠中）のサマリ	睡眠中
2	心拍変動(睡眠中)	睡眠中
3	呼吸数(睡眠中) のサマリ	睡眠中
4	推定酸素変動（睡眠中）	睡眠中
5	心拍の平均・分散等	睡眠中，運動中，通常
6	手首温度の平均・分散等	睡眠中，運動中，通常
7	歩数の平均・分散等	睡眠中，運動中，通常

(4) COVID-19 を検知するための特徴量の方針

COVID-19 については感染の特徴や罹患時の症状など一般的な医学的知見[5,6]が存在し，それらの情報を特徴量として取り込む事で検知精度の向上に繋がる事が想定される． COVID-19 罹患時には以下の A)～D)の特徴があるとされており，A)B)C)は Fitbit で取得できるバイタルデータに特徴が表れると考えられる (表 2)．

A)　息苦しさ(呼吸困難)
B)　強いだるさ(倦怠感)
C)　重症化の際にみられる肺炎の症状
D)　感染後から症状発症までの期間の長さ(平均 5-6 日)

また，症状を自覚する数日前から徐々に変化するバイタルデータを捉えるため，データの取得日と過去の i 日前を起点としてそれぞれ j，k 日間の平均 x_k, y_{ij} の差分を特徴量として作成した（図 2）．

「Discussion on the Utilization of Machine Learning and Data Engineering to Control Infection in Nursing Care Site by Detecting Physical Condition Changes」
[†]「Fujitsu Limited」
[‡]「Faculty of Engineering Yokohama National University」
[*]「YRP International Alliance Institute」

a)　Fitbit および Fitbit Sense のロゴは，米国およびその他の国における Fitbit, Inc の商標または登録商標です．

Copyright ©2023 Information Processing Society of Japan.
All Rights Reserved.

表 2　取得バイタル情報

項番	取得データ	COVID-19 特徴
1	SpO_2 (睡眠中)のサマリ	C)
2	心拍変動(睡眠中)	B)
3	呼吸数(睡眠中) のサマリ	A)
4	推定酸素変動（睡眠中）	C)

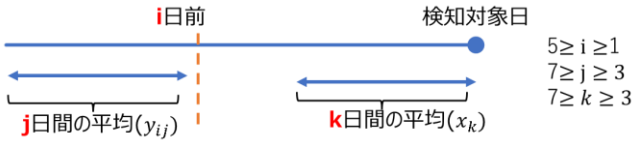

図 2　徐々に変化するバイタルデータの特徴

作成した特徴量のうち，COVID-19 罹患者のデータで正常時と感染期間で異なる傾向が見られた．図 3 は i=5，j=3，k=3 の心拍変動(睡眠中)，i=1，j= 7，k=7 の呼吸数(睡眠中)に関する縦軸が日数，横軸がバイタルデータの値のグラフである．

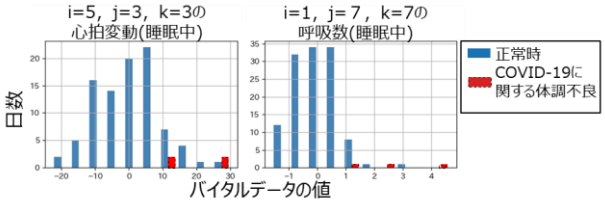

図 3　COVID-19 罹患者のデータにみられた特徴

3.　機械学習による評価実験

(1)　実施条件

体調不良の検知では 30 名のデータから COVID-19 罹患のデータが存在した 3 名のデータを活用し，体調不良が検知できるか検証した．体調不良のラベルは問診表を活用し，①正常の状態，②COVID-19 の体調不良，③COVID-19 以外の体調不良にラベリングした(表 3).

また，体調不良発生前のデータは体調不良の兆候があるデータが含まれる可能性があるため以下の条件で正常の状態から削除した．

- 体調不良が発生した前日，2 日前，3 日前のデータ
- 体調不良が発生した翌日，2 日後，3 日後のデータ

表 3　機械学習の入力データ

	データ件数	正常	COVID-19	COVID-19 以外
被験者 1	130	127	3	0
被験者 2	68	60	7	1
被験者 3	150	140	9	1

(2)　検証方法

本稿では検証に用いる機械学習の手法として説明性・解釈性から Isolation Forest[7]を採用した．正常時のデータを用いて学習し，10fold 交差検定で精度を評価した．また精度検証のための指標として，Receiver Operating Characteristic(ROC)曲線と ROC 曲線から算出した Area Under the Curve（AUC）を採用した．

(3)　機械学習の実施結果

図 4 は COVID-19 に関する体調不良の検知における被験者 1，2，3 の ROC 曲線，表 4 は ROC 曲線から AUC を計算した結果である．AUC はいずれも 0.5 以上の値であり，異常を検知できている事が分かる．また COVID-19 に対する精度のほうが高い事から，COVID-19 の観点で考察した特徴量が有効に働いている事が推測できる．

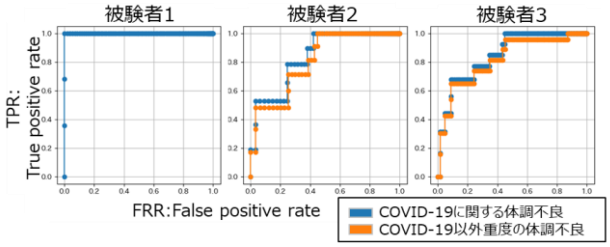

図 4　被験者 1，2，3 の ROC 曲線

表 4　被験者 1，2，3 の ROC 曲線における AUC 値

	被験者 1	被験者 2	被験者 3
COVID-19	1.00	0.907	0.888
COVID-19 以外		0.889	0.826

この結果により，罹患時の症状に関する特徴量を作成し学習することで検知精度を向上できる可能性がある事を確認できた．

4.　まとめと今後の予定

本研究では，介護現場にて収集したバイタルデータで体調不良の検知を実施し，COVID-19 罹患者の病状に特化した特徴量の作成が検知精度の向上に有効であることを確認した．今後は，COVID-19 感染判断の重要な指標と考えられるにもかかわらず今回は特徴がみられなかった SpO_2 について BAN を用いて細かい粒度でデータを取得し，さらなる検知精度の向上を試行する．COVID-19 感染以外による体調不良との差別化に加え，既往症や他の疾病の症状などを観点に含めた特徴量の作成を行う等，実用化を目指した検討を実施していく予定である．

謝辞　本研究成果は，国立研究開発法人情報通信研究機構の委託研究(222A02)により得られたものであり，本研究にご協力頂いた皆様に，謹んで感謝の意を表する．

参考文献

[1]　小林 匠, 河野 隆二, 杉本 千佳,渡邊 祥太, 蔦谷 雄一, 石黒 由紀,長島 久美子, 長坂 知美,無線ボディエリアネットワークと B5G を見据えた広域データ回線網を活用した高齢者見守り・介護支援ネットワークプラットフォームの実験的検討，IEICE ソサイエティ大会，BS-5-5，2022.

[2]　渡邊 祥太, 蔦谷 雄一, 長島 久美子, 長坂 知美, 杉本 千佳,小林 匠, 河野 隆二, ICT を活用した介護現場感染症対策支援に関する取り組み, 研究報告電子化知的財産・社会基盤, 2022-EIP-95, No. 12, pp.1 - 7, 2022.

[3]　Alice Zheng, Amanda Casari, 株式会社ホクソエム,機械学習のための特徴量エンジニアリング ―その原理と Python による実践,2019

[4]　fitbit. https://www.fitbit.com/global/jp/home, (参照 2023-1-4)

[5]　厚生労働省　新型コロナウイルスに関する Q&A https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/dengue_fever_qa_00001.html

[6]　who.int　Coronavirus disease (COVID-19) https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19

[7]　Charu C. Aggarwal. Outlier Analysis, 2013

Copyright ©2023 Information Processing Society of Japan.
All Rights Reserved.