

Title	睡眠活動データを用いた認知状態尺度推定
Author(s)	杉浦, 伸一
Citation	
Issue Date	2024-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18885
Rights	
Description	Supervisor: 岡田 将吾, 先端科学技術研究科, 修士(情報科学)

修士論文

睡眠活動データを用いた認知状態尺度推定

杉浦 伸一

主指導教員 岡田 将吾

北陸先端科学技術大学院大学
先端科学技術研究科
(情報科学)

令和6年3月

Abstract

Dementia is a disease that interferes with daily life due to the deterioration of memory and thinking ability. As the world's population ages, the impact of dementia on society is expected to increase. Although there are many diseases that cause dementia, there is still no fundamental cure for most of them, and the only effective treatment is symptomatic relief in the early stages of the disease. Therefore, it is important to detect dementia in its early stages, and research on automatic dementia detection systems that utilize various types of data is ongoing.

It is known clinically and physiologically that dementia affects sleep in various ways. In this study, we collected activity data during sleep using mattress-type sensors installed in four welfare facilities for the elderly in Japan with the aim of constructing a system to automatically estimate dementia from data collected during sleep. In addition, the Mini Mental State Examination (MMSE), a screening test for dementia, was administered to some of the participants, and their cognitive status was estimated according to the results.

To confirm the statistical nature of the obtained sleep data set, we statistically confirmed that there was a predominant difference between the high and low scoring groups on the dementia scale score for a portion of the activity measured during their sleep.

Next, using data collected from 144 participants, we compared several bicategory prediction systems for dementia scale scores based on LSTM, Support Vector Machine, and Random Forest, using only the amount of activity during sleep. All systems achieved an accuracy of over 0.62 on the Macro F1 score, indicating that it is possible to estimate the cognitive state scale to some extent even from sleep data alone. The LSTM model achieved the highest accuracy of 0.67 on the Macro F1 score. Next, by applying SHAP to the model using Random Forest, we confirmed that the feature with the largest contribution to the estimation was the depth of sleep.

Next, feature extraction methods based on X-Means, Auto Encoder, and PCA were compared. We confirmed that all feature extraction methods improved in accuracy, and the feature extraction method using X-Means achieved the highest accuracy of 0.79 on the Macro F1 score. Finally, we constructed a discrimination system using an abnormality detection model to separate sleep data into typical healthy sleep and typical unhealthy sleep. The highest Macro F1 score of 0.91 was achieved in the subsequent discrimination experiments.

The mattress-type sensor used for data collection is a commercially available product and is inexpensive compared to specialized medical equipment. It is difficult for people to recognize the signs of dementia by themselves, and it is not always possible for all people to have a third party who can recognize the trend of

cognitive decline. This study is part of a clinical study on the relationship between dementia and sleep, and suggests that sleep information can be used as part of an automatic dementia detection system that has been studied using various modalities. The study is part of a clinical study of the association between dementia and sleep.

概要

認知症は、記憶力や思考力の低下により、日常生活に支障をきたす病気である。世界的に進む高齢化社会に伴い、認知症が社会に与える影響も増大することが見込まれている。認知症を引き起こす疾患は多数あるが、多くの場合において根本的な治療法は未だに存在せず、有効な治療法はその初期段階での症状緩和のみである。従って認知症の初期段階での検知は重要であり、そのために、様々なデータを活用した認知症の自動探知システムの研究が進んでいる。

認知症は睡眠に様々な影響を与えることが臨床・生理学的に知られており、本研究では、睡眠中に収集したデータから認知症を自動的に推定するシステムの構築を目指して、日本の4つの高齢者福祉施設に設置したマットレス型のセンサーを用いて睡眠中の活動量データを収集した。さらに、参加者の一部に、認知症スクリーニング検査である Mini Mental State Examination(MMSE)を実施し、その結果に従って参加者の認知状態を推定した。

得られた睡眠データセットの統計的な性質を確認するために、認知症尺度スコアにおいて、高得点群と低得点群にはその睡眠中の計測された活動量の一部について、優位な差があることを統計的に確認した。

次に、144名の参加者から集計したデータを用いて、LSTM,Support Vector Machine, Random Forest を基盤とした睡眠中の活動量のみを用いた複数の認知症尺度スコアの二分類予測システムを比較検討した。全てのシステムにおいて Macro F1 スコアで 0.62 を超える精度を達成し、睡眠データのみからでも認知状態尺度の推定が一定程度可能であることを示した。また、LSTM を使用したモデルにおいて最高精度 Macro F1 スコアで 0.67 を達成した。次に、Random Forest を使用したモデルにおいて SHAP を適用することで、推定に寄与が大きい特徴量が、睡眠の深さであることを確認した。

次に、X-Means,Auto Encoder, PCA に基づく特徴量抽出手法を比較した。全ての特徴量抽出手法において精度の向上を確認し、X-Means を用いた特徴量抽出手法において最高精度 Macro F1 スコアで 0.79 を達成した。最後に、異常検知モデルを用いた識別システムを構築し、睡眠データを典型的な健康的睡眠/典型的な不健康睡眠に分離した。その後実施した識別実験において、最高で Macro F1 スコアで 0.91 を達成した。

データ収集の際に使用したマットレス型のセンサは市販品であり、医療専用機器と比較して安価であると言える。認知症の兆しに自分自身で気付くのは困難であり、また、認知機能低下の傾向に気が付き得る日常的に触れ合う第三者の存在は、必ずしもすべての人にとって可能ではない。ここに、人間の手を介さない認知症自動推定システムの必要性が存在し、この研究は、認知症と睡眠の関連に関する臨床的な研究の一部であると同時に、様々なモーダルを用いて研究されている認知症の自動検出システムの一部に、睡眠情報を用いることができることを示唆するものである。

目次

第1章	はじめに	1
第2章	関連研究	3
2.1	身体活動データから認知症を自動的に検出する	3
2.2	睡眠と認知症の関連の解析	3
2.3	関連研究と本研究の位置づけ	4
第3章	研究方法	5
3.1	データセット	5
3.1.1	認知症傾向推定の枠組み	5
3.1.2	認知症傾向の調査	6
3.1.3	マットレス型センサによる集計データ	7
3.1.4	MMSEによる推定ラベルで分けた参加者の統計量	8
3.2	LSTM,SVM,Random Forest を用いた機械学習モデル	10
3.2.1	LSTM を用いた機械学習モデル	13
3.2.2	SVM,Random Forest を用いた機械学習モデル	14
3.3	K-Means, PCA, Auto Encoder を用いた特徴量抽出	15
3.4	異常検知モデルを用いたデータ選別	17
第4章	実験	21
4.1	統計検定	21
4.2	LSTM,SVM,Random Forest を用いた機械学習モデル	23
4.2.1	LSTM を用いた実験	24
4.2.2	SVM ,Random Forest を用いた実験	24
4.2.3	実験結果	24
4.3	K-Means, PCA, Auto Encoder を用いた特徴量抽出の比較実験	25
4.3.1	実験結果	25
4.4	異常検知によるデータ選別	27
4.4.1	実験 1: 実験結果	27
4.4.2	実験 2: 実験結果	29

第5章 考察	30
5.1 LSTM,SVM,Random Forest を用いた機械学習モデル	30
5.2 K-Means, PCA, Auto Encoder を用いた特徴量抽出	31
5.3 異常検知によるデータ選別	31
5.3.1 理想的なデータセットを使用した実験	36
5.3.2 異常値ごとに異なるモデルを用いた実験	36
5.4 ラベルノイズの存在	36
第6章 おわりに	39
付録A データ	41

目次

3.1	マットレス型センサを用いた認知症傾向推定の枠組み	5
3.2	年齢 (x 軸) に対する MMSE 得点 (y 軸) の分布	6
3.3	Active Sleep ANALYZER	7
3.4	計測データ例 (標準化後)	8
3.5	総参加日数	9
3.6	総参加日数 (相対頻度)	10
3.7	Average act per participant	11
3.8	Average rrt per participant	11
3.9	Average hrt per participant	12
3.10	Average sas per participant	12
3.11	Average plm per participant	13
3.12	Overview:LSTM model	14
3.13	Overview:SVM, Random Forest model	14
3.14	概要: 複数の特徴量抽出手法の比較実験	15
3.15	個人ごとの特徴量作成手法:PCA, AE の場合	16
3.16	個人ごとの特徴量作成手法:X-Means の場合	16
5.1	ホテリング理論による異常値: ラベル無しグループ	32
5.2	ホテリング理論による異常値: 認知的に健康と推定されるグループ	32
5.3	ホテリング理論による異常値:MCI グループ	33
5.4	ホテリング理論による異常値: 認知症が疑われるグループ	33
5.5	LOF による異常値: ラベル無しグループ	34
5.6	LOF による異常値: 認知的に健康と推定されるグループ	34
5.7	LOF による異常値:MCI グループ	35
5.8	LOF による異常値: 認知症が疑われるグループ	35
A.1	睡眠時間 (分)	42
A.2	入眠時間 (分)	42
A.3	睡眠効率	43
A.4	中途覚醒時間 (分)	43
A.5	離床回数 (回)	44
A.6	在床時間 (分)	44

A.7 離床時間 (分)	45
A.8 日中睡眠時間 (分)	45
A.9 日中在床時間 (分)	46

表目次

3.1	計測した特徴量	8
4.1	act の統計量および S-W 検定の p 値	21
4.2	hrt の統計量および S-W 検定の p 値	21
4.3	rrt の統計量および S-W 検定の p 値	22
4.4	sas の統計量および S-W 検定の p 値	22
4.5	plm の統計量および S-W 検定の p 値	22
4.6	検定の結果	23
4.7	MMSE スコアの高/低分類タスクのマクロ F1 スコア	24
4.8	LSTM による各参加者の MMSE 得点の高/低の分類結果の混合行列	25
4.9	実験結果: 次元 10	26
4.10	実験結果: 次元 30	26
4.11	実験結果: 次元 50	26
4.12	実験結果: 次元 100	26
4.13	ホテリング理論によるデータ選抜実験 1(次元 10)	27
4.14	ホテリング理論によるデータ選抜実験 1(次元 30)	27
4.15	ホテリング理論によるデータ選抜実験 1(次元 50)	27
4.16	ホテリング理論によるデータ選抜実験 1(次元 100)	28
4.17	LOF によるデータ選抜実験 1(次元 10)	28
4.18	LOF によるデータ選抜実験 1(次元 30)	28
4.19	LOF によるデータ選抜実験 1(次元 50)	28
4.20	LOF によるデータ選抜実験 1(次元 100)	28
4.21	結果: データ選抜実験 2. 異常データ識別	29
4.22	結果: データ選抜実験 2. 正常データ識別	29
5.1	健康/認知症 識別精度比較	37
5.2	健康/その他 識別精度比較	37
5.3	健康/MCI 識別精度比較	37

第1章 はじめに

認知症は、一度正常に達した認知機能が後天的な脳の障害によって持続的に低下し、日常生活や社会生活に支障をきたすようになった状態と定義されている [1]. 社会の高齢化に伴い、認知症が社会に与える影響も増加が見込まれており、認知症患者は2050年までに世界中で1億5千万人に達すると推定されている [2].

認知症を発症する疾患は多数あるが、多くの場合根本的な治療法は発見されていない。従って、認知症に対する一般的な臨床治療は緩和ケアであるが、アルツハイマー病などの認知症を引き起こす代表的な疾患に対しては、早期発見により症状の進行の遅延を見込める疾患であることも報告されている [3]. この観点から、認知症の早期発見を目的とした研究は数多く行われており、とりわけ、ウェアラブルデバイスや計算機の実用発展に伴い、機械学習的なアプローチに基づいた、日常行動センシングからの認知症推定に関する研究や、歩行速度や身体活動の特徴に注目した認知症推定システムを提案するものや、睡眠中の活動量を活用して認知状態を予測に取り入れた研究が提案されている [4][5][6][7][8][9][10][11].

本研究では、認知症の早期発見を目的として、マットレス型のセンサから取得した睡眠中のデータのみから、認知症傾向を予測するシステムを構築することを目的とする。マットレス型センサからのデータ収集は参加者に与える負担が殆どなく、これによる継続的なデータ収集の容易さや、データ集計場所の普遍性など、認知症自動推定システムの社会への普及のために重要である利点を多数備えている。睡眠中の活動データから認知症推定を目指す本研究を支える医学的な根拠として、認知症は患者の睡眠パターンに影響を及ぼすことが統計的に知られていることが挙げられる。臨床的な観点から、認知症と睡眠障害の指標の関係を示す研究が報告されている [12]. 加えて、生理学的な観点からは、認知症の原因疾患と睡眠障害の関係を説明する研究が多く報告されている [13][14][15][16].

データ集計は、日本の4つの高齢者向け養護施設で行われた。参加者の睡眠中の活動データを計測するために、施設内の参加者のベッドにマットレス型のセンサを設置した。同時に、一部の参加者には認知機能の程度を推定するためのスクリーニングテストを実施し、それに基づくラベルを与えた。結果として、404名のラベルの付いた睡眠データ、およびラベルの付いてない681人の睡眠データを集計した。これを利用して認知症スケールの分類問題に取り組むことを目標に据えた。

一部の参加者に対して、Mini Mental State Estimation(MMSE)[17]と呼ばれる認知テストを実施した。この結果に基づいて、参加者を正常な認知状態にあるグループと、認知状態が低下しているグループ、およびMCI(Mild Cognitive Impairment)

と呼ばれる中間的な状態のグループに分類した。本研究では主だった実験が二つある。MMSEの結果によるラベルの付いたデータセットを用いて、一つ目のものでは、認知的に健康であると推定されるグループと認知症疑いがあるグループの二分類を予測するための機械学習モデルとして時系列解析モデル、およびランダムフォレストモデルとサポートベクトルマシン識別機を提案し、それらの結果を比較検討した。二つ目の実験では、一晩中の睡眠データから有効な特徴量を抽出することを目的として、複数の特徴量抽出手法を実施した。本実験でデータ集計のために用いたセンサは市販されている比較的安価なものであり、導入コストは少なく、ユーザーへの負荷も軽い。このような特徴を有するセンサは個人の睡眠の長期的モニタリングを可能とし、このセンサから得られた睡眠行動データから認知症傾向が推定可能であることが実証されると、段階的に変化する認知状態に付随する睡眠の様子の特ラッキングが可能となる。本研究の主な成果は以下の通りである:

認知症尺度を推定するための睡眠データセットの統計的な解析:

我々は、集計した睡眠中のデータを分析し、異なる認知状態にあると推定されるグループ間の、睡眠の様子の違いを統計的に調べた。この解析によって、認知症尺度の違いによって、参加者の睡眠中の心拍数、活動量、および周期性体動に統計的に有意な差があることを確かめた。

時系列解析モデルの推定への有効性の確認:

複数の機械学習モデルによる認知症尺度推定モデルを構築し、その結果、時系列解析モデルが最も有効であることを確認した。

クラスタリングに基づくデータ抽出手法の有効性の確認:

我々は、睡眠データに対し複数の特徴量抽出手法を適用し、個人ごとに特徴量ベクトルを作成した。それを認知症尺度推定モデルに適用し、Auto Encoderや主成分分析による特徴量抽出手法に比して、クラスタリングに基づくデータ抽出手法が有効であったことを確かめた。

第2章 関連研究

2.1 身体活動データから認知症を自動的に検出する

認知症の診断には、問診、身体検査、神経学的検査が行われ、認知症の有無とその症状が判断される。Jarroldら[5]は、音響的特徴や言語から音声から認知症のタイプを検出した。Hodgesら[6]は、ウェアラブルデバイス（RFID ブレスレット）とRFID タグの付いた物体を用いて、認知症などの認知機能障害の指標を検出した。Hayesら[18]は、モーションセンサーシステムを用いた活動特徴の抽出により、歩行速度、参加者の住居での日常活動量、軽度認知障害（MCI、高または低）のレベルとの関連性を調査した。Dawadiら[19]は、参加者のアクティビティを完了する能力と健康評価（認知症か認知的に健康か）の関係を調査した。Robbenら[20]は、参加者の住居でセンサデータを収集するアンビエントセンサモニタリングシステムを開発し、日常行動の変化を定量化するアルゴリズムを提案した。Matsubaraら[9][10]はアルツハイマー型認知症は概日リズムの調整機能を持つメラトニンの分泌を阻害するという事実に着目して、マットレス型センサから集計したデータから推定した概日リズムに基づく、アルツハイマー病検知システムを提案した。Khosroozadら[11]はアルツハイマー病に一般に見られる睡眠障害に着目し、マットレス型のセンサから集計したデータから推定した睡眠の質に基づく、アルツハイマー病認知症の初期段階の検知システムを提案した。

2.2 睡眠と認知症の関連の解析

生化学、神経学的な見地からのアルツハイマー型認知症と睡眠に関する研究は、[21][22][23][24]で詳細に議論されている。認知症と睡眠障害の関連を統計的に調べた研究として、[25][26]などがある。我々のアプローチと関連して、臨床的な観点から、機械学習を用いてアルツハイマー病と睡眠の関連を調べた研究には[27]があり、ここでは教師なし学習を用いてアルツハイマー病による認知状態の低下の識別に有用な睡眠に関連する特徴量に関する詳細な議論が展開されている。睡眠時の心拍データが、認知症の有効な周辺症状であることを示した研究として、[28]がある。この研究では、認知症の病状の変化により短期間に入退院を繰り返した一人の被験者の睡眠時の心拍データから推定した睡眠データから、認知症による認知状態の低下の検出に、睡眠データが有用であることが示されている。

2.3 関連研究と本研究の位置づけ

2.1 節で述べたように，認知機能に関連する様々な行動に着目し，認知機能を推定するシステムは多数提案されている．本研究では，2.2 節で紹介した，認知症と睡眠障害との医学的な関連を基礎として，睡眠データのみから認知症尺度を推定することを目的としている．睡眠データを活用した認知機能推定システムは複数報告されているが，本研究で使用するデータセットの量はいずれと比較しても膨大であり，それを背景とした時系列解析的アプローチや，大規模データ処理によるアプローチを本研究では主たる方法としている．

第3章 研究方法

本章では、マットレス型のセンサを用いて収集したデータおよび教師ラベルとしてのMMSEについて、また、それらを用いて実施する認知症尺度の二値分類タスクの手法について説明する。

3.1 データセット

3.1.1 認知症傾向推定の枠組み

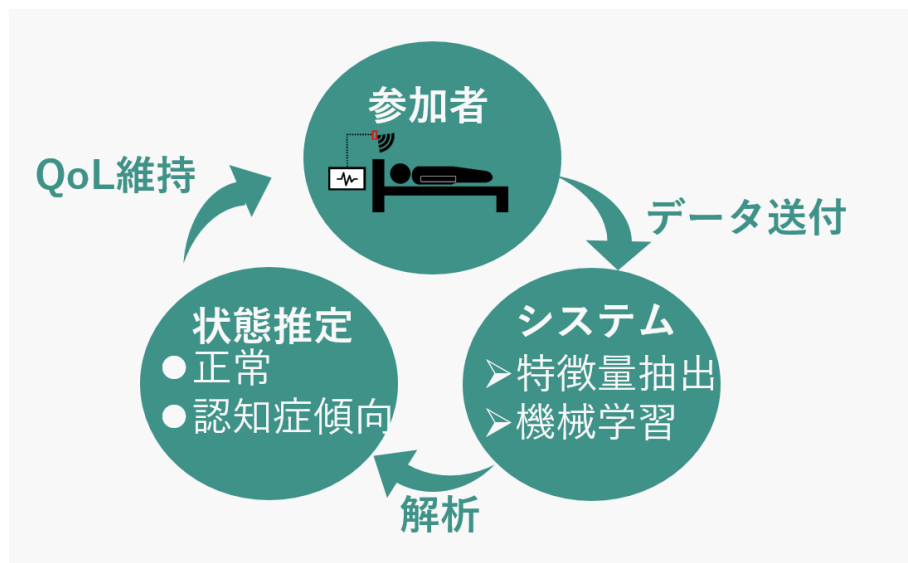


図 3.1: マットレス型センサを用いた認知症傾向推定の枠組み

本章における認知症傾向の目的、概要を図3.1に示す。この枠組みでは、使用者はマットレス型のセンサの上で睡眠をとる。センサから自動的に計測されるデータから認知症傾向を推定するシステムを構築し、使用者に推定される認知状態をフィードバックする。このシステムは使用者個人で完結して用いることが可能であり、推定は一度の睡眠ごとに行われる。

3.1.2 認知症傾向の調査

本実験では4か所の高齢者福祉施設を対象として集められた高齢者の睡眠中の行動データをデータセットとして使用する。データ集計は参加者によって1日から、長いもので2年強に渡る期間行われた。一部の参加者は年齢および性別の公開に同意し、その同意が得られた50名の平均年齢は87.8歳(±7.32歳)であり、そのうち8人が男性で42人が女性であった。

参加者の一部は、認知症スクリーニング検査であるMini Mental State Examination(MMSE)を彼らの居住する高齢者福祉施設において受けた。MMSEにおいては、28~30点が健康、23点以下が認知機能が低下していると判断される。それらの中間の24~27点は、Mild Cognitive Impairment(MCI)と呼ばれ、認知症とは言えないが、健康とも呼べない状態と判定される。図3.2は、年齢の公開に同意した参加者の、年齢に対するMMSEスコアをプロットしたものである。

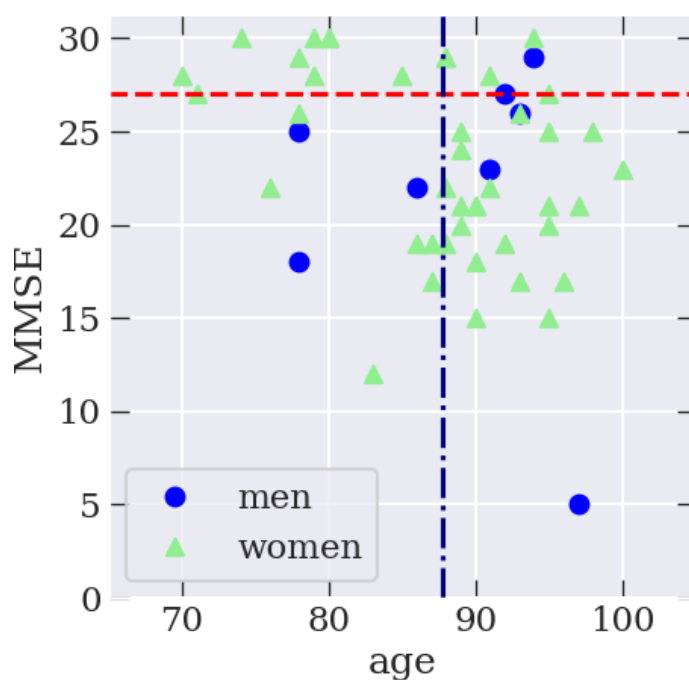


図 3.2: 年齢 (x 軸) に対する MMSE 得点 (y 軸) の分布

ここで注意すべき点として、MMSEは認知症の診断の際に世界標準的に使用される認知機能検査の一つであるが、この結果が認知症の診断結果と必ずしも符合するとは限らないというものがある。一般的に用いられている代表的な認知症の診断基準は米国精神医学会によるもの[29]や、WHOによるもの[30]などがあるが、そのいずれにおいても、単に記憶力の低下をもって認知症と診断するのではなく、それに伴う社会生活上の困難が生じていることも診断基準の基本的な要点に含ま

れている。つまり、認知症の診断は単なる認知機能検査だけをもって下されることはなく、生活機能検査の障害についての問診や、血液検査、脳のCT、MRIなどによる脳機能障害の可能性の排除や、原因疾患の特定などの流れを経て行われる。つまり、通常の診療過程において、MMSEなどの認知機能スクリーニングテストの結果のみで認知症と診断されることはない。従って本研究ではMMSEの低スコアラベルによって認知症であるとみなすことはなく、そのグループをMMSEによって認知症と推定されるグループと呼称する。

3.1.3 マットレス型センサによる集計データ

睡眠中の活動量データセットは、PARAMOUNT BED CO., LTD.製のマットレス型計測器であるActive sleep ANALYZER(図3.3)を使用して集計された。



図 3.3: Active Sleep ANALYZER

このマットレスの内部にはマイクロ波ドップラ方式センサ (78.0 × 24.5 × 1.8cm, 1kg) が使用されており、このセンサを用いてセンシングおよび睡眠中の行動の推定が実施される。測定対象とした睡眠行動は、ベッドでの睡眠状態と覚醒状態、ベッドからの離床回数、睡眠情報、一分あたりの心拍数、呼吸数、活動量、呼吸障害、および周期性体動の発生回数である。表 3.1. に、計測項目とその説明を示す。図 3.4. は、active sleep ANALYZER から推定された、ある被験者の一晩の心拍数、呼吸数、活動量のグラフである。

Active Sleep ANALYZER は一種の圧力センサであり、その計測原理は Cole-Kripke Algorithm[31] という腕時計型のデバイスから集計した活動量を従属変数とした覚醒/睡眠判定のための回帰アルゴリズムが元となっている。Cole-Kripke Algorithm は腕時計型の活動計から心拍数や呼吸数などを推定する手法の gold standard であり、Active Sleep ANALYZER から集計した活動量から心拍数などの測定項目を推定するアルゴリズムは Cole-Kripke Algorithm と同様の重回帰分析である。そ

表 3.1: 計測した特徴量

Feature	Detail
rrt	一分あたりの呼吸数の指標
hrt	一分あたりの心拍数の指標
act	一分あたりの活動量の指標
sas	一分あたりの呼吸障害イベントの指標
plm	一分あたりの周期性体動イベントの指標

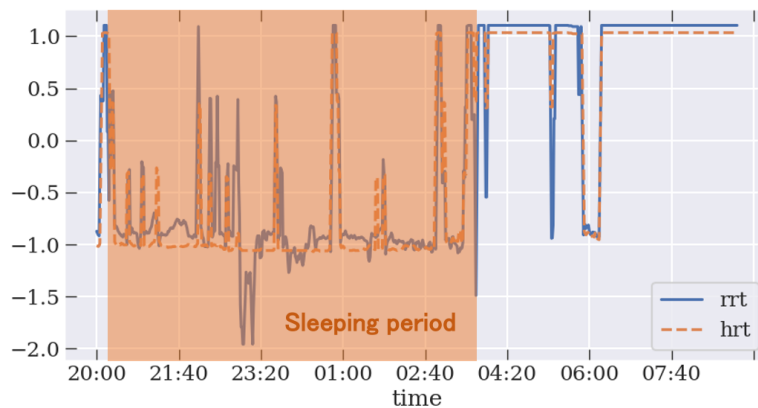


図 3.4: 計測データ例 (標準化後)

の精度は、腕時計型圧力センサと Cole-Kripke Algorithm の組み合わせと同程度であることが確認されている [31].

3.1.4 MMSE による推定ラベルで分けた参加者の統計量

本節では、MMSE による推定ラベルでグループ分けした参加者の統計的な情報について説明する。

参加者は全体で 1,085 人であり、そのうち、MMSE によるラベルが付いていない参加者が 681 人、MMSE により認知的に健康であると推定されたグループが 280 人、MCI の状態であると推定されたグループが 85 人、認知症であると推定されたグループが 39 人であった。図 3.5 にグループごとの参加日数の累計を、および図 3.6 にグループごとの総参加日数の相対頻度を示す。MMSE によるラベルの付いていない参加者全体の総参加日数が 65,668 日で最も多く、それに続いて、認知的に健康と推定されたグループの 26,312 日、MCI と推定されたグループの 14,248 日、認知症と推定されるグループの 11,166 日と続いた。この順序は参加人数の多い順であり、参加人数が多いグループの方が総参加日数が大きくなるのは順当であるが、グループごとに参加日数の傾向に特徴的な点があることが見て取れた。図 3.6

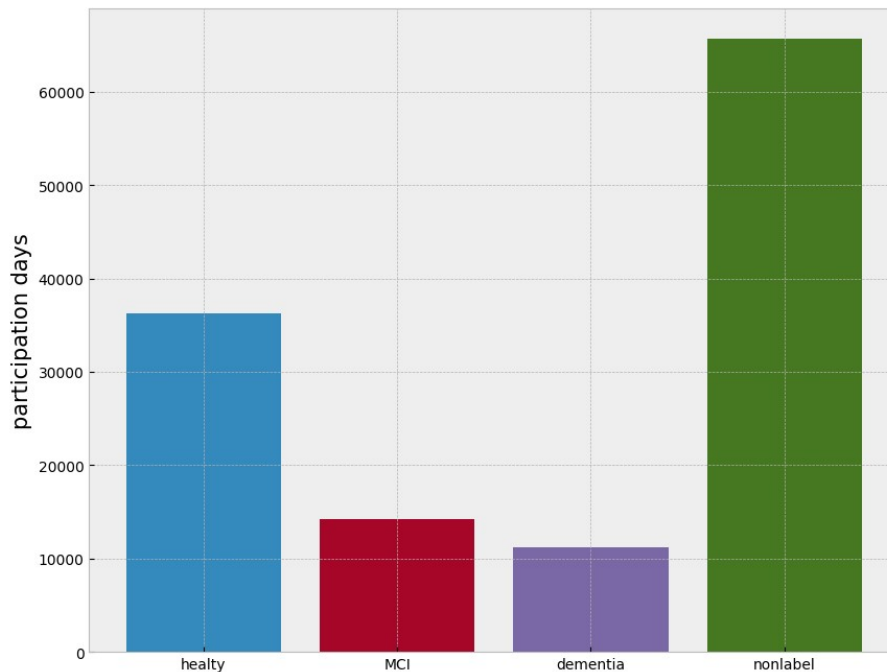


図 3.5: 総参加日数

を見ると、MMSE によるラベルの付いていないグループ、および認知的に健康と推定されたグループでは、参加日数が 100 日未満の参加者が最も多く、参加日数が 400 日以上の参加者は極めて少ない割合を占めることが分かる。一方で、MCI と推定されたグループでは、参加日数が 200 日を超える参加者もある程度存在し、認知症の疑いがあると推定されたグループでは、参加日数が 500 日弱の参加者が最多であった。本研究においてデータ収集は連続した期間において実施されたため、MMSE のスコアが低いグループほどデータ収集期間が長い参加者が多い理由として、MMSE スコアが低い参加者ほど介護を要する状態に割合が多く、そのような参加者は介護施設に滞在する期間も長くなるためだと推察される。また、MMSE スコアによるラベルが付いていないグループの参加者の参加日数の傾向が認知的に健康であると推定されたグループと近い理由は、認知的に健康である参加者は長期に渡って介護を要する状態にあることが比較的少ないため、MMSE を受けずに介護施設を出ることが多いからだと推察される。

図 3.7~3.10 に、MMSE のラベルが付いていないグループ、ラベルが付いているグループの、参加者毎の睡眠中の各計測項目 (表 3.1) の平均のプロット図を示す。これらの図から推察されることとして、

1. 認知的に健康と推定されるグループと MMSE によるラベルの付いていないグループの睡眠中の各計測項目の平均の傾向は近い。
2. 健康と推定されるグループから MCI、認知症と推定されるグループと MMSE

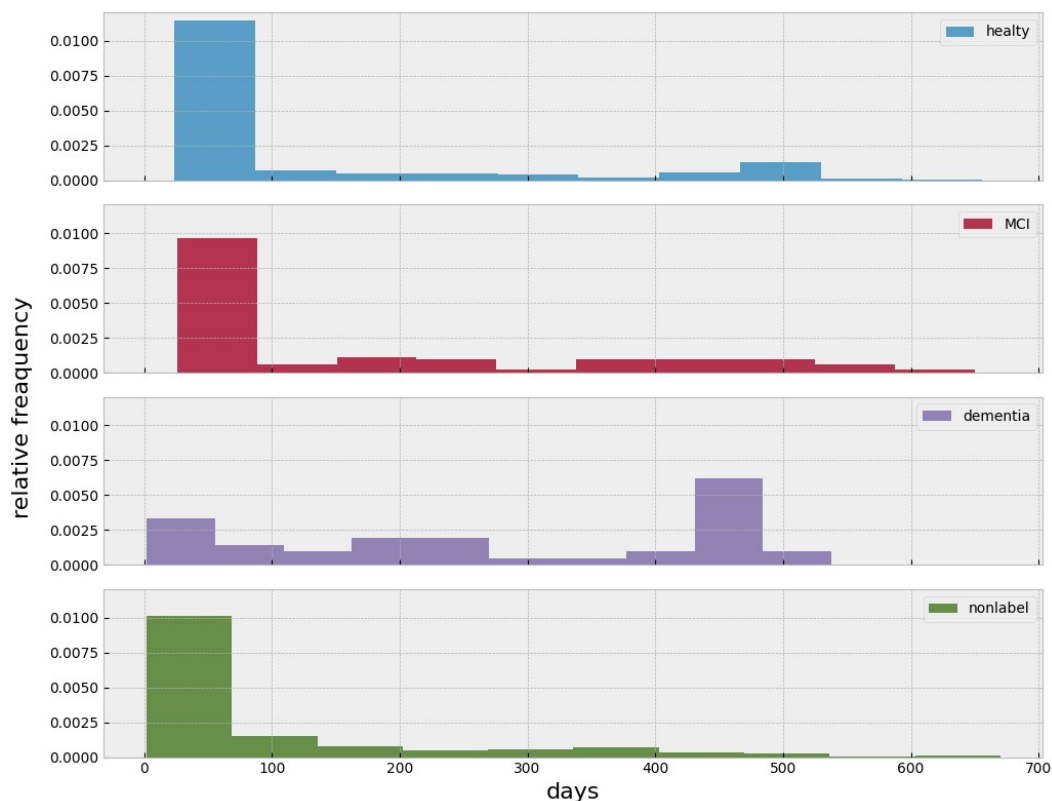


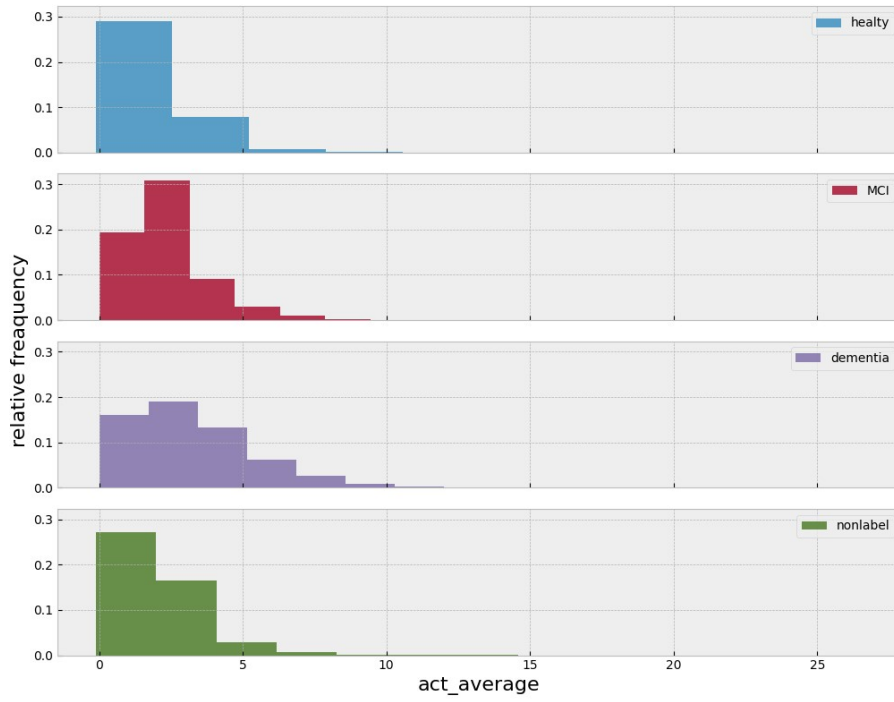
図 3.6: 総参加日数 (相対頻度)

スコアが低くなるほど，睡眠中の各計測項目の平均はなだらかに連続的に変化する。

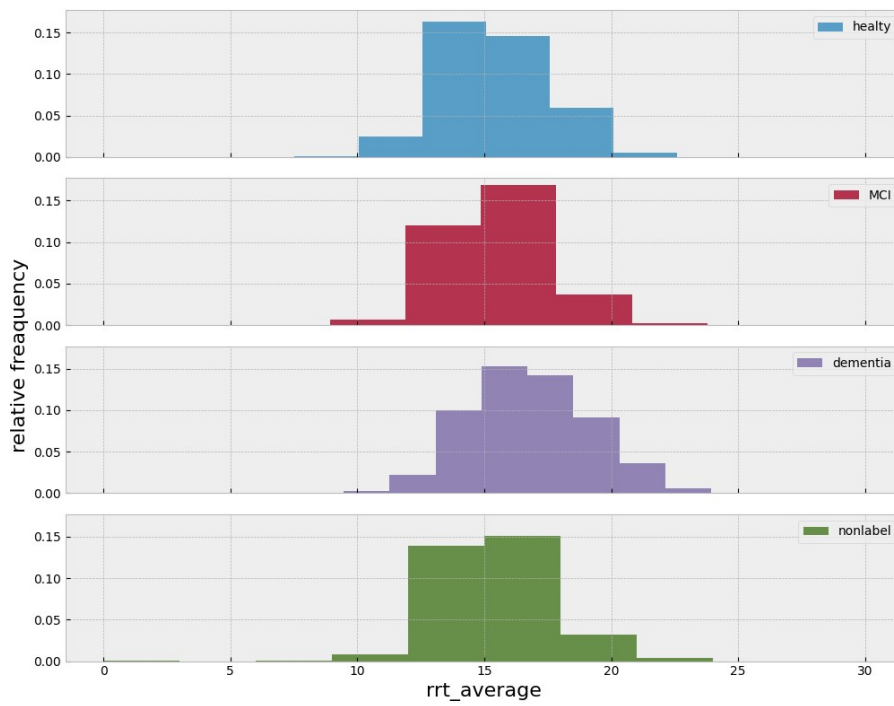
ことが挙げられる．[25][26]で示されている通り，認知症と睡眠障害には関連があることが統計的に確認されており，図 3.7~3.10 に示した本研究において計測した項目の睡眠中の平均値の図は，この統計的な差異を示していると推察される．これを確認するために，統計解析の手法を用いて，グループ間の計測項目の平均に有意な差があるかを確認する．

3.2 LSTM,SVM,Random Forest を用いた機械学習モデル

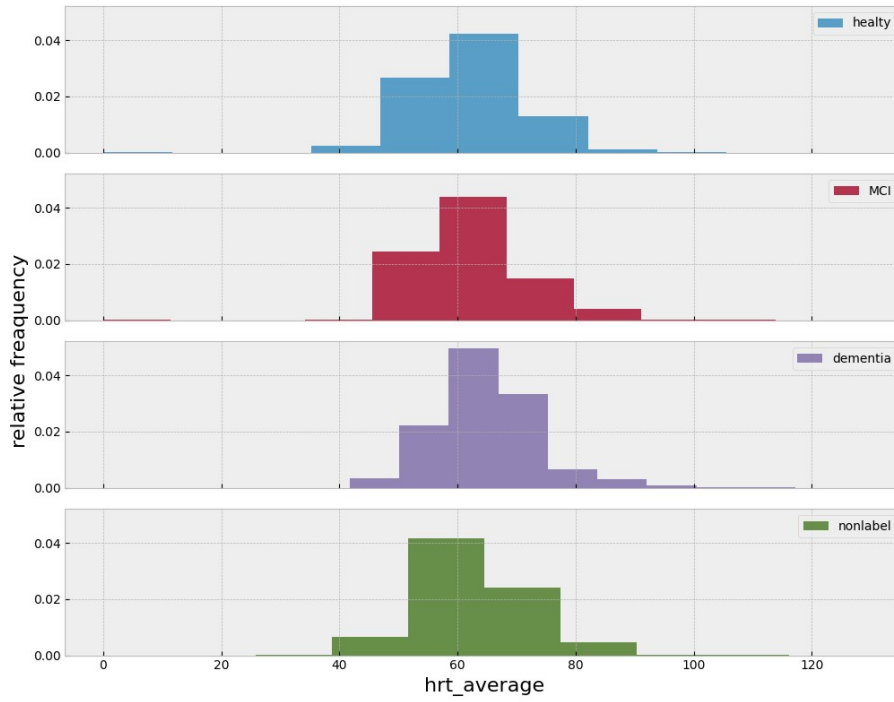
この節では時系列モデルである LSTM[32]，および Support Vector Machine[33]，Random Forest[34] を使用した特徴量抽出手法について説明する．



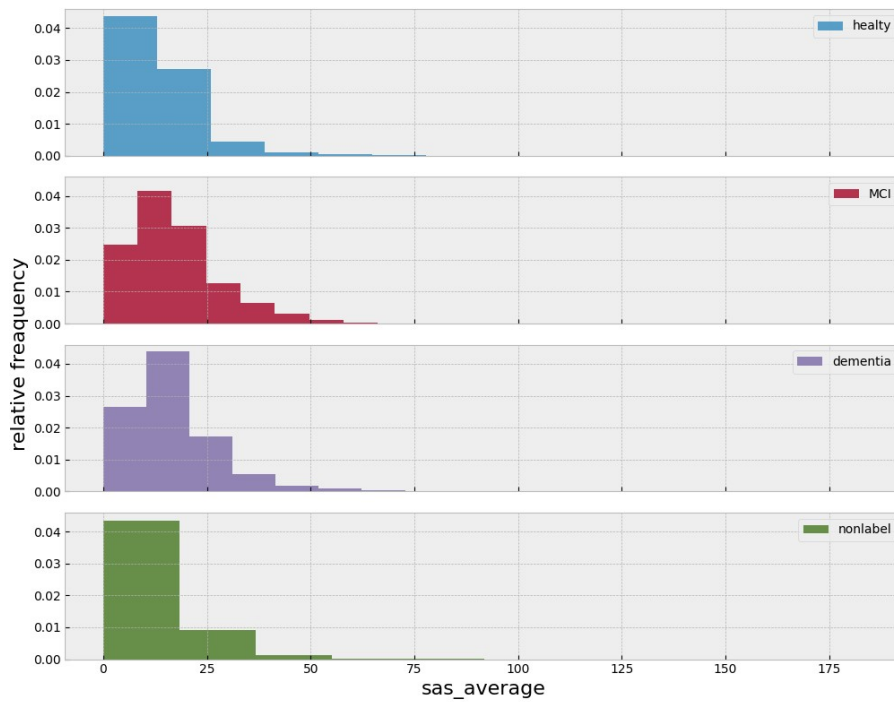
☒ 3.7: Average act per participant



☒ 3.8: Average rrt per participant



☒ 3.9: Average hrt per participant



☒ 3.10: Average sas per participant

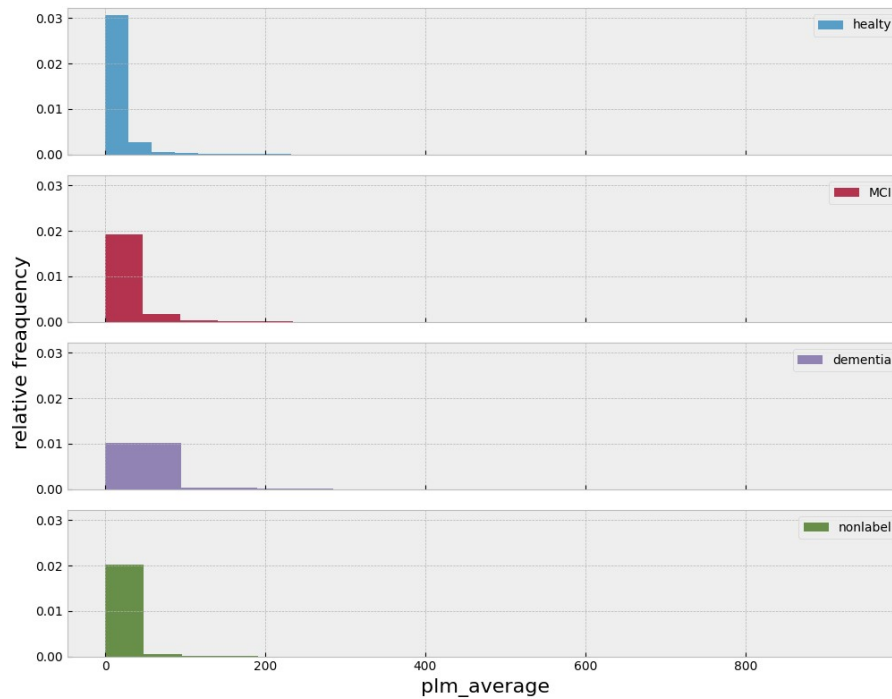


図 3.11: Average plm per participant

3.2.1 LSTM を用いた機械学習モデル

計測されたデータは時系列データであり，そこに注目した特徴量抽出を LSTM を用いて実施する．特徴量抽出の流れは以下のとおりである．

1. 参加者の一日分の睡眠データの標準化を行う．この際，異なる参加者のデータを用いて標準化されることがないように，同じ参加者の同じ日付のデータのみを用いて標準化する．
2. 得られた標準化された一日分のデータを，2層の LSTM を用いてエンコードする．
3. 上記の工程を 1 人の参加者に対して計測日数分行う．
4. 得られる計測日数分の LSTM の最終状態量ベクトルを平均プーリングして，線形識別層へ入力し，認知症尺度の 2 値分類を行う

特徴量抽出レイヤである LSTM と識別層である線形識別器は連結されており，誤差逆伝播による学習は同時に行われる．

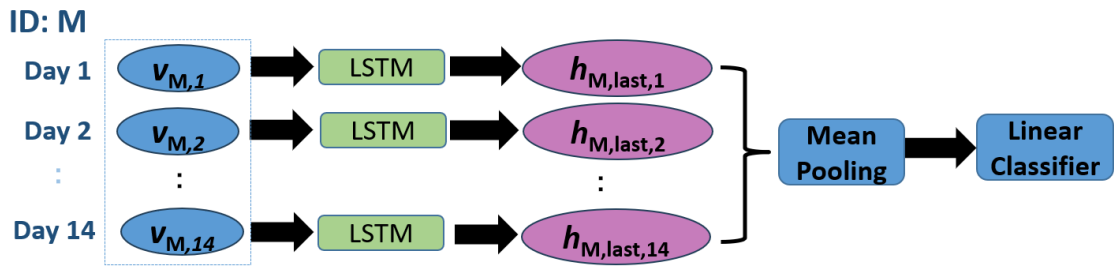


図 3.12: Overview:LSTM model

3.2.2 SVM,Random Forest を用いた機械学習モデル

SVM と RF には、1 人の参加者の認知症スケールの 2 分類を予測するために、その参加者のすべての睡眠データから計算される各計測項目の平均値や分散などの統計値を入力として用いる。その概要を図 3.13 に示す。特徴量抽出の手順は以下の通りである。

1. それぞれの参加者がもつ 14 日分のデータに対して、そのそれぞれの一晩のデータから、各種計測項目の平均値、分散を抽出する。
2. 計測日数分あるそれぞれのベクトルを結合することで、1 人に対して一つの特徴量を作成する。
3. その特徴量を Random Forest,SVM に入力し、認知症尺度の 2 値分類を行う

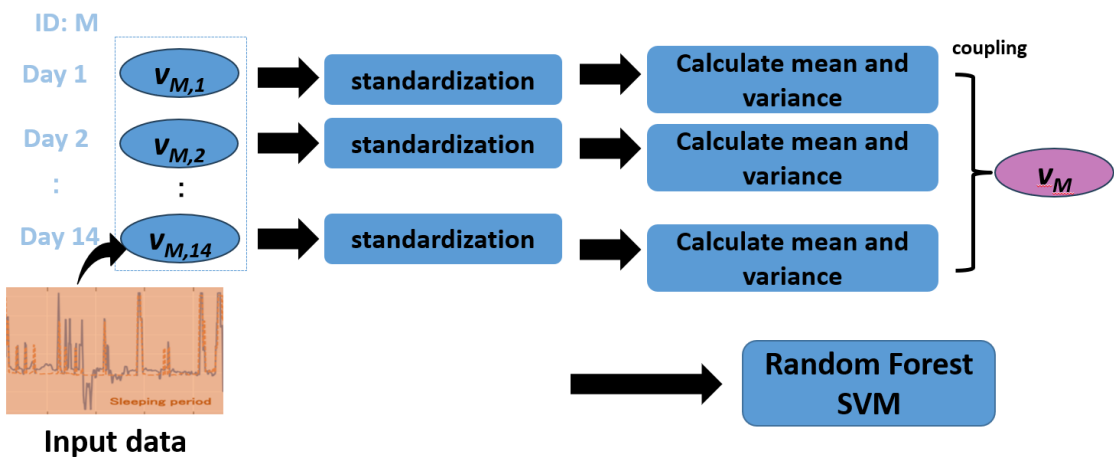


図 3.13: Overview:SVM, Random Forest model

3.3 K-Means, PCA, Auto Encoder を用いた特徴量抽出

この節では、K-Means, PCA, Auto-Encoder を用いた特徴量抽出手法について説明する。

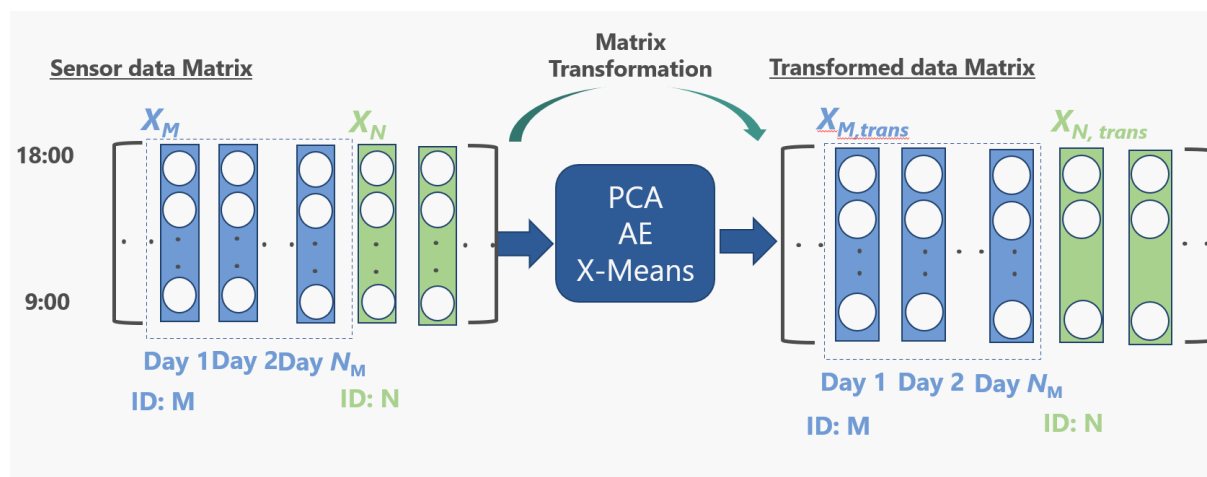


図 3.14: 概要: 複数の特徴量抽出手法の比較実験

この手法では、[7]を参考に特徴量手法を実施する。[7]では、高齢者介護施設に設置された睡眠センサ、ドアセンサ、人感センサから収集された行動データから認知症尺度を推定するために、行動パターンに基づく特徴量抽出を実施した。これは以下の手順によって実施される。

1. 一分間のセンサーの測定値に基づき、18:00-9:00 までのセンサー測定値を計算する。
2. センサー計測値に対して、睡眠パターンに対応した睡眠特徴量を抽出する。
3. 1日の行動特徴量を個人ごとの特徴量に変換する。

睡眠特徴量抽出の手法として、次元削減アルゴリズムである主成分分析 (PCA), オートエンコーダー (AE), X-Means[35] を実施した。PCA は線形の次元削減手法であり、データ間に線形的な関係が強いとき、効果的な特徴量抽出が可能になることが期待できる。オートエンコーダーはニューラルネットワークを使用した非線形の次元削減手法であり、データ間に非線形的な関係が強いとき、識別に有用な特徴量を抽出することが期待できる。X-means はクラスタリング手法の一種であり、k-Means と異なる点として、k-Means では初めからクラスタの数を設定する必要があるが、X-means では自動的に最適なクラスタ数を発見するアルゴリズムであるという点である。本研究では、一晩あたりのデータを単位データとし、全て

のデータを X-Means を用いてクラス分類し，それに基づき参加者一人あたりの特徴量を作成する。

これらの手法を実施するために，全てのセンサー測定値を連結し，データ行列を作成する．この行列の形状は(参加者全員の各日の計測時間の計測値の全て×全参加者の参加日数の総和)となっている．この睡眠データ行列に対し前述の次元削減手法およびクラスタリング手法を実施することで，睡眠パターンに対応した一日ごとの典型睡眠ベクトル集合を求める．その概要を図 3.14 に記す。

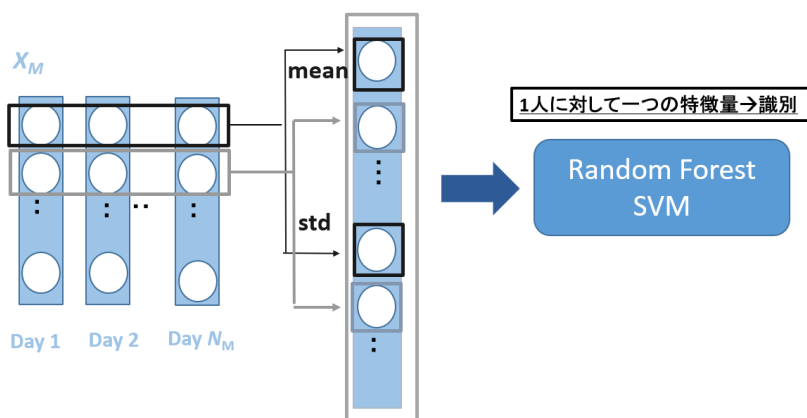


図 3.15: 個人ごとの特徴量作成手法:PCA, AE の場合

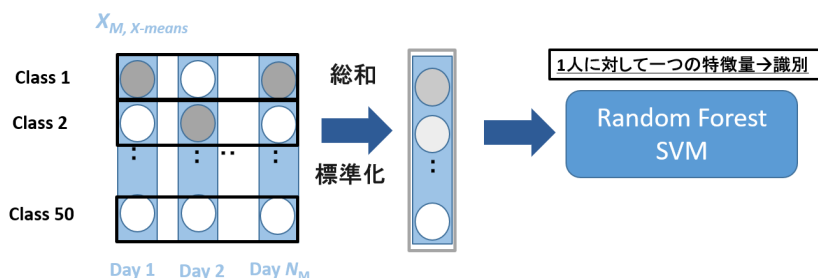


図 3.16: 個人ごとの特徴量作成手法:X-Means の場合

各々の特徴量抽出手法を実施し次元数が削減された毎晩の睡眠データベクトルは，その後個人を識別する id に基づいて集計され，統計処理により個人に対し一つの代表ベクトルを形成する．その概要を図 3.15,3.16 に示す。

AE 特徴量，PCA 特徴量に対しては，参加者一人の全ての変換後の睡眠データベクトルの全ての要素の平均と分散を取ることで個人を代表する特徴量を作成する．X-Means 特徴量に対しては，どこかのクラスに分類される一晩の睡眠データベクトルを，そのベクトルが所属するクラスに対応する次元のみ 1，その他の要素を 0 とする one-hot ベクトルに変換し，その後，参加者一人の全ての変換後の睡

眠データベクトルの、同一要素の平均をとることで、個人を代表する特徴量を作成する。

このように、それぞれの特徴量抽出手法を経て構成された個人ごとの特徴量、および比較のために作成する、単に平均と分散からなる統計特徴量を SVM および Random Forest に適用する。

3.4 異常検知モデルを用いたデータ選別

この節では、異常検知モデルを用いたデータ選別手法について説明する。本研究で使用するデータには MMSE の結果による認知状態の推定ラベルが付いているものと、ラベルが付いていないものが存在する。ラベルには、認知的に健康、MCI、認知症の疑い、とそれぞれ推定されるものの3種類があるが、識別タスクの精度を高めるためには、この3つのグループにはそれぞれ異なる典型的な形態があることが望ましい。認知症は睡眠の様態に影響を与えることは生理学的にも ([21][22][23][24]) 臨床的にも ([25][26]) 知られているが、個人の睡眠の様態は認知症を引き起こす疾患からだけでなく、生活習慣や身体的特性など、個人を取り巻く多様な要素から影響を受ける。従って睡眠データから、認知症が睡眠に与える本質的な影響を抽出するのは困難であり、MMSE ラベルにより分けられたグループの、典型的な睡眠形態を見つけることは必ずしも簡単ではない。参加者の中には、認知的には健康であるが睡眠の様態は認知症患者に近い個人や、認知症が疑われるが睡眠の様態としては健康な個人と一見変わらない個人も存在する。そのような参加者を発見し、認知症自動推定のシステムにおいて適切に扱うために、異常検知モデルを用いたデータ選別を実施した。

使用した異常検知モデルはホテリングの T^2 乗法および Local Outlier Factor (LOF) 法 [36] である。

ホテリング理論では、 M 次元正規分布 $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ から N 個の独立標本に基づき、標本平均 $\hat{\boldsymbol{\mu}}$ を式 (3.1) で、標本共分散 $\hat{\boldsymbol{\Sigma}}$ を式 (3.2) で定義する。そして、負の対数尤度 $-\ln \mathcal{N}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ の2倍をもとに、異常度を式 (3.3) で定義する [37]。

$$\hat{\boldsymbol{\mu}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}^n \quad (3.1)$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}^n - \hat{\boldsymbol{\mu}})(\mathbf{x}^n - \hat{\boldsymbol{\mu}})^\top \quad (3.2)$$

$$a(\mathbf{x}') = (\mathbf{x}^n - \hat{\boldsymbol{\mu}})^\top \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}^n - \hat{\boldsymbol{\mu}}) \quad (3.3)$$

$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ から新たな標本 \mathbf{x}' を独立に観測したとき、以下が成立する。

1. $\mathbf{x}' - \hat{\boldsymbol{\mu}}$ は、平均 $\mathbf{0}$ 、共分散 $\frac{N+1}{N} \boldsymbol{\Sigma}$ の M 次元正規分布に従う。

2. $N\hat{\Sigma}$ は, $\mathbf{x}' - \hat{\boldsymbol{\mu}}$ と統計的に独立に, 自由度 $N - 1$, スケール行列 Σ の M 次元ウィシャート分布に従う.
3. 式 (3.3) から作られる統計量 $\frac{N-M}{(N+1)M}a(\mathbf{x}')$ は, 自由度 $(M, N - M)$ の F 分布に従う.
4. $N \gg M$ の場合には, $a(\mathbf{x}')$ は, 近似的に, 自由度 M , スケール因子 1 のカイ二乗分布に従う.

この研究において, データの次元数 M は表 3.1 に示した 5 であり, 一サンプルを一人の参加者とした場合, サンプル数 N は全参加者数の 1,085 である. 従って $N \gg M$ が成立する. この近似が成立する場合のホテリングの T^2 乗法を使用した異常検知の手順は以下である.

1. パーセント値 α を決め, カイ二乗分布から

$$1 - \alpha = \int_0^{\alpha_{th}} du \chi^2(u|M, 1) \quad (3.4)$$

を計算し, 閾値 α_{th} を求める.

2. 正常分布が圧倒的多数を占めると信じられるデータから標本平均 (3.1) および標本分散 (3.2)

$$\hat{\boldsymbol{\mu}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}^n$$

$$\hat{\Sigma} = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}^n - \hat{\boldsymbol{\mu}})(\mathbf{x}^n - \hat{\boldsymbol{\mu}})^\top$$

を計算する.

3. 新たな観測値 \mathbf{x}' を得るたび, 異常度としてのマハラノビス距離 (3.3)

$$a(\mathbf{x}') = (\mathbf{x}^n - \hat{\boldsymbol{\mu}})^\top \hat{\Sigma}^{-1} (\mathbf{x}^n - \hat{\boldsymbol{\mu}})$$

を計算する.

4. $a(\mathbf{x}') > a_{th}$ なら, 異常判定をする.

LOF法における基本的なアイデアは、異常度を計算する各サンプルの k 個近傍点との距離をもとに、異常度 $a_{LOF}(\mathbf{x}')$ を

$$a_{LOF}(\mathbf{x}') = \frac{1}{k} \sum_{\mathbf{x} \in N_k(\mathbf{x}')} \frac{d_k(\mathbf{x}')}{d_k(\mathbf{x})} \quad (3.5)$$

で定義する [37]. ただし, $d_k(\mathbf{u})$ は一般に近傍有効距離を u の周りの k 近傍に渡り平均したもので,

$$d_k(\mathbf{u}) = \frac{1}{k} \sum_{\mathbf{u}' \in N_k(\mathbf{u})} l_k(\mathbf{u} \rightarrow \mathbf{u}') \quad (3.6)$$

と定義される. ただし, 近傍距離 $l_k(\mathbf{u} \rightarrow \mathbf{u}')$ は以下で定義される.

$$l_k(\mathbf{u} \rightarrow \mathbf{u}') \equiv \begin{cases} e_k(\mathbf{u}') & (\mathbf{u}' \in N_k(\mathbf{u}) \text{ かつ } \mathbf{u} \in N_k(\mathbf{u}')) \\ d(\mathbf{u}, \mathbf{u}') & (\text{それ以外の全ての場合}) \end{cases} \quad (3.7)$$

本研究においては, 距離 $d(\mathbf{u}, \mathbf{u}')$ はミンコフスキー距離を用いた.

本研究において異常度を算出するための標本平均, 標本分散を使用するために使用したデータは, 図 3.7~3.10 に示した参加者の睡眠中の各計測項目の平均のうち, ラベルが付いていないものである. 異常検知によるデータ選別を施す枠組みでは, データ選別の後段に識別タスクがある. 識別タスクにおいてはラベルの付いたデータしか用いることが出来ないため, データ選別のための異常度算出にラベルの付いたデータを用いると, リーキングの観点から適切ではない. MMSE によるラベルの付いていないデータは, 睡眠中の各計測項目や参加日数の傾向から, 認知的に健康である参加者が多いと推定され, ラベル無しデータに異常検知の手法を用い, ラベル付きのデータの異常度を算出するための前提条件の一つである, 圧倒的多数が正常であると信じられるというものを満たしていると考えた. データ選別を実施する実験は以下の 3 種類である. それぞれの結果が 3.3 で説明する, 何らのデータ選別を実施しない手法の結果と比較される. このいずれにおいても, 使用する機械学習モデルは 3.3 のものと同じである.

実験 1

正常と判定された健康推定グループのデータ, および異常と判定された認知症または MCI と推定されるグループのデータを, 学習にもテストにも使用する.

実験 2

正常と判定された健康推定グループのデータ, および正常と判定された認知

症または MCI と推定されるグループのデータで学習しモデルを作る。また、異常と判定された健康推定グループのデータ、および異常と判定された認知症または MCI と推定されるグループのデータで学習しモデルを作る。テストデータは選別せず、識別タスクの初段階において正常/異常判定を実施し、それに基づく適切なモデルに接続し識別する。

第4章 実験

この章では、3.1で説明した使用したデータセットに対して実施した統計解析について説明する。その後、3.2,3.3,3.4節において説明した実験の設定および結果について説明する。

4.1 統計検定

異なる MMSE ラベルが付いたグループ間の計測データ間に統計的な有意差があるかを調べるために、統計的検定を実施した。すべてのデータは Python のライブラリである Scipy(Ver:1.10.1) を用いて分析し、平均値(±標準偏差)で示した。すべての計測項目は、まずシャピロ・ウィルク検定(S-W 検定)を用いて正規分布に従っているかを確認した。その結果を表 4.1~4.5 に示す。有意水準は全て 5%未満とした。

表 4.1: act の統計量および S-W 検定の p 値

MMSE ラベル	平均(±分散)	S-W 検定の p 値
認知的に健康 (N=280)	1.85(±0.805)	4.09×10^{-11}
MCI(N=85)	2.21(±1.04)	4.79×10^{-4}
認知症疑い (N=39)	3.09(±1.61)	4.37×10^{-3}

表 4.2: hrt の統計量および S-W 検定の p 値

MMSE ラベル	平均(±分散)	S-W 検定の p 値
認知的に健康 (N=280)	62.0(±6.80)	3.00×10^{-3}
MCI(N=85)	62.9(±8.22)	5.41×10^{-2}
認知症疑い (N=39)	65.3(±7.42)	1.53×10^{-1}

検定の結果,

- act は、認知症が疑われるグループのみ正規分布であり、認知的に健康と推定されるグループと MCI は正規分布ではない。

表 4.3: rrt の統計量および S-W 検定の p 値

MMSE ラベル	平均 (± 分散)	S-W 検定の p 値
認知的に健康 (N=280)	15.2(±1.86)	4.73×10^{-3}
MCI(N=85)	15.3(±2.16)	3.42×10^{-1}
認知症疑い (N=39)	16.6(±2.30)	7.61×10^{-1}

表 4.4: sas の統計量および S-W 検定の p 値

MMSE ラベル	平均 (± 分散)	S-W 検定の p 値
認知的に健康 (N=280)	12.5(±6.40)	3.30×10^{-14}
MCI(N=85)	14.6(±7.92)	4.08×10^{-5}
認知症疑い (N=39)	16.8(±8.97)	1.55×10^{-4}

- hrt は、MCI と認知症が疑われるグループにおいて正規分布であり、認知的に健康と推定されるグループは正規分布ではない。
- rrt は、MCI と認知症が疑われるグループにおいて正規分布であり、認知的に健康と推定されるグループは正規分布ではない。
- sas は、全てのグループにおいて正規分布ではない。
- plm は、全てのグループにおいて正規分布ではない。

と判断した。この結果を踏まえて、

- act において、認知的に健康と推定されるグループと MCI に有意差があるかマン・ホイットニーの U 検定
- hrt は、MCI と認知症が疑われるグループに有意差があるか、F 検定およびスチューデントの t 検定
- rrt は、MCI と認知症が疑われるグループに有意差があるか、F 検定およびスチューデントの t 検定
- sas は、群間の有意差を確認するため、スティーラ・ドゥワス検定

表 4.5: plm の統計量および S-W 検定の p 値

MMSE ラベル	平均 (± 分散)	S-W 検定の p 値
認知的に健康 (N=280)	11.3(±11.5)	4.90×10^{-21}
MCI(N=85)	16.7(±19.3)	1.51×10^{-11}
認知症疑い (N=39)	22.8(±22.4)	3.94×10^{-6}

- plm は、群間の有意差を確認するため、スティール・ドゥワス検定を実施した。その結果を表 4.6 に示す。

表 4.6: 検定の結果

検定名	特徴量	グループ	結果
マン・ホイットニーの U 検定	act	認知的に健康,MCI	有意差なし
F 検定	hrt	MCI, 認知症疑い	等分散
F 検定	rrt	MCI, 認知症疑い	等分散
スチューデントの t 検定	hrt	MCI, 認知症疑い	有意差なし
スチューデントの t 検定	rrt	MCI, 認知症疑い	有意差あり
スティール・ドゥワス検定	plm	全てのグループ	健康および認知症推定群で有意差あり
スティール・ドゥワス検定	sas	全てのグループ	健康および認知症推定群で有意差あり

以上の結果をまとめると、

- act において、認知症が疑われるグループのみ正規分布であり、認知的に健康と推定されるグループと MCI グループに有意差はない。
- hrt は、認知症が疑われるグループと MCI のグループは有意差のない正規分布であり、認知的に健康と推定されるグループは正規分布ではない。
- rrt は、認知症が疑われるグループと MCI のグループは有意差のない正規分布であり、認知的に健康と推定されるグループは正規分布ではない。
- sas は、全てにおいて正規分布ではなく、認知的に健康と推定されるグループと認知症が疑われるグループに有意差がある。
- plm は、全てにおいて正規分布ではなく、認知的に健康と推定されるグループと認知症が疑われるグループに有意差がある。

となる。

4.2 LSTM,SVM,Random Forest を用いた機械学習モデル

この節では 124 名の参加者のデータを用いて実施する、時系列モデルである LSTM[32], および Support Vector Machine[33], Random Forest[34] を使用した特徴量抽出手法について説明する。MMSE によって 124 名の被験者のうち 61 人が認知的に健康と推定されるグループ、20 人が MCI, 43 人が認知症が疑われるグループと判定された。この実験においては、61 人の認知的に健康と推定されるグループを MMSE スコアの高群、MCI および認知症が疑われるグループを MMSE スコアの低群とし、2 クラス識別タスクを実施した。

4.2.1 LSTMを用いた実験

計測されたデータは時系列データであり，そこに注目した特徴量抽出を LSTM を用いて実施した．データセットを有効に使うために，20Fold の Stratified Cross Validation を実施し，各 Fold の Test set の結果の平均を全体の結果として採用した．また，training set 内に別途 validation set を設けた．この validation set に対して，過学習を抑制するために 30 を閾値とする early stopping を実施した．ここで，test set で結果を採用する epoch は，最後の epoch から 30 を引いた，early stopping がカウントされる直前の epoch とした．LSTM の内部次元数は 32 であり，dropout ratio は 0.75 に設定した．

4.2.2 SVM ,Random Forest を用いた実験

LSTM を使用した実験と同様に，20Fold の Stratified Cross Validation を実施した．この際，nest した 20Fold の crossvalidation を実施することでグリッドサーチを実施し，それぞれのモデルにおいて最適なパラメータを探索した上で結果を取得した．SVC のカーネルには，RBF を採用し，コストパラメータ C と γ についてグリッドサーチを実施した．その探索範囲については， C と γ の両者とも [0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000] である．ランダムフォレストについては，分割に使用するフィーチャー数と作成するツリーの数についてグリッドサーチを実施し，その探索範囲はフィーチャー数については [1, 2, 3, 4, 5] で，ツリーの数については [100, 1000] である．

4.2.3 実験結果

LSTM,SVM,Random Forest を用いた機械学習モデルの実験結果を表 4.7,4.8 に示す．評価指標として macro-F1 スコアを採用し，クロスバリデーションの各 fold の最終結果の平均およびその分散を示した．

表 4.7: MMSE スコアの高/低分類タスクのマクロ F1 スコア

Model	Macro F1 score(\pm STD)
LSTM	0.67(\pm 0.25)
Random Forest	0.63(\pm 0.22)
Support Vector Machine	0.62(\pm 0.25)

LSTM を採用したモデルにおいて，最高精度の平均 macro-F1 スコア 0.67 を達成した．

表 4.8: LSTM による各参加者の MMSE 得点の高/低の分類結果の混合行列

		Estimated	
		High	Low
Actual	High	37	23
	Low	16	45

4.3 K-Means, PCA, Auto Encoder を用いた特徴量抽出の比較実験

この実験において使用したデータセットには、MMSE テストにおいて健常と判定された参加者が 279 人、MCI と判定された参加者が 85 人、認知症疑いと判定された参加者が 43 人が含まれている。表 3.1 に示した各測定項目の 18:00-9:00 までの全てのセンサー測定値を一時間ごとに平均をとったものを解析データ行列として使用した。元々のデータ行列の行数、つまり一晩あたりのデータベクトルの次元数は次元数 7×15 時間 = 135 次元であり、これを X-Means, PCA, AutoEncoder で次元削減した。識別モデルとして、SVC と Random Forest を使用し、両モデルにおいて、4.2 節と同様に、nest した Cross Validation によりグリッドサーチを実施した。パラメータの探索範囲も同様に設定した。オートエンコーダーは 3 層の線形層および活性化層としての ReLu から構成された。

4.3.1 実験結果

それぞれの手法において次元を 10,30,50,100 に圧縮し、SVC と Random Forest で 2 クラス分類実験を実施した。識別タスクは、認知的に健康と推定されるグループと MCI または認知症が疑われるグループを識別するタスク、認知症が疑われるグループと認知的に健康と推定されるグループまたは MCI を識別するタスク、認知的に健康と推定されるグループと認知症が疑われるグループを識別するタスク、認知的に健康と推定されるグループと MCI を識別するタスク、MCI と認知症が疑われるグループを識別するの 5 種類存在し、評価指標は macro-F1 スコアとした。20Fold の nest したクロスバリデーションを実施し、テストスコアの平均を最終結果とした。実験結果を表 4.9~4.12 に示す。

表 4.9: 実験結果: 次元 10

特徴量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.598	0.692	0.737	0.542	0.562
	RF	0.585	0.598	0.718	0.555	0.614
Auto-Encoder	SVC	0.536	0.654	0.693	0.452	0.702
	RF	0.578	0.678	0.700	0.442	0.738
PCA	SVC	0.594	0.720	0.741	0.507	0.639
	RF	0.637	0.608	0.661	0.506	0.675

表 4.10: 実験結果: 次元 30

特徴量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.558	0.677	0.671	0.519	0.708
	RF	0.568	0.669	0.741	0.499	0.680
Auto-Encoder	SVC	0.598	0.672	0.673	0.502	0.740
	RF	0.549	0.512	0.680	0.468	0.667
PCA	SVC	0.577	0.714	0.709	0.540	0.720
	RF	0.617	0.559	0.708	0.499	0.690

表 4.11: 実験結果: 次元 50

特徴量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.560	0.713	0.795	0.542	0.694
	RF	0.592	0.612	0.691	0.509	0.707
Auto-Encoder	SVC	0.593	0.780	0.806	0.534	0.534
	RF	0.628	0.629	0.699	0.515	0.676
PCA	SVC	0.586	0.725	0.748	0.510	0.728
	RF	0.584	0.523	0.648	0.512	0.655

表 4.12: 実験結果: 次元 100

特徴量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.555	0.729	0.757	0.493	0.745
	RF	0.558	0.621	0.721	0.471	0.652
Auto-Encoder	SVC	0.579	0.634	0.753	0.477	0.709
	RF	0.551	0.587	0.663	0.444	0.621
PCA	SVC	0.533	0.711	0.753	0.499	0.690
	RF	0.567	0.569	0.682	0.453	0.628

4.4 異常検知によるデータ選別

4.4.1 実験1: 実験結果

MMSEによるラベルの付いていないデータの圧倒的多数が正常であると仮定したとき、異常判定の枠組みで正常であると判定されるデータは認知的に健康であり、異常と判定されるデータはMCIまたは認知症の疑いがあると期待される。実験1では、モデルの学習に用いるデータとモデル評価に用いるテストデータの両者に、正常と判定された健康推定グループのデータ、および異常と判定された認知症またはMCIと推定されるグループのデータを使用する。このデータセットは識別に理想的なデータセットであることが期待される。

4.3と同様に、それぞれの手法において次元を10,30,50,100に圧縮し、SVCとRandom Forestで5組の2クラス分類実験を実施した。ホテリング理論によるデータ選別の実験結果を表4.13~4.16に、LOF法によるデータ選別の実験結果を4.17~4.20にそれぞれ示す。

表 4.13: ホテリング理論によるデータ選抜実験1(次元10)

特微量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.830	0.755	0.777	0.778	0.726
	RF	0.698	0.611	0.733	0.565	0.699
Auto-Encoder	SVC	0.740	0.592	0.697	0.612	0.573
	RF	0.749	0.552	0.708	0.675	0.570
PCA	SVC	0.826	0.696	0.863	0.782	0.566
	RF	0.794	0.503	0.759	0.740	0.442

表 4.14: ホテリング理論によるデータ選抜実験1(次元30)

特微量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.822	0.731	0.830	0.734	0.620
	RF	0.558	0.621	0.721	0.471	0.652
Auto-Encoder	SVC	0.766	0.523	0.654	0.688	0.466
	RF	0.758	0.584	0.709	0.674	0.378
PCA	SVC	0.851	0.691	0.914	0.734	0.589
	RF	0.754	0.534	0.726	0.705	0.462

表 4.15: ホテリング理論によるデータ選抜実験1(次元50)

特微量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.807	0.650	0.784	0.712	0.585
	RF	0.665	0.519	0.688	0.612	0.656
Auto-Encoder	SVC	0.856	0.568	0.757	0.727	0.678
	RF	0.797	0.531	0.724	0.698	0.489
PCA	SVC	0.818	0.788	0.905	0.782	0.562
	RF	0.748	0.568	0.710	0.681	0.378

ホテリング理論によるデータ選別では、どの結果においても4.2の結果を上回る精度を得られたが、LOFによるデータ選別では精度向上が見られなかった。

表 4.16: ホテリング理論によるデータ選抜実験 1(次元 100)

特徴量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.790	0.485	0.519	0.662	0.679
	RF	0.644	0.519	0.586	0.508	0.539
Auto-Encoder	SVC	0.796	0.674	0.716	0.662	0.540
	RF	0.709	0.673	0.733	0.597	0.476
PCA	SVC	0.842	0.720	0.864	0.758	0.516
	RF	0.742	0.531	0.710	0.645	0.478

表 4.17: LOF によるデータ選抜実験 1(次元 10)

特徴量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.679	0.600	0.676	0.655	0.575
	RF	0.679	0.532	0.757	0.602	0.474
Auto-Encoder	SVC	0.616	0.519	0.651	0.522	0.447
	RF	0.638	0.520	0.659	0.554	0.488
PCA	SVC	0.717	0.497	0.707	0.662	0.468
	RF	0.695	0.506	0.626	0.604	0.419

表 4.18: LOF によるデータ選抜実験 1(次元 30)

特徴量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.653	0.732	0.585	0.585	0.513
	RF	0.642	0.481	0.631	0.556	0.523
Auto-Encoder	SVC	0.639	0.532	0.551	0.626	0.601
	RF	0.700	0.569	0.630	0.623	0.417
PCA	SVC	0.689	0.658	0.704	0.672	0.587
	RF	0.666	0.491	0.720	0.602	0.473

表 4.19: LOF によるデータ選抜実験 1(次元 50)

特徴量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.640	0.600	0.643	0.571	0.514
	RF	0.669	0.482	0.580	0.535	0.481
Auto-Encoder	SVC	0.667	0.481	0.630	0.569	0.445
	RF	0.627	0.488	0.690	0.650	0.407
PCA	SVC	0.743	0.573	0.777	0.694	0.562
	RF	0.661	0.552	0.677	0.562	0.559

表 4.20: LOF によるデータ選抜実験 1(次元 100)

特徴量抽出手法	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
X-Means	SVC	0.712	0.482	0.478	0.674	0.424
	RF	0.712	0.485	0.529	0.533	0.486
Auto-Encoder	SVC	0.642	0.491	0.680	0.581	0.472
	RF	0.592	0.523	0.674	0.518	0.477
PCA	SVC	0.698	0.508	0.735	0.716	0.378
	RF	0.645	0.520	0.669	0.555	0.433

4.4.2 実験2: 実験結果

実験2では、異常検知によって正常と判定されたデータ同士、異常と判定されたデータ同士で個別に識別モデルを学習する。実験1では、異常検知において異なる判定を下されたデータの識別を実施しているが、この実験では同じ判定が下されたデータ同士を識別するタスクを実施する。実験2では、

- 異常検知で同じ判定を下されるデータ同士でも、MMSE ラベルの違いによって睡眠形態の差異は存在し、それを適切に捉える識別モデルは訓練可能である

という仮定を検証する。テストにおいては、テストデータは初めに異常検知モデルによって判定が下され、その後、異常判定されたデータは異常判定されたデータで学習したモデルへ接続され、正常判定されたデータは正常判定されたデータで学習したモデルへ接続される。

この実験においては、特徴量抽出手法として3.3節のPCAを採用した。特徴量の次元を10,30,50,100に圧縮し、識別器としてのSVCおよびRandom Forestに接続した。異常検知として、ホテリング理論のみを使用した。

実験結果を図4.21,4.22示す。

表 4.21: 結果: データ選抜実験2. 異常データ識別

圧縮次元	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
10	SVC	0.380	0.607	0.502	0.422	0.583
	RF	0.527	0.495	0.518	0.407	0.492
30	SVC	0.341	0.607	0.612	0.247	0.703
	RF	0.500	0.517	0.598	0.324	0.639
50	SVC	0.767	0.719	0.847	0.672	0.696
	RF	0.655	0.545	0.738	0.474	0.549
100	SVC	0.441	0.626	0.677	0.298	0.513
	RF	0.575	0.412	0.490	0.357	0.456

表 4.22: 結果: データ選抜実験2. 正常データ識別

圧縮次元	モデル	健康/その他	認知症/その他	健康/認知症	健康/MCI	MCI/認知症
10	SVC	0.581	0.690	0.775	0.477	0.698
	RF	0.519	0.610	0.604	0.480	0.707
30	SVC	0.584	0.676	0.705	0.514	0.641
	RF	0.518	0.582	0.635	0.469	0.691
50	SVC	0.529	0.685	0.733	0.489	0.673
	RF	0.516	0.587	0.584	0.469	0.708
100	SVC	0.558	0.685	0.759	0.464	0.662
	RF	0.479	0.523	0.587	0.453	0.671

第5章 考察

5.1 LSTM,SVM,Random Forest を用いた機械学習モデル

2クラス分類の実験結果について、表4.7, 4.8に示されているとおり、いずれのモデルでも Macro F1 score は0.6を超え、分類タスクは概ねうまく進んだことが分かる。我々はグループ分けにおいて、MMSEスコアの27点を閾値としたので、この結果は、認知症がもたらす睡眠障害は、MMSEにおいて明確に認知状態が低下していると分類される状態だけでなく、軽度の認知障害がある、と推定される段階においても既に現れていることを示唆している。

MMSEによる認知症推定尺度の高/低群の二値分類実験において、表4.7, 4.8から分かる通り、いずれのモデルにおいても Macro F1 スコアにおいて0.6を超える精度を達成した。この結果は異なるデータセットを用いた [9][10][11]らの結果と10~20%程度劣るものの、依然として睡眠データの認知症尺度推定への有用性を一定程度示すものである。一方で、[7]で使用したデータと本研究で使用したデータは一部共通しているが、[7]の結果と比較して本実験の結果は20%程度劣るものであった。この理由として、[7]では睡眠データだけでなく、高齢者養護施設内に設置されたドアセンサー、人感センサーから集計したデータから生活全体を表現する行動ベクトルを形成し、認知症による生活への影響をより高次元で表現することに成功しているからだと考えられる。この実験において実施した3つの機械学習推定モデルにおいて、LSTMを使用したものが最高精度を達成した。この結果は、認知症による睡眠への影響はある特定のイベントとして現れ、その特徴を時系列データ解析モデルであるLSTMが上手くとらえることが出来たからだと推測できる。認知症による睡眠への影響は睡眠全体に全般的におよぶのではなく、むしろ特定の呼吸障害に伴う身体活動の活発化など、特定のイベントとして存在するという事は[11]において指摘されており、本実験の結果と符号する。

5.2 K-Means, PCA, Auto Encoder を用いた特徴量抽出

3.3節で説明した複数の特徴量抽出手法を比較する実験について、表4.9~4.12に示されている通り、認知症と疑われるグループと認知的に健康と推定されるグループまたはMCIを識別するタスク、認知症が疑われるグループと認知的に健康と推定されるグループを識別するタスク、MCIと認知症が疑われるグループを識別するタスクにおいて、概ねmacro-F1スコアは0.7を超え、識別が上手く進んだことが分かる。最高精度はオートエンコーダーで次元を50まで圧縮し、識別機にSVCを適用したときの認知的に健康と推定されるグループと認知症が疑われるグループの識別タスクにおける0.806であった。

先述の実験と比較して結果が向上した一因として、データ数の増加が考えられるが、3.2節の複数の機械学習モデルを比較した実験においては、一晩毎の睡眠データをモデルで処理し、その後モデルの出力ベクトルを平均することで最終的な結果を出したが、同じ参加者であっても、睡眠の様子は必ずしも一様ではなく、従って一晩毎のデータに対する機械学習モデルの結果は大きく異なることがあり、結果として識別に有効な参加者を代表する特徴量ベクトルを作成できなかったことが考えられる。一方で、3.3節で説明した複数の特徴量抽出手法を比較した実験においては、手法自体の有効性に加え、全ての参加者の一晩のデータをまず初めにX-means,Auto Encoder またはPCAで処理し、その後参加者毎に変換された一晩の睡眠データベクトルの平均および分散をとることで、個人を代表する特徴量を作成した。この手順により、個人に対して存在する睡眠行動のブレや偏差を吸収し、より本質的な睡眠行動を表現するベクトルを獲得できたことで、認知症尺度を推定精度が向上したことが考えられる。

5.3 異常検知によるデータ選別

ホテリング理論を用いて算出した、それぞれのラベルの異常値を図5.1~5.4に示す。ラベル無しグループにおいて、圧倒的多数が正常データであり、ラベル無しデータの多数が正常データであったことが分かる。正常データにおいては、約6%の参加者が異常判定され、これらの参加者のデータがMCI,認知症が疑われるグループとの識別を困難にしていると考えられる。MCIデータにおいては約70%のデータが、認知症が疑われるグループにおいては約61%の参加者が正常判定され、これらの参加者は、認知能力が衰えている傾向があるものの、睡眠様態としては健康であることが推定され、睡眠データから認知症を推定するタスクを困難にしているデータだと考えられる。これらの結果から、認知症が睡眠様式として必ずしも顕在化してはいないことが推察される。

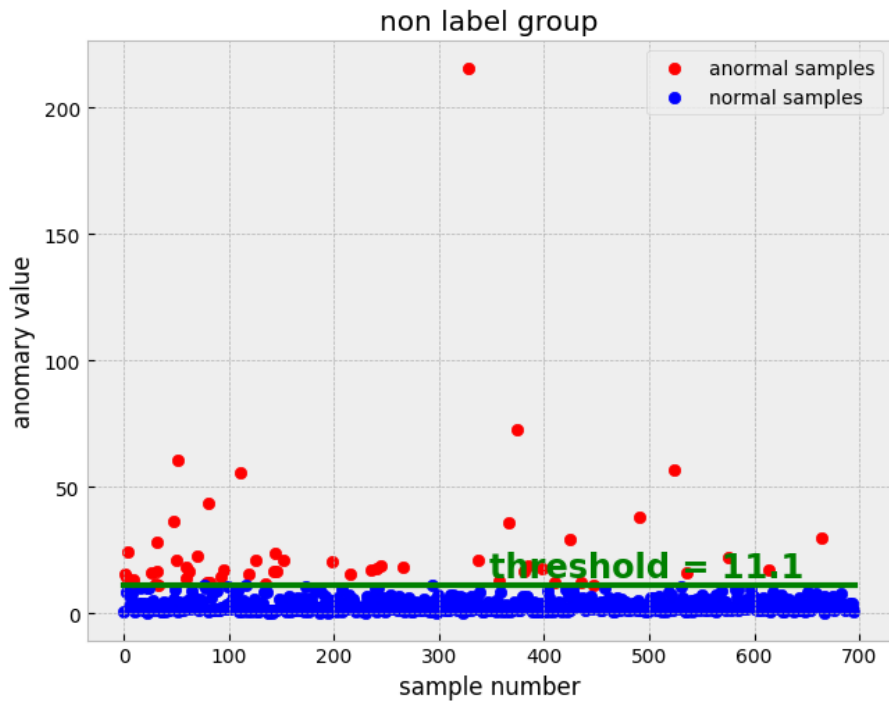


図 5.1: ホテリング理論による異常値: ラベル無しグループ

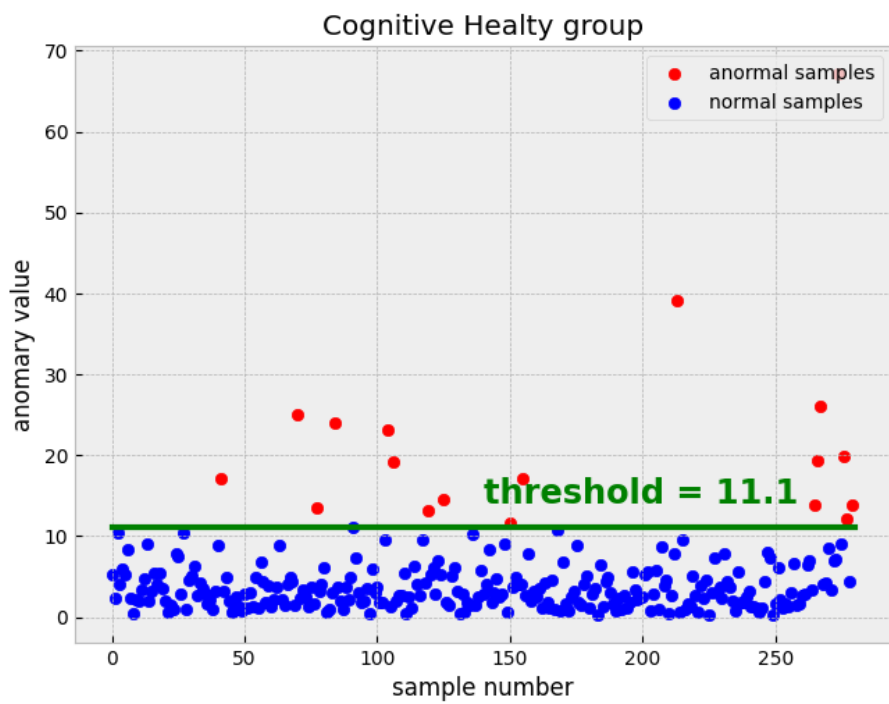


図 5.2: ホテリング理論による異常値: 認知的に健康と推定されるグループ

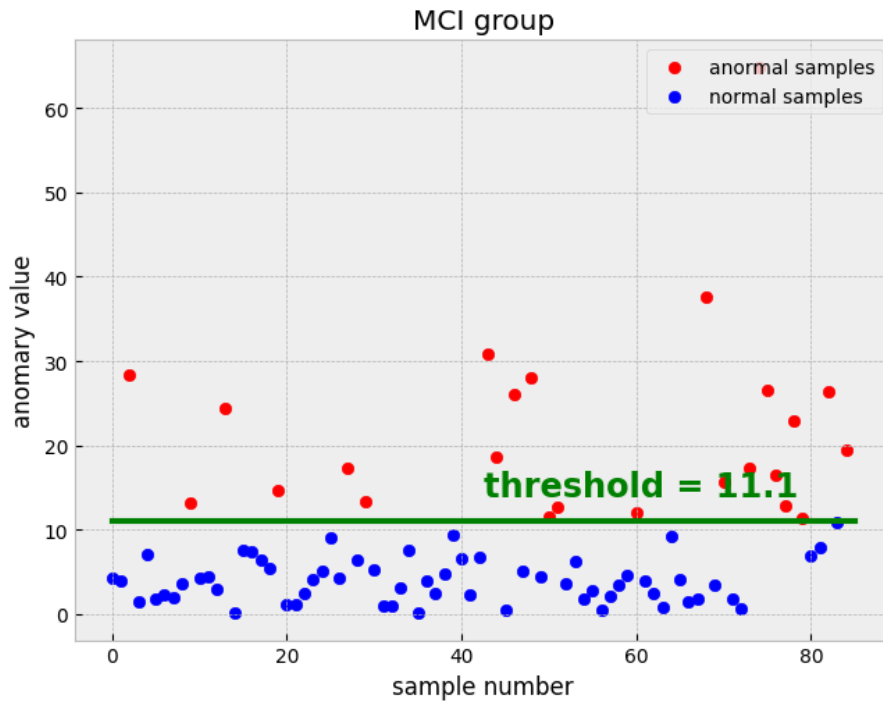


図 5.3: ホテリング理論による異常値:MCIグループ

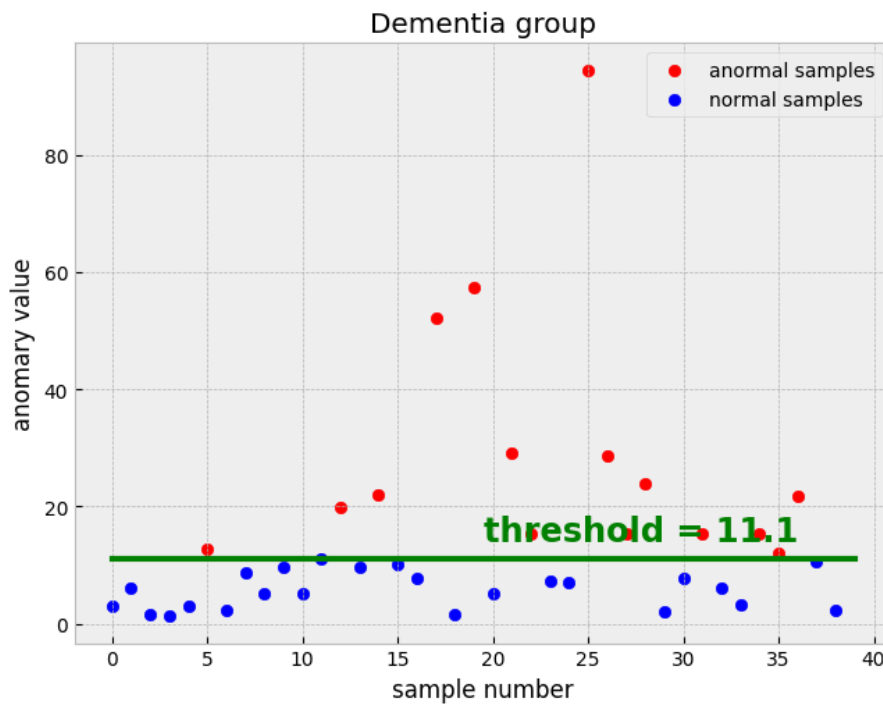


図 5.4: ホテリング理論による異常値: 認知症が疑われるグループ

次に、LOF 理論を用いて算出した異常値を 5.1~5.4 に示す。なお、これらの図においては、異常値は 0-1 に範囲にスケールしている。

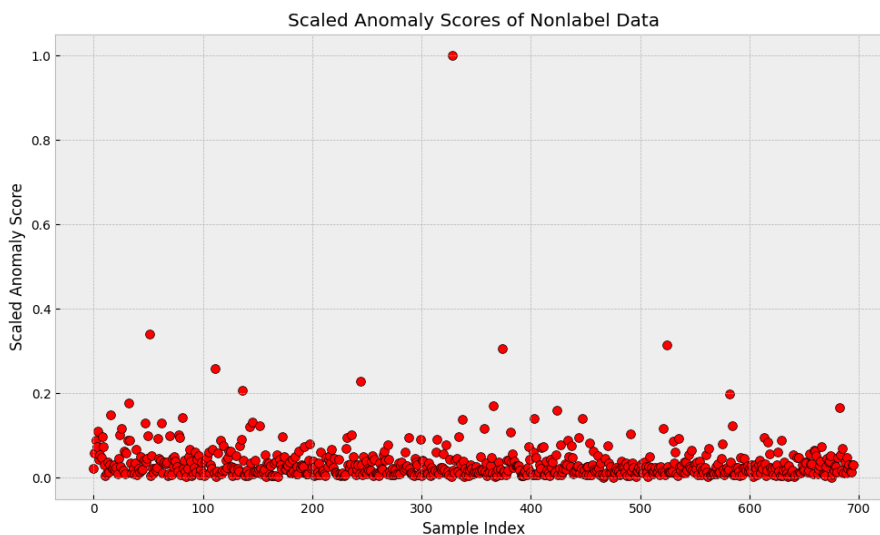


図 5.5: LOF による異常値: ラベル無しグループ

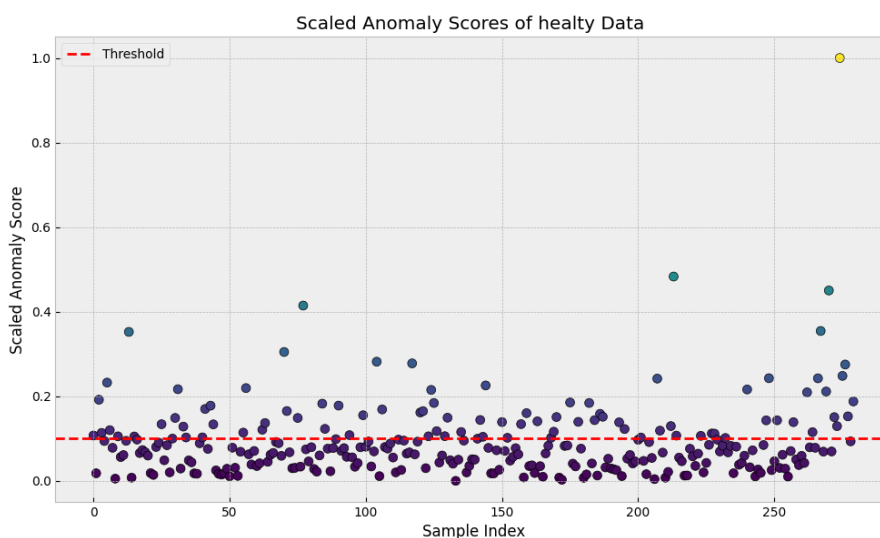


図 5.6: LOF による異常値: 認知的に健康と推定されるグループ

LOF による異常値算出の結果においても、ホテリング理論で計算したものと同様の傾向が見て取れる。ラベル無しデータの圧倒的多数は正常サンプルである一方で、認知的に健康と推定されるグループの中にも異常判定されるサンプルも少なからず存在し、MCI, 認知症が疑われるグループの中にも、正常判定されるサンプルが多数存在した。LOF によるデータ選別実験の精度がデータ選別なしの結果

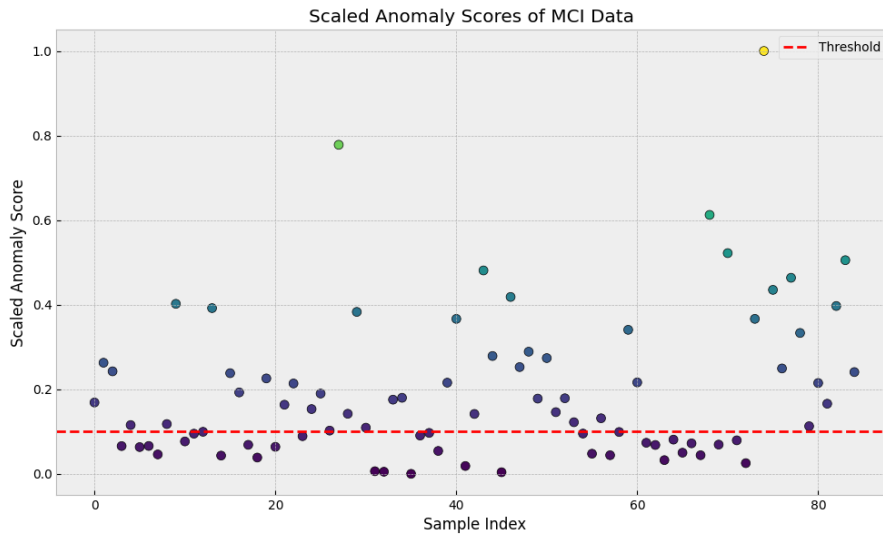


図 5.7: LOF による異常値:MCI グループ

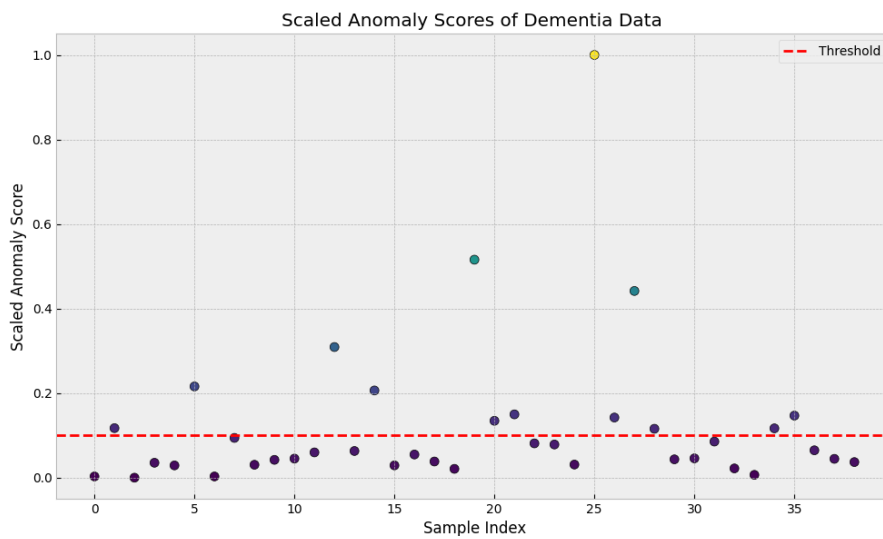


図 5.8: LOF による異常値: 認知症が疑われるグループ

を上回らなかった要因として、LOF法における適切なパラメータの選定が為されなかったことが考えられる。

5.3.1 理想的なデータセットを使用した実験

表4.13~4.16に示している通り、ホテリング理論によるデータ選別を行った実験の多くの場合において3.3節の実験の精度を上回る結果が出た。最高精度はPCAにおいて次元を50に圧縮し、識別器としてSVCを用いた認知的に健康と推定されるグループと認知症が疑われるグループの識別タスクにおける0.914であった。この実験において、健康と推定されるグループは正常判定された参加者のみのデータ、MCI、認知症が疑われるグループにおいては異常判定された参加者のみのデータを用いた。認知症が睡眠に様々な悪影響を与えることは知られており、それは本研究の生理学的根拠であるが、臨床において睡眠様態と推定認知状態は必ずしも符合せず、認知症が睡眠に与える影響は必ずしも顕在化しない。本実験においては、ある種の典型的な健康的睡眠、典型的な不健康な睡眠をまず異常検知システムにより測り、それらを識別するタスクに取り組んだと考えることが出来る。その結果、最高推定精度は約10%向上し、機械学習モデルで認知状態に基づく典型的な睡眠傾向を識別することが可能であることを示した。

5.3.2 異常値ごとに異なるモデルを用いた実験

表4.21, 4.22に示している通り、識別タスクの結果は概ね実験1より悪化した。本実験では、異常検知システムによって異常判定/正常判定されたデータ同士の識別タスクに取り組んだが、それは実験1で実施した典型的な健康/不健康な睡眠を識別するよりも困難であったことが推察される。しかしながら、50次元に圧縮した認知的に健康と推定されるグループと認知症が疑われるグループの識別タスクにおいて、識別モデルにSVCを用いたものでは、異常データ同士の識別精度が0.847、正常データ同士の識別精度が0.733であり、3.3節の実験の結果と遜色ないものであった。本実験においては、実験1よりもより実際の状況に近いデータを用いて解析したと考えられるが、本実験の結果は、睡眠特徴量のより詳細なモードを解析可能なシステムを用いることができれば、必ずしも顕在化はしない認知症が睡眠に与える本質的な影響を抽出し、一見同一な睡眠特徴量からでも、認知状態を識別するシステムの作成可能性を示唆するものである。

5.4 ラベルノイズの存在

表5.1~5.3に、データの選別の有無による、識別タスクの最高精度の比較を載せる。認知的に健康と推定されるグループと認知症が疑われるグループの識別タ

スクにおいて、認知的に健康と推定されるグループは正常判定されたサンプルのみ、認知的に健康と推定されるグループは異常判定されたサンプルのみを選別するデータ選別を施した実験において、最高精度 macro-F1 スコア 0.914 を達成し、これはデータ選別を実施しない実験の精度よりも約 10% 上であった。健康/その他の識別タスク、健康/MCI の識別タスクにおいても、データ選別を実施した実験の精度が、データ選別を実施しない実験の精度よりもおおよそ 20% 程度向上しており、これらの結果は、MMSE によるラベルと、典型的な睡眠形態が符号していない参加者の存在を示唆するものである。

表 5.1: 健康/認知症 識別精度比較

データ選別	モデル	圧縮次元	健康/認知症
なし	SVC	50	0.806
あり	SVC	30	0.914
異常サンプルのみ	SVC	50	0.847
正常サンプルのみ	SVC	10	0.775

表 5.2: 健康/その他 識別精度比較

データ選別	モデル	圧縮次元	健康/その他
なし	RF	30	0.617
あり	SVC	50	0.856
異常サンプルのみ	SVC	50	0.767
正常サンプルのみ	RF	30	0.584

表 5.3: 健康/MCI 識別精度比較

データ選別	モデル	圧縮次元	健康/MCI
なし	RF	10	0.555
あり	SVC	10	0.782
異常サンプルのみ	SVC	50	0.672
正常サンプルのみ	RF	30	0.514

認知症が睡眠に影響を与えることは臨床的に確かではあるが、個人の睡眠形態は、認知症を引き起こす疾患だけではなく、様々な複合的な要因により形成されるものである。認知的には健康であるが、睡眠形態としては認知症傾向のあるグループに近い参加者や、認知症が疑われるが、睡眠形態として認知症が顕在化しない参加者が存在し、これらはラベルとデータが符号しない、つまり識別タスクにおいてラベルにノイズが混じっている状態であると考えることが出来る。本研究のデータ収集において、参加者は4つの高齢者福祉施設に居住する日本の高齢

者であり、参加に際し条件は設けず、多様な参加者が集まった。睡眠データから認知症尺度を推定するタスクにおいて、ラベルにノイズがある参加者、つまり睡眠形態と認知症ラベルが合致しない参加者は、その睡眠形態から認知状態を推定することは非常に困難であると考えられ、データ選別を実施した実験では、そのような識別困難な参加者を除外したことが精度の向上に寄与したと考えられる。

一方で、異常検知において異常サンプルのみを選別し実施した識別タスクでは、データ選別を実施しない実験と比較していずれの識別タスクに於いても推定精度が向上しており、このことは、典型的な睡眠様態から外れた参加者内においても、認知状態によって睡眠の形態には差異があり、従って睡眠データから認知状態は識別可能であることを示している。また、異常検知において正常サンプルのみを選別した識別タスクにおいても、データ選別を実施しない実験と数%劣る程度の精度は示しており、正常クラス内においても識別がある程度は可能であると考えることが出来る。これらの結果は、認知症の睡眠形態への影響の現れ方が、睡眠クラスタにより異なるものを示唆するものである。

第6章 おわりに

我々は、本研究においてまず、認知症尺度スコアにおいて、高得点群と低得点群にはその睡眠中の計測された活動量の一部について、優位な差があることを統計的に確認した。次に、睡眠中の活動量のみを用いた複数の認知症尺度スコアの二分類予測システムを比較検討し、最高で Macro F1 スコアで 0.67 を達成した。次に、複数の特徴量抽出手法を比較し、最高で Macro F1 スコアで 0.81 を達成した。最後に異常検知モデルを用いた識別システムを構築し、典型的な健康/不健康な睡眠状態を計算し、理想的なデータセットにおいて最高精度で 0.91 を達成した。

この研究は、認知症と睡眠の関連に関する臨床的な研究の一部であると同時に、様々なモダルを用いて研究されている認知症の自動検出システムの一部に、睡眠情報を用いることができることを示唆するものである。本研究の結果は、様々なモダルを用いて研究されている認知症の自動検出システムの一部に、睡眠情報を用いることができる可能性を示唆するものであり、精度改善の必要性はあるものの、MMSE のより精緻な分類が可能となれば、診断の補助や認知状態のモニタリングシステムに応用できる可能性がある。本研究でデータ収集のために使用したセンサーは、安価で購入可能な市販品であり、かつ利用者への負荷が非常に軽い。これらの要素は睡眠データを用いた認知症自動推定システムの普及のために非常に重要な要素である。認知症は自分自身でその兆候に気付くのが困難な病であり、また、認知機能低下の兆しに気付き得る日常的に触れ合う第三者の存在は、必ずしもすべての人にとって可能ではない。今後ますます認知症が社会に与える影響が増大することが見込まれる現在において、認知症の殆ど唯一の有効な治療法である早期治療に繋げるために、認知症の自動推定システムの普及は重要であり、本研究がそれに寄与することを期待する。

睡眠の様態は必ずしも認知状態を反映せず、様々な個人的な環境を背景としている現代人から、認知症が睡眠に与える本質的な影響のパターンを抽出することは必ずしも容易ではない。睡眠様態を単に異常/正常で分けするだけでは、睡眠データから認知状態を推定することは困難であり、より詳細な睡眠モードの解析、認知症が睡眠に与える本質的な影響のパターンを抽出するシステムを構築することが求められる。これらは今後の課題であり、とりわけ、認知症の早期発見のために、MCI の早期発見を可能とするためのネットワーク設計や、データの処理方法について検討を続けていく所存である。

謝辞

本研究の遂行にあたり，多くの皆様のご助力を賜りました．貴重なデータを収集し共有してくださった西日本電信電話株式会社及び，ジョージ・アンド・ショー
ン株式会社の関係者の皆様に感謝申し上げます．岡田研究室において，同様の取
り組みを行っていらっしゃった南澤さん，安藤さん，村瀬さんには，論文やゼミで
の議論を通して数多くの助言を頂きました．また，日ごろのゼミを通じての議論
や，研究室内外の活動を通じて関わってくださった岡田研究室の皆様にご感謝申
し上げます．岡田将吾准教授には，常に的確な指摘や提案を頂き，本研究の実施に
際し多大なご助力を頂きました．感謝申し上げます．

付録A データ

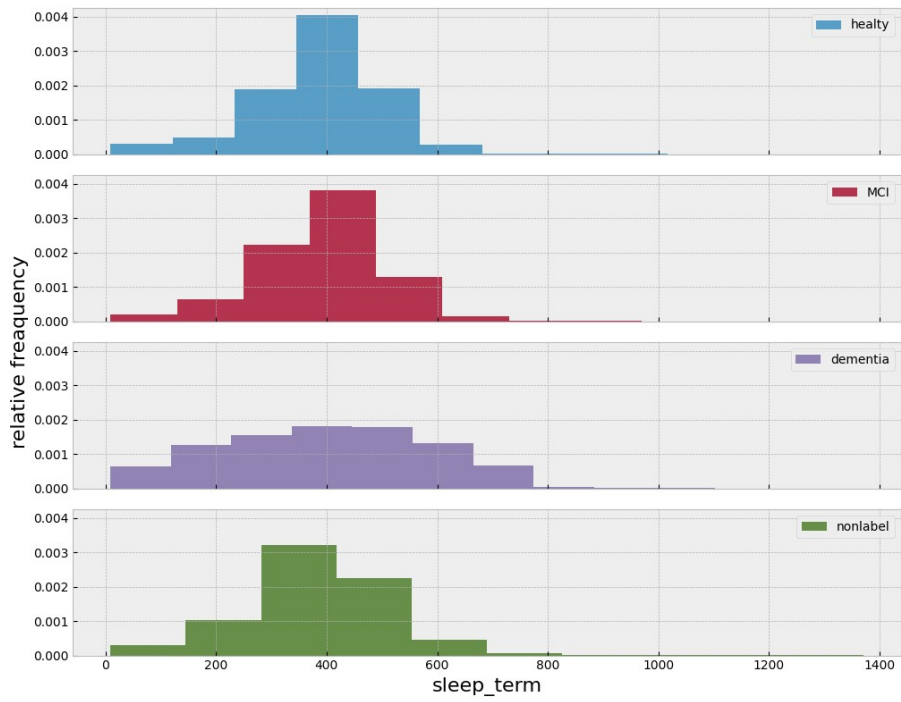


図 A.1: 睡眠時間 (分)

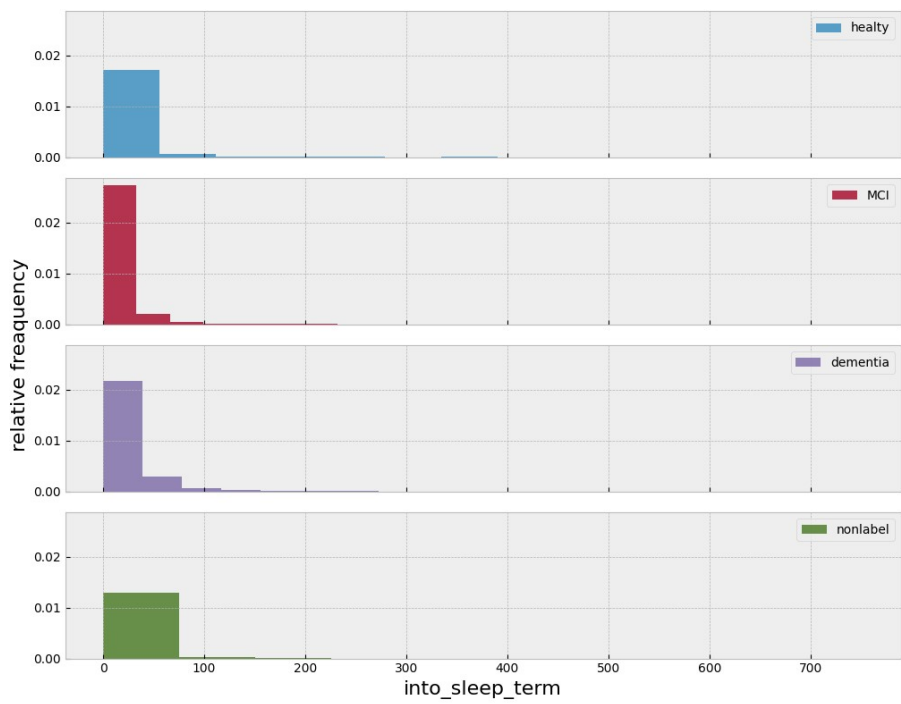


図 A.2: 入眠時間 (分)

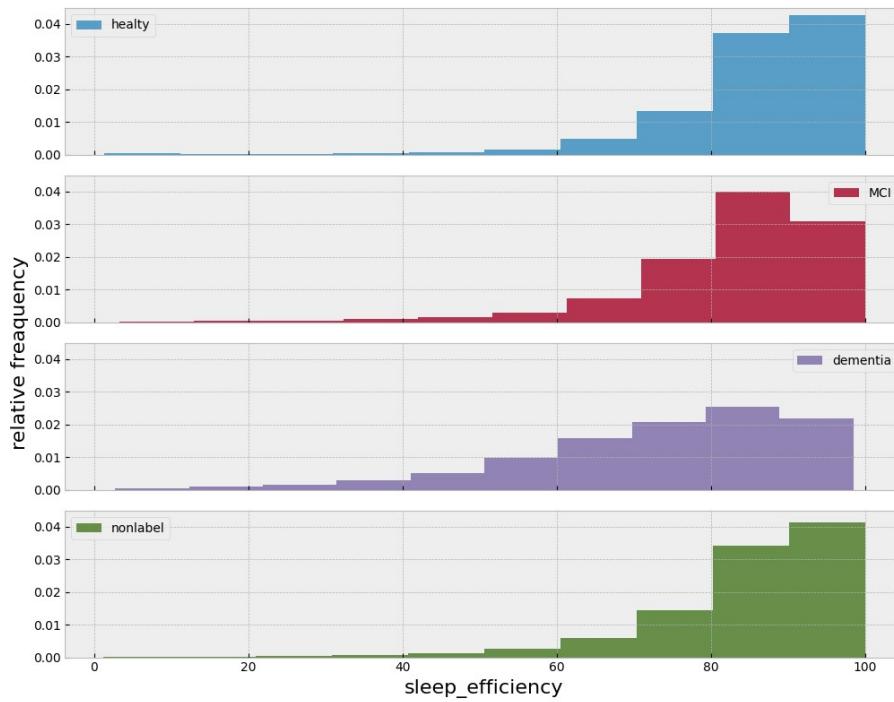


图 A.3: 睡眠效率

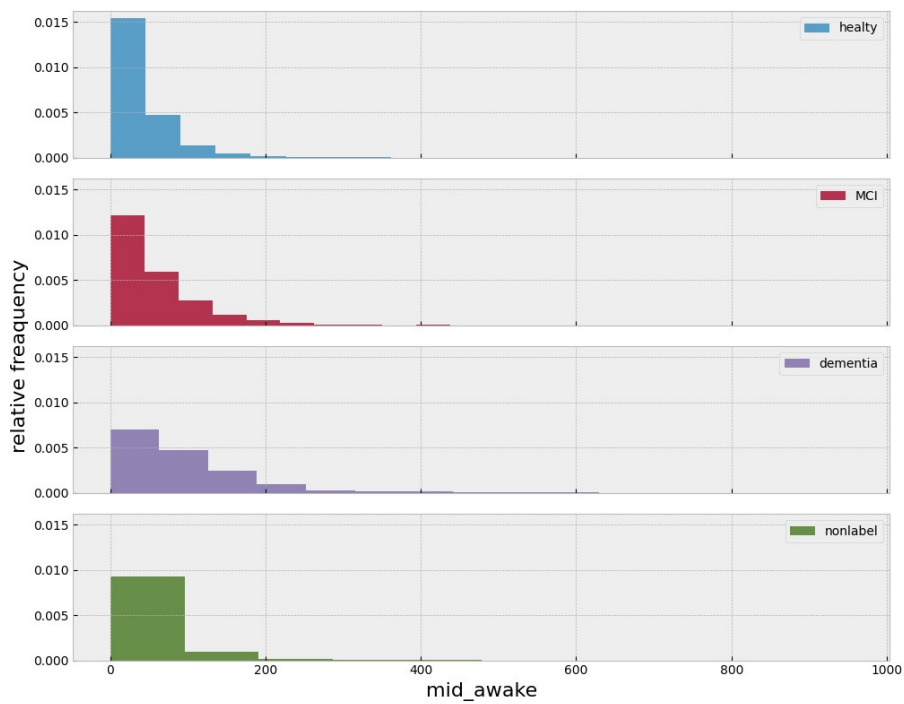


图 A.4: 中途觉醒时间 (分)

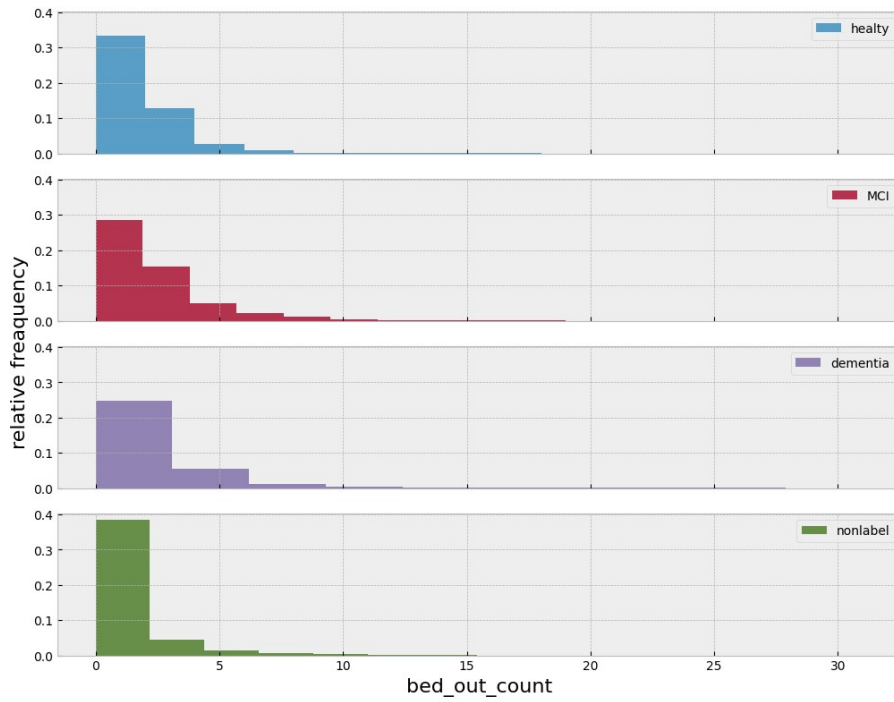


図 A.5: 離床回数 (回)

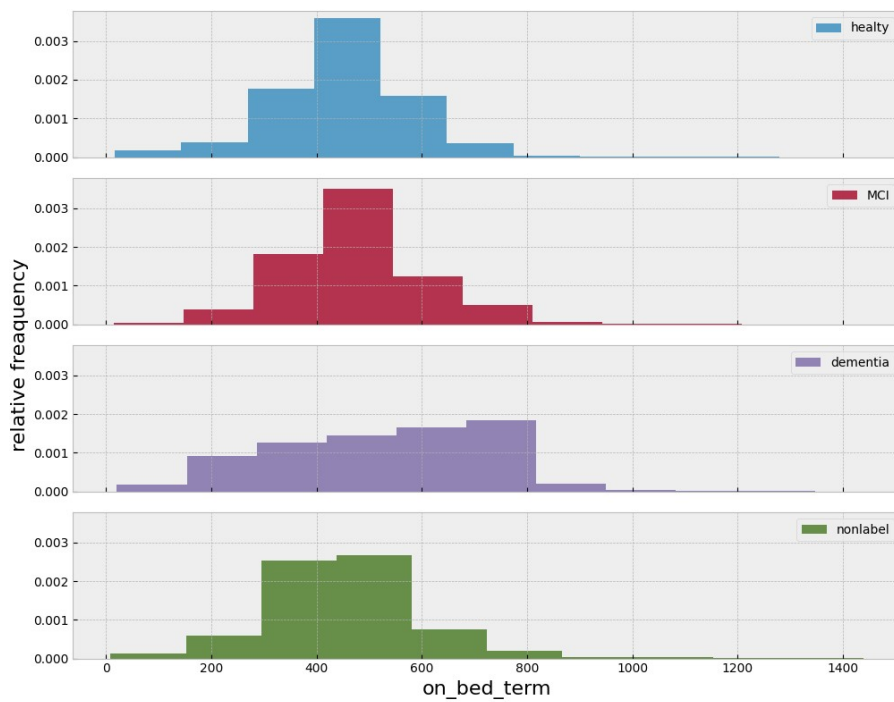


図 A.6: 在床時間 (分)

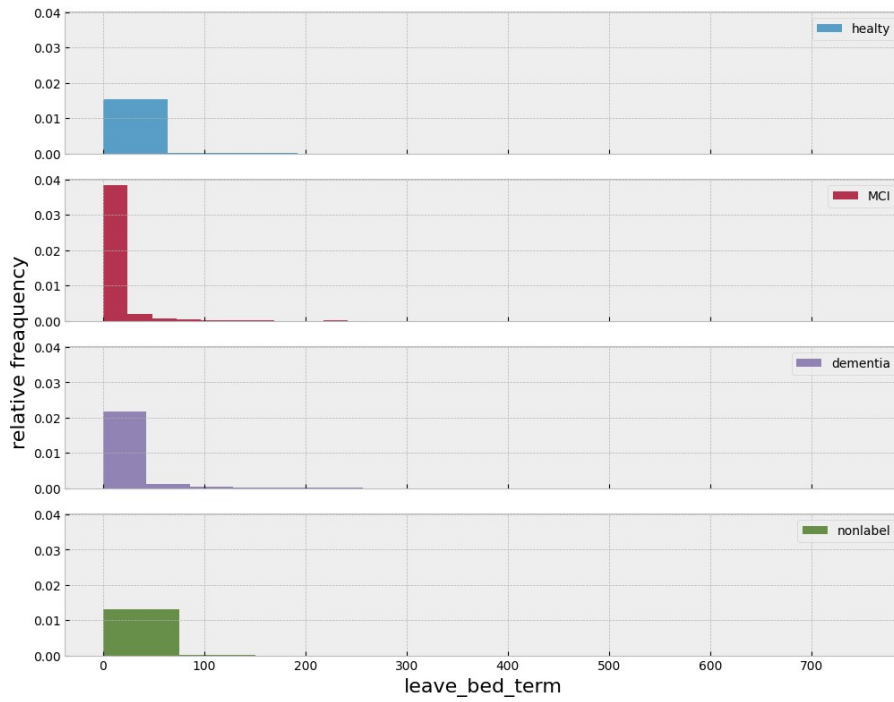


図 A.7: 離床時間 (分)

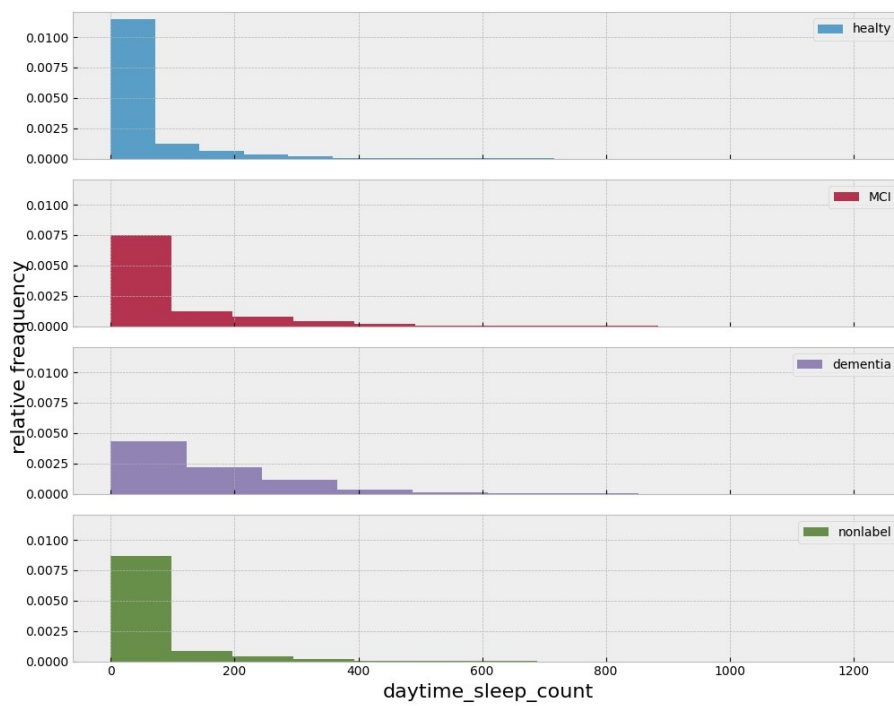


図 A.8: 日中睡眠時間 (分)

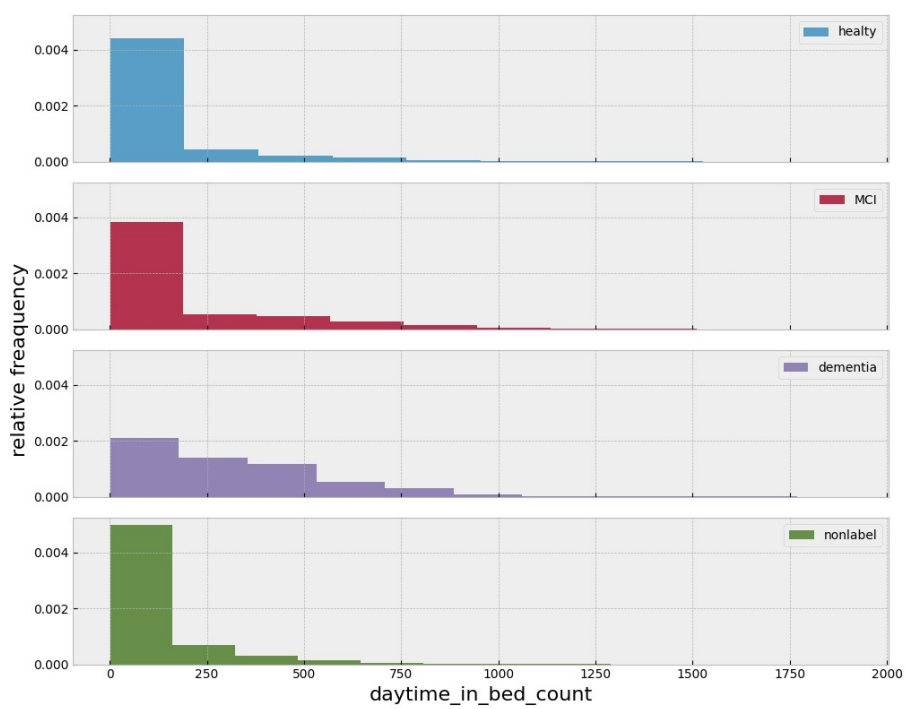


図 A.9: 日中在床時間 (分)

参考文献

- [1] 日本神経学会. 認知症疾患診療ガイドライン 2017, 2017.
- [2] Fact sheets: Dementia, World Health Organization, September 2020. <http://www.who.int/en/news-room/fact-sheets/detail/dementia>.
- [3] Konstantina G Yiannopoulou and Sokratis G Papageorgiou. Current and future treatments for alzheimer ’ s disease. *Therapeutic advances in neurological disorders*, Vol. 6, No. 1, pp. 19–33, 2013.
- [4] Chathurika Palliya Guruge, et al. Advances in multimodal behavioral analytics for early dementia diagnosis: A review. In *Proceedings of the 2021 International Conference on Multimodal Interaction*, pp. 328–340, 2021.
- [5] William Jarrold, et al. Aided diagnosis of dementia type through computer-based analysis of spontaneous speech. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pp. 27–37, Baltimore, Maryland, USA, June 2014. Association for Computational Linguistics.
- [6] Mark R. Hodges, et al. Pervasive Computing, 8th International Conference, Pervasive 2010, Helsinki, Finland, May 17-20, 2010. Proceedings. *Lecture Notes in Computer Science*, pp. 192–209, 2010.
- [7] Akira Minamisawa and Others. Dementia Scale Score Classification Based on Daily Activities Using Multiple Sensors. *IEEE Access*, Vol. 10, pp. 38931–38943, 2022.
- [8] Shogo Okada, et al. Dementia scale classification based on ubiquitous daily activity and interaction sensing. *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 192–198, 2019.
- [9] Naoya Matsuda, Iko Nakari, and Keiki Takadama. Alzheimer dementia detection based on unstable circadian rhythm waves extracted from heartrate. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pp. 4473–4476, 2021.

- [10] Naoya Matsuda, Iko Nakari, and Keiki Takadama. Unstable circadian rhythm of heart rate of alzheimer dementia based on biological data of mattress sensor. Vol. 2022, pp. 1129–1132, 07 2022.
- [11] Somayeh Khosroazad, Ali Abedi, and Marie Hayes. Sleep signal analysis for early detection of alzheimer’s disease and related dementia (adrd). *IEEE Journal of Biomedical and Health Informatics*, Vol. PP, pp. 1–12, 01 2023.
- [12] Alexandra Wennberg. Sleep Disturbance, Cognitive Decline, and Dementia: A Review. *Seminars in Neurology*, Vol. 37, No. 04, pp. 395–406, 2017.
- [13] Shireen Sindi, et al. Sleep disturbances and dementia risk: A multicenter study. *Alzheimer’s & Dementia*, Vol. 14, No. 10, pp. 1235–1242, 2018.
- [14] James M. FitzGerald, et al. Sleep-wake cycle disturbances in elderly acute general medical inpatients: Longitudinal relationship to delirium and dementia. *Alzheimer’s & Dementia: Diagnosis, Assessment & Disease Monitoring*, Vol. 7, No. 1, pp. 61–68, 2017.
- [15] Youngsin Jung, et al. Phenoconversion from probable rapid eye movement sleep behavior disorder to mild cognitive impairment to dementia in a population-based sample. *Alzheimer’s & Dementia: Diagnosis, Assessment & Disease Monitoring*, Vol. 8, No. 1, pp. 127–130, 2017.
- [16] Akira Minamisawa, et al. Dementia Scale Score Classification Based on Daily Activities Using Multiple Sensors. *IEEE Access*, Vol. 10, pp. 38931–38943, 2022.
- [17] Ingrid Arevalo-Rodriguez, et al. Mini Mental State Examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews*, Vol. 2021, No. 7, p. CD010783, 2021.
- [18] Tamara L Hayes, et al. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. *Alzheimer’s & Dementia*, Vol. 4, No. 6, pp. 395–405, 2008.
- [19] Prafulla Nath Dawadi, et al. Automated cognitive health assessment from smart home-based behavior data. *IEEE journal of biomedical and health informatics*, Vol. 20, No. 4, pp. 1188–1194, 2015.
- [20] Saskia Robben, et al. Delta features from ambient sensor data are good predictors of change in functional health. *IEEE journal of biomedical and health informatics*, Vol. 21, No. 4, pp. 986–993, 2016.

- [21] Brianne A. Kent, et al. Sleep and hippocampal neurogenesis: Implications for Alzheimer 's disease. *Frontiers in Neuroendocrinology*, Vol. 45, pp. 35–52, 2017.
- [22] Michael R Irwin, et al. Implications of sleep disturbance and inflammation for Alzheimer's disease dementia. *The Lancet Neurology*, Vol. 18, No. 3, pp. 296–306, 2019.
- [23] Joseph R. Winer, et al. Waking Up to the Importance of Sleep in the Pathogenesis of Alzheimer Disease. *JAMA Neurology*, Vol. 75, No. 6, p. 654, 2018.
- [24] Jerrah K. Holth, et al. Sleep in Alzheimer's Disease—Beyond Amyloid. *Neurobiology of Sleep and Circadian Rhythms*, Vol. 2, pp. 4–14, 2017.
- [25] MairavCohen - Zion, Carl Stepnowsky, Marler, Tamar Shochat, Daniel F. Kripke, SoniaAncoli - Israel. Changes in Cognitive Function Associated with Sleep Disordered Breathing in Older People. *Journal of the American Geriatrics Society*, Vol. 49, No. 12, pp. 1622–1627, 2001.
- [26] Terri Blackwell, Kristine Yaffe, Sonia Ancoli-Israel, Jennifer L Schneider, Jane A Cauley, Teresa A Hillier, Howard A Fink, Katie L Stone, and Study of Osteoporotic Fractures Group. Poor Sleep Is Associated With Impaired Cognitive Function in Older Women: The Study of Osteoporotic Fractures. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, Vol. 61, No. 4, pp. 405–410, 2006.
- [27] Alberto Corbi, et al. Connection between sleeping patterns and cognitive deterioration in women with Alzheimer's disease. *Sleep and Breathing*, Vol. 26, No. 1, pp. 361–371, 2022.
- [28] Shingo Tomura, et al. Estimating Surrounding Symptom Level of Dementia Patient by Sleep Stage. *2015 9th International Symposium on Medical Information and Communication Technology (ISMICT)*, pp. 190–194, 2015.
- [29] PRACTICE GUIDELINE FOR THE Treatment of Patients With Alzheimer 's Disease and Other Dementias. <https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/Clinical%20Practice%20Guidelines/alzheimers.pdf>.
- [30] RISK REDUCTION OF COGNITIVE DECLINE AND DEMEN- TIA . <https://iris.who.int/bitstream/handle/10665/312180/9789241550543-eng.pdf?sequence=17>.

- [31] 鳴田千晶, 松下大輔. ミリ波レーダーによる臥床時非接触睡眠判定手法. 日本建築学会技術報告集, Vol. 28, No. 68, pp. 527–532, 2022.
- [32] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, Vol. 9, No. 8, pp. 1735–1780, 1997.
- [33] Corinna Cortes and Vladimir Naumovich Vapnik. Support-vector networks. *Machine Learning*, Vol. 20, pp. 273–297, 1995.
- [34] L Breiman. Random forests. *Machine Learning*, Vol. 45, pp. 5–32, 10 2001.
- [35] Dan Pelleg and Andrew Moore. X-means: Extending k-means with efficient estimation of the number of clusters. *Machine Learning*, p, 01 2002.
- [36] Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. Lof: identifying density-based local outliers. *SIGMOD Rec.*, Vol. 29, No. 2, p. 93–104, may 2000.
- [37] 井手剛. 入門 機械学習による異常検知 R による実践ガイド. コロナ社.