

デジタルバイオマーカーの時系列分析に基づく 摂食障害予測モデルの開発

中村 真穂[†] Siriaraya Panote^{††} Cheng Asaka Lan^{†††} 中島 伸介^{†††}

^{†††} 京都産業大学先端情報学研究科 〒 603-8555 京都府京都市北区上賀茂本山

^{††} 京都工芸繊維大学 情報工学・人間科学系 〒 606-8585 京都府京都市左京区松ヶ崎橋上町

^{†††} 京都産業大学情報理工学部 〒 603-8555 京都府京都市北区上賀茂本山

E-mail: [†]{g2254261,a007532,nakajima}@cc.kyoto-su.ac.jp, ^{††}spanote@kit.ac.jp

あらまし 摂食障害の従来の診断手法は、自己申告や臨床観察に依存しており、想起バイアスや過少申告といった主観的要因の影響を受けやすい。そのため、日常生活における行動パターンから摂食障害に特徴的な行動シグネチャーを抽出し、早期発見のための予測手法を確立することが重要な課題となっている。本研究では、スマートフォンのセンサデータを用いて、摂食障害の有無を識別する機械学習モデルを構築し評価することを目的とする。具体的には、摂食障害群および健常対照群から収集した加速度、位置情報、アプリ使用履歴などのデータを用いて、時系列分析が可能な LSTM モデルを構築し、24 時間や 12 時間などの時間単位で分割した特徴量に対して分類精度を比較する。将来的には、本研究で得られた知見とモデルをもとに、予測機能を搭載したスマートフォンアプリケーションを開発し、日常的な使用を通じて摂食障害の兆候を検出し、通知できるシステムの実現を目指す。

キーワード デジタルバイオマーカー、摂食障害、機械学習、スマートフォン

1 はじめに

摂食障害は中枢性摂食異常症とも呼ばれ厚生労働省の特定疾患（難病）に指定される精神疾患である。主に拒食症と過食症に分けられ、どちらも食行動異常とそれに伴う認知や情動の障害を主徴としている。日本では、年間約 2 万 3 千人の患者がおり（図 1 [1]）、摂食障害の罹患率は増加傾向にあるため、その診断と治療の重要性が高まっている [2]。特に、拒食症は早期発見が遅れることで症状が悪化しやすく、他の精神疾患よりも死亡率が高いというデータがある [3]。また、摂食障害の罹患数は、10 代、20 代に多いというデータもあるため、大学生の年齢の間までに早期発見、治療が望まれている [4]。従来の摂食障害の診断では、臨床医による観察や患者の自己申告が主に用いられており、これらの方法は主観に依存するため、記憶の偏りや過少申告といった問題が指摘されている（図 2）。すなわち、従来の方法では摂食障害の早期の発見および治療の実現は容易ではない。このような状況に対し、治療法の一つとして、食事や感情の記録を通じて自己理解や行動変容を支援するモバイルヘルスアプリの開発 [5] といった取り組みがある。ただし、興味深い取り組みではあるものの、行動に着目した摂食障害の早期の発見および治療に対する効果的なシステムの開発には至っていない。

そこで我々は、スマートフォンを利用したデジタルバイオマーカーに着目した。デジタルバイオマーカーとは、埋め込み型環境センサやウェアラブルなどのデジタルデバイスによって収集および測定される、客観的で定量化可能な生理学および行動的データである [6]。これらは行動データをリアルタイムで

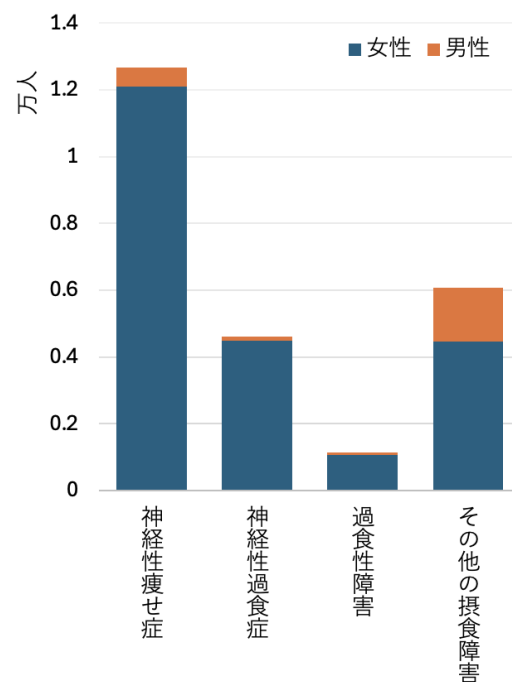


図 1 国内の摂食障害患者数（年間）[1]

収集し、従来の方法では得られない特徴を捉えることができる。たとえば、過食症の患者は、不規則な食事時間、頻繁なファストフード店の訪問、または身体活動レベルの低下といった特異な行動を示す可能性がある。したがって本稿では、スマートフォンのセンサデータを活用し、摂食障害に関連する行動パターンを検出し、機械学習を用いた予測モデルの構築を目的としている。そこで、本稿ではスマートフォンのセンサデータを使用し

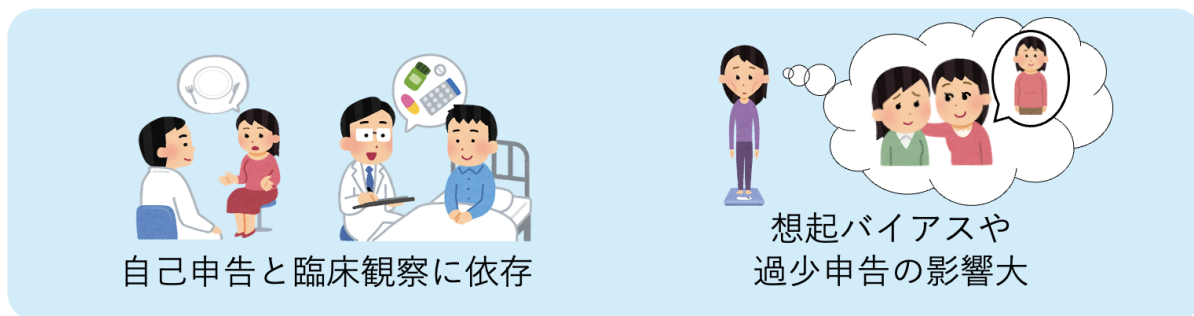


図 2 従来の摂食障害診断の問題点

て機械学習を行い、構築したモデルの分類精度を評価したため報告する。

本稿の構成は以下の通りである。2 節では関連研究を紹介する。3 節では提案手法について詳細を説明する。4 節ではモデルの分類精度の評価実験について報告する。最後に 5 節ではまとめを記述する。

2 関連研究

摂食障害の症状の客観的な定量化についての研究は、これまでにソーシャルメディアを活用したものが主に挙げられる。

Chancellor らは Instagram¹ を使用し、ソーシャルメディアにおける精神疾患の重症度 (MIS) を定量化した。Latent Dirichlet Allocation (LDA) を利用して投稿内容をトピックモデル化し、専門家と非専門家によるトピック評価を行い、トピックごとの MIS をスコア付けした。また、過去 7 か月間の MIS データを基に、翌月の MIS を予測する回帰モデルを構築している [7]。

また、Wang らは Twitter(現 X)² を使用し、摂食障害に苦しむ個人のコミュニティ構造と相互作用について調査している。サンプリングされた摂食障害ユーザと、非摂食障害ユーザに対して収集された 2 つの参照データセットとの社会的ステータス、行動パターン、心理測定特性における差異を定量化し、摂食障害ユーザと非摂食障害ユーザを分類する予測モデルを構築している [8]。

摂食障害と関連する精神疾患として議論される、うつ病については以下のような研究がある。

阪口らはアンサンブル学習を用いて、スマートフォンの会話、位置情報、ロックなどのデータから 9 つの特徴量を抽出し、うつ状態を判別するモデルを構築している。結果、会話、位置情報、スマートフォンのロックから抽出した特徴量はうつ状態の判別に有用であるとしている [9]。

田近らは開発したスマートフォンアプリとウェアラブルデバイスによる日々の記録のデータから、うつ病の再発予測モデルを構築している [10]。

Moshe らはスマートフォンとウェアラブルデバイスを用いてうつ病と不安障害の症状を予測可能か評価する研究を行った。GPS とウェアラブルデバイスのデータでメンタルヘルスを予測

できること、GPS データのみの場合と比較して、GPS とウェアラブルデータを組み合わせたモデルでは、うつ病や不安の症状を予測する能力が大幅に向上したという結果であった [11]。

上述の通り、摂食障害の診断について客観的指標を用いた定量化は行われつつあるが、まだ一部のソーシャルメディアの活用に限られている。また、うつ病などの他の精神疾患ではデジタルバイオマーカーを用いた予測モデル開発が研究されており、これは摂食障害にも応用できると考える。そこで本研究では、スマートフォンのセンサデータを活用することで、より一般的な客観的指標を用いた摂食障害の予測モデルを構築し、その評価を行った。

3 デジタルバイオマーカーを用いた摂食障害の予測モデルの構築

本節では、デジタルバイオマーカーを用いた摂食障害の予測モデルの構築方法について説明する。摂食障害と、正常な食事をしている対照群から、スマートフォンを使用し、加速度計の測定値、GPS の位置情報、アプリケーションの使用状況などのデータを収集する。これらのデータを統合して機械学習モデルに組み込むことにより、高度な行動パターンの分類が可能となるのではないかと考える (図 3)。

3.1 予測モデル構築に利用するデータ

予測モデルの開発に使用するデータは、スマートフォンアプリ、AWARE³ を使用してイギリスで収集された 40 人分のデータである。収集期間は 2024 年 5 月 6 日から 2024 年 5 月 12 日の 7 日間であり、40 人のうち摂食障害患者は 24 人であった。データには位置情報、スクリーンのオン/オフ時間、ジャイロセンサデータ、加速度センサデータ、スマートフォンアプリの使用状況、キーボード入力状況などが含まれる。収集データの内容を表 1 に示す。

位置情報からは行動パターンや訪問場所の分析が可能である。例えば、過食症患者の特異な食事行動を特定することができると考えられる。また、加速度センサデータを活用することで、身体活動の低下や活動リズムの変調を検出することができる。スマートフォンアプリをどのように使用し、どのくらいの頻度で使用しているかは、その人の精神的健康や行動について

1 : Instagram <https://www.instagram.com/>

2 : X <https://x.com/>

3 : AWARE <https://awareframework.com/>

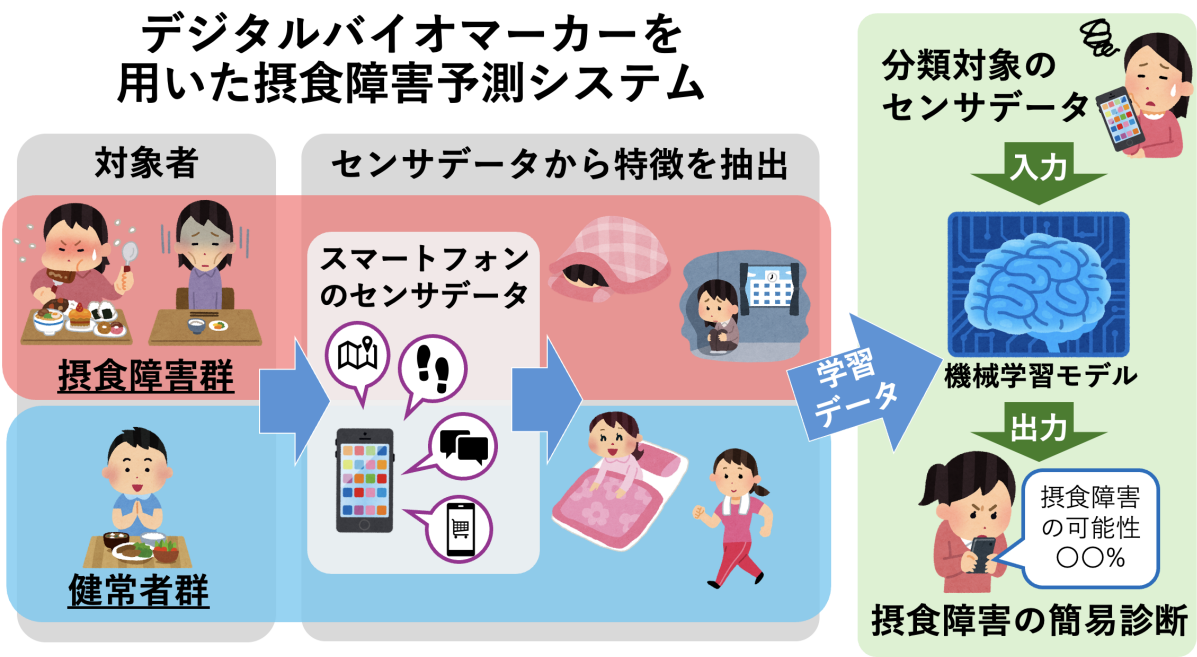


図 3 デジタルバイオマーカーを用いた摂食障害の予測システムの概要

表 1 収集したセンサデータ一覧

スマートフォンのセンサデータ
ジャイロセンサ
加速度計
位置情報
スクリーンのオン/オフ時間
フォアグラウンドでのアプリの使用状況
アプリの通知
Bluetooth
バッテリーの充電状況
バッテリーの放電状況
バッテリーの状態
画面のタッチ情報
画面の明るさ
キーボード入力情報

貴重な手掛かりになると予想する。例えば深夜に SNS を長時間見る、過剰な頻度でスマートフォンを使用する (充電を頻繁に行う)、といった行動が関連する可能性があるのではないかと考える。

3.2 機械学習を用いた予測モデルの構築

分析を行うために、カテゴリ項目はベクトル化し、機械学習モデルが処理できるように前処理を行う。例えば、ヘルスケア、SNS、ショッピング、エンタメ、といったカテゴリでスマートフォンアプリを分類することで、アプリの使用状況にどのような傾向があるのかが予測できると考える。そして、LSTM を使用し、24 時間、12 時間、6 時間など、データを時間単位で区切り、それぞれの時間粒度で、交差検証を行うことでモデルに対する評価を行う。これにより、夜間に SNS を見る頻度が高いといった特徴や、平日と休日の行動の違いなどが分析でき、摂

食障害かどうかの予測に役立つのではないかと考える。

4 摂食障害予測モデル評価実験

本実験では、スマートフォンから取得した行動特徴量の時系列データを用い、時間粒度の違いが摂食障害の分類性能に影響を与えるか検証した。特に、行動データを集約する時間単位を変化させた場合に、識別性能がどのように変化するか比較を行った。また、特徴量の寄与度を可視化し、どのセンサデータが摂食障害の判別に寄与しているかを確認した。

4.1 実験概要

データは 2024 年 5 月 6 日から 2024 年 5 月 12 日の 7 日間でスマートフォンアプリ AWARE を使用してイギリスで収集された 40 人分のデータに対して分析を行う。40 人のうち摂食障害患者は 24 人であった。センサデータはジャイロセンサ、加速度計、スクリーンのオン/オフ時間、フォアグラウンドでのアプリの使用状況、アプリの通知、バッテリーの充電状況、バッテリーの放電状況、バッテリーの状態、画面のタッチ情報、画面の明るさキーボード入力情報を使用した。時間粒度として、1 時間、2 時間、3 時間、4 時間、6 時間、12 時間、24 時間の 7 条件を設定した。分類モデルには、時系列データの時間的依存関係を捉えることが可能な Long Short-Term Memory (LSTM) ネットワークを用いる。学習および評価には、交差検証を用いた。評価指標は、分類性能を多角的に評価するため、AUC、Accuracy、Precision、Recall、F1-score を用いた。なお、前処理として、学習データに基づく欠損補完および標準化を行い、評価データには同一の変換を適用した。

また、各時間粒度でどの特徴量に注目して予測を行っているか、またどの時間帯が摂食障害の予測に大きく寄与してい

るかを可視化するため、fold が 1 の時の SHAP 値（SHapley Additive exPlanations）を用いた解析を行った。

4.2 結果

4.2.1 各時間粒度での LSTM の学習結果

表 2 各時間粒度の LSTM での学習結果

時間粒度	AUC	Accuracy	F1-score	Precision	Recall
1 時間	0.538	0.564	0.560	0.723	0.503
2 時間	0.641	0.614	0.660	0.691	0.642
3 時間	0.661	0.623	0.665	0.719	0.641
4 時間	0.646	0.601	0.652	0.693	0.644
6 時間	0.660	0.580	0.610	0.672	0.595
12 時間	0.586	0.567	0.628	0.575	0.756
24 時間	0.588	0.523	0.526	0.575	0.544

表 2 に各時間粒度における分類性能の比較結果を示す。評価は被験者単位の交差検証に基づいて行い、AUC, Accuracy, Precision, Recall, F1-score の平均値および標準偏差を算出した。2 時間から 4 時間の粒度において AUC および F1-score が高く、特に 3 時間粒度で最も高い AUC(0.661) と F1-score(0.665) が得られた。一方、1 時間の粒度では AUC が 0.538, F1-score が 0.560 と低く、また 24 時間粒度では AUC が 0.588, F1-score が 0.526 とどちらも低下する傾向が確認された。

4.2.2 SHAP による寄与度可視化の結果

時間粒度を 1 時間、6 時間、12 時間、24 時間とした場合の SHAP ヒートマップを図 4, 図 5, 図 6, 図 7 に示す。



図 4 時間粒度 1 時間における SHAP ヒートマップ (時間*特徴量)



図 5 時間粒度 6 時間における SHAP ヒートマップ (時間*特徴量)

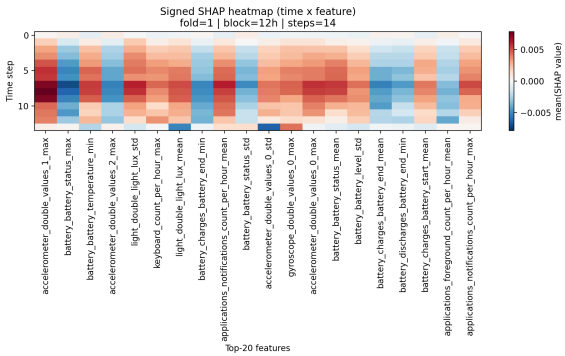


図 6 時間粒度 12 時間における SHAP ヒートマップ (時間*特徴量)

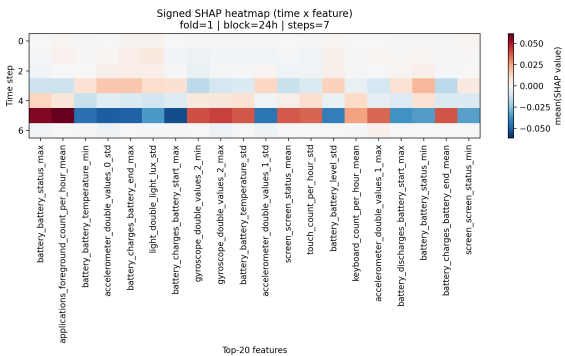


図 7 時間粒度 24 時間における SHAP ヒートマップ (時間*特徴量)

1 時間の粒度では、時間ステップ数が多く（168 step）、寄与は時間方向に広く分散しており、特定の時間帯や特徴量に強く集中する傾向は確認されなかった。モデルがどの時間情報を重視しているかを明確に捉えることは困難であった。バッテリー関連や加速度センサ、ジャイロセンサといった特徴量が上位に現れていた。

6 時間の粒度では、系列後半において SHAP 値の絶対値が大きくなる傾向が明確に確認された。特に、バッテリー関連特徴量、加速度センサ、通知・アプリ利用に関する特徴量が上位に多く現れていた。1 時間粒度と比較してバッテリー関連特徴量が上位に多く見られた。

12 時間の粒度では、時間ステップ数が減少したことにより、寄与が比較的明瞭に可視化され、特徴量間の差がやや強調される傾向が見られた。6 時間と同じく、バッテリー関連特徴量、加速度センサ、通知・アプリ利用に関する特徴量が上位に多く現れていた。

24 時間の粒度では、特定の 1 ステップにおいて強い寄与が集中する様子が確認された。時間方向の解像度が低く、寄与が発生した具体的な時間帯を詳細に把握することは困難である。バッテリー関連特徴量、加速度センサ、ジャイロセンサといった特徴量が上位に多く現れていたが、通知・アプリ利用に関する特徴量が少なくなった。

4.3 考察

4.3.1 各時間粒度での LSTM の学習結果に対する考察

時間粒度が分類性能に与える影響を分析した結果、時間粒度

が細かすぎる場合と粗すぎる場合のいずれにおいても性能が低下する傾向が確認された。1 時間粒度では、行動の変化を細かく捉えることができる反面、短時間の行動変動やノイズの影響を受けやすく、安定した特徴表現が困難であったと考えられる。一方、12 時間および 24 時間粒度では、摂食障害に特有な短時間の行動変化や時間帯依存の特徴が平均化され、摂食障害に関連する行動パターンが十分に表現されなかった可能性が考えられる。これに対し、2 時間から 6 時間程度の時間粒度では、ノイズの抑制と行動変化の保持が適切に両立され、安定した分類性能が得られたと考えられる。2 時間、3 時間、4 時間といった時間幅は、外出や授業や休息といった、日常行動のまとまりを適切に表現できる粒度であり、摂食障害に関連する行動パターンの差異を効果的に捉えられた可能性がある。6 時間単位では生活リズムに基づく大まかな行動傾向を捉えられるものの、時間解像度の低下により、細かな行動差異が失われた可能性がある。その結果、AUC は高いものの、分類精度全体としてはやや低下したと考える。これらの結果は、摂食障害の予測において、日常行動の短期的な変動を適切な時間幅で捉えることが重要であることを示唆している。

4.3.2 SHAP による寄与度可視化の結果に対する考察

SHAP による可視化結果から、時間粒度の違いが、モデルが捉える「重要な時間情報の構造」に大きく影響することが明らかとなった。また、モデルが系列全体を一様に利用するのではなく、一部の行動情報を重視して予測を行っていることが示唆された。1 時間の粒度では、時間方向の解像度が高い一方で、寄与が時間全体に分散し、ノイズ的な影響を受けやすい状態となっていた。このことは、短時間の行動変動が必ずしも摂食障害の判別に直結しない可能性を示唆している。12 時間および 24 時間粒度では、時間方向の情報が大きく集約されるため、寄与は特定の時間ステップに集中しているが、行動変化の細かな推移を捉えることが困難である。特に 24 時間粒度では、どの時間帯の行動が重要であったかを詳細に解釈することは難しく、解釈性の面では制約が生じると考えられる。どの時間粒度でも共通して、バッテリー関連特徴量、加速度センサ、通知・アプリ利用に関する特徴量、ジャイロセンサが上位に見られた。そのため、摂食障害のユーザの端末使用状況や活動量、生活リズムに関する情報が予測に重要な役割を果たしていると考えられる。これらは、摂食行動や生活習慣の変化と間接的に関連している可能性がある。さらに、今回上位に現れた特徴量に絞って学習を行うことで、摂食障害と関連が薄い特徴量によるノイズを抑制でき、モデルの精度が向上する可能性があると考えられる。

5 おわりに

本稿では、スマートフォンのセンサデータを用いて摂食障害の予測モデルの構築とその分類精度の評価を行った。結果として、時間粒度の違いによって分類性能に差異が生じ、3 時間から 6 時間での分類精度が高いことが確認された。一方、1 時間粒度では性能が低下し、24 時間粒度では情報の過度な集約によ

り分類性能が低下する傾向が確認された。さらに、SHAP を用いた可視化により、バッテリー関連特徴量、加速度センサ、通知・アプリ利用に関する特徴量、ジャイロセンサといった特徴量が影響していることが確認できた。

今後は、SHAP で選定された特徴量の数を細かく調整し、最適な特徴量数を探索する予定である。これにより、性能向上とモデル簡素化の両立が期待される。他の時系列を考慮するモデルを検討することでも精度向上や、出力に対する説明性の向上を図っていきたい。

謝 辞

本研究の一部は、京都産業大学先端科学技術研究所（人間情報学研究センター）共同研究プロジェクト（M2301）および京都工芸繊維大学 KYOTO AGORA 2023 プロジェクト（生成型 AI と人間の共創推進勉強会）の助成を受けたものである。ここに記して謝意を表す。

文 献

- [1] 厚生労働省. 厚生労働科学研究成果データベース 摂食障害の診療体制整備に関する研究. <https://mhlw-grants.niph.go.jp/project/26089>, 2016.
- [2] 田上哲也 and 葛谷英嗣. 中枢性摂食異常症. *医療*, 55(12):624–627, 2001.
- [3] Agneta M. Rosling, Pär Sparén, Claes Norring, and Anne-Liis von Knorring. Mortality of eating disorders: A follow-up study of treatment in a specialist unit 1974 – 2000. *International Journal of Eating Disorders*, 44(4):304–310, 2011.
- [4] 山中学, 宮坂 菜穂子, 吉内 一浩, 佐々木 直, 野村 忍, and 久保木 富房. 大学生の摂食障害 (大学生のメンタルヘルスと心身症). *心身医学*, 40(3):215–219, 2000.
- [5] Anjali Devakumar, Jay Modh, Bahador Saket, Eric P. S. Baumer, and Munmun De Choudhury. A review on strategies for data collection, reflection, and communication in eating disorder apps. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.
- [6] Antoine Piau, Katherine Wild, Nora Mattek, and Jeffrey Kaye. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild alzheimer disease and implications for clinical care: Systematic review. *J Med Internet Res*, 21(8):e12785, Aug 2019.
- [7] Stevie Chancellor, Zhiyuan Lin, Erica L. Goodman, Stephanie Zerwas, and Munmun De Choudhury. Quantifying and predicting mental illness severity in online pro-eating disorder communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, page 1171 – 1184, New York, NY, USA, 2016. Association for Computing Machinery.
- [8] Tao Wang, Markus Brede, Antonella Ianni, and Emmanouil Mentzakis. Detecting and characterizing eating-disorder communities on social media. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, WSDM '17, page 91 – 100, New York, NY, USA, 2017. Association for Computing Machinery.
- [9] 阪口 航太, 佐久間 拓人, and 加藤 昇平. アンサンブル学習を用いたスマートフォンログからのうつ状態判別モデル. In 第 85 回

全国大会講演論文集, volume 2023, pages 645–646, feb 2023.

- [10] 田近 亜蘭, 熊谷 成将, and 古川 壽亮. スマートフォンアプリとウェアラブルデバイスを用いた、寛解期のうつ病患者の再発予測. *予防精神医学*, 7(1):3–12, 2022.
- [11] Isaac Moshe, Yannik Terhorst, Kennedy Opoku Asare, Lasse Bosse Sander, Denzil Ferreira, Harald Baumeister, David C. Mohr, and Laura Pulkki-Råback. Predicting symptoms of depression and anxiety using smartphone and wearable data. *Frontiers in Psychiatry*, 12, 2021.