

LightGBM 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

阪口 航太[†] 佐久間 拓人[†] 加藤 昇平[†][†]名古屋工業大学大学院工学研究科工学専攻

1 はじめに

昨今の COVID-19 のパンデミックの影響で、精神疾患の罹患者が急激に増加している。Kessler ら [1] によると、精神疾患の罹患経験のある者のうち 3/4 は 24 歳までに発症しており、特に若い世代にとって精神疾患に関する問題は重大である。自身でうつ状態のセルフチェックを行う方法として Patient Health Questionnaire-9 (PHQ-9) がある。しかし PHQ-9 は自記式の質問票であることから、回答者自身が質問票に回答する意思が必要であるため、より非介入的なスクリーニングシステムであることが好ましい。

医療機関への受診を拒む者や、うつ状態の自覚がない者への有効なスクリーニング手法を目指し、スマートフォンで取得したデータからうつ状態を予測する研究がある。スマートフォンはユーザの行動や心理状態によって使用傾向が異なるため、スマートフォンで取得したライフログ (スマートフォンログ) によるうつ状態のスクリーニングの可能性が示唆されている [2]。

Wahle ら [3] はスマートフォンの使用傾向、加速度、Wi-Fi、GPS に関連する特徴量から、RBF カーネル付き Support Vector Machine (SVM) を用いたうつ病の判別モデルを提案している。また山本ら [4] はスマートフォンで取得したセンシングデータや端末ログからストレスと関連する特徴量を生成し、アンサンブル学習を用いたストレス状態の判別モデルを提案している。本稿では、高速かつ高性能な勾配ブースティング手法として知られる LightGBM を用いてスマートフォンログからのうつ状態判別モデルを構築し、モデルの性能評価について関連研究との比較により考察する。

2 方法

2.1 データセット

本研究では、スマートフォンを用いてメンタルヘルス、学業成績、行動傾向を調査した研究で収集されたオープンデータセット StudentLife dataset [5] を用いる。Wang ら [5] はダートマス大学の 2013 年の春学期におけるコンピュータサイエンスの授業を受講した学生のうち、任意で研究に参加した 60 名を実験協力者とした。スマートフォンログの計測は 10 週間に及び、各実験協力者に貸与した Nexus4 の端末を使用してデータを計測した。センシングデータとして、Bluetooth、会話、位置情報、スマートフォンのロック、Wi-Fi などのスマートフォンログを収集した。スマートフォン

Depression Discrimination Model from Smartphone Logs Using LightGBM

Kota SAKAGUCHI[†], Takuto SAKUMA[†], Shohei KATO[†][†]Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology[†]Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

{ksakaguchi, sakuma, shohey}@katolab.nitech.ac.jp

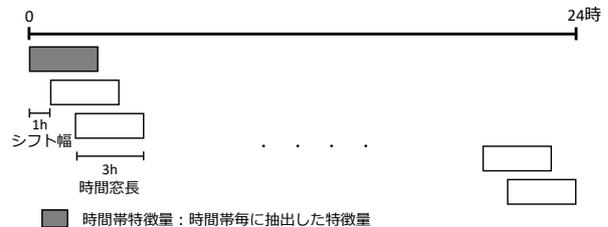


図 1: 時間帯特徴量の抽出方法

ログの収集及びクラウドへのアップロードはユーザの操作を介さず、一時的にスマートフォンに保存された後に自動的にクラウドへアップロードされた。また実験協力者は 10 週間の計測期間の前後で PHQ-9 に回答した。

本研究では StudentLife dataset のうち、スマートフォンログの計測期間終了時における PHQ-9 の回答直前で収集した 38 名の 14 日間を解析対象とする。解析対象における会話、位置情報、スマートフォンのロックに関する特徴量と計測期間終了後の PHQ-9 の回答データから、本研究で用いるデータセットを再構成する。解析対象の各 1 名について、1 日あたりのスマートフォンログを 1 サンプルとし、後述の計 117 次元の特徴量を抽出する。続いて、PHQ-9 において一般的に用いられる基準 (スコアが 10 以上か否か) により、各サンプルを「うつ状態」「非うつ状態」の 2 群でアンテーションする。欠損値については、PHQ-9 のデータあるいは特徴量すべてが欠損しているサンプルを解析対象から除外し、部分的に欠損しているサンプルは欠損値を 0 で補完している。以上の手続きにより、7 名から構成されたうつ状態群は 69 サンプル、31 名から構成された非うつ状態群は 308 サンプルからなるデータセットへ再構成した。

2.2 特徴量

本節ではスマートフォンログから抽出した特徴量について詳述する。会話に関する特徴量は、会話の時間情報から会話時間および会話回数の 2 次元を抽出する。位置情報に関する特徴量は、GPS の緯度・経度を基に移動回数、滞在時間、位置座標変位の 3 次元を抽出する。スマートフォンのロックは 1 時間以上ロックされている間に、ロックの継続時間が記録されている。そこでスマートフォンのロックに関する特徴量は、ロックの継続時間からロック時間およびロック回数の 2 次元を抽出する。抽出し前述の 7 次元の特徴量に対して、図 1 に示すように時間帯ごとの特徴量 (時間帯特徴量) と統計量 (分散、傾き、1 日あたりの総和) を抽出する。以上の計算処理により、スマートフォンログから計 117 次元の特徴量を抽出する。

表 1: ハイパーパラメータの探索範囲

ハイパーパラメータ	探索範囲
reg_alpha	0.0001
reg_lambda	0.0001, 0.001
num_leaves	2, 4, 8
colsample_bytree	0, 3, 6
subsample	0.7, 0.8, 0.9
subsample_freq	4, 6
min_child_samples	2, 5
feature_fraction	0.9

2.3 判別モデル

本研究では LightGBM を用いたうつ状態の判別モデルを構築し、層化 5 分割交差検証によりモデルの性能を評価する。本研究で用いるデータはうつ状態群（サンプル数： n_1 ）が非うつ状態群（サンプル数： n_2 ）に比べてサンプル数が少なく（ $n_1 < n_2$ ）不均衡である。そのため、学習用データにおけるうつ状態群と非うつ状態群に含まれるサンプルに対し、それぞれ式 (1) の $weight_1$, $weight_2$ の重みを付与して学習することで、データの偏りを是正する。ただし非うつ状態群に比べ、うつ状態群の判別を重視するため $\alpha = 2$ とする。

$$weight_1 = \alpha \frac{1}{n_1} \quad weight_2 = \frac{1}{n_2} \quad (1)$$

表 1 にハイパーパラメータのグリッドサーチの範囲を示す。ハイパーパラメータは、層化 5 分割交差検証のホールド毎に探索する。117 次元の特徴量全てがうつ状態の判別に寄与するとは限らない。そこで、特徴重要度を基に特徴量選択を実施し、モデルを構築した。

3 結果と考察

3.1 モデルの性能評価に関する結果と考察

モデルの性能評価指標として、G-mean、感度、特異度を算出する。G-mean は不均衡なデータを用いた判別モデルの総合的評価指標にしばしば用いられ、感度と特異度の積による平方根で算出される。表 2 に関連研究との比較結果を示す。関連研究と比較すると G-mean は本研究のものが最も良好であり、うつ状態判別モデルの有効性が示唆された。また感度と特異度の差について着目すると、関連研究の 2 手法においては 0.16 以上であるが、本研究では差が見られなかった。本研究の判別モデルが、うつ状態群と非うつ状態群の判別性能において最もバランスが良いことが示唆された。

3.2 特徴量選択に関する結果と考察

表 3 に、選択された 32 次元の特徴量のうち特に重要度の高い特徴量を示す。上位 5 次元の特徴量のうち、3 次元が「位置座標変位」に関わる特徴量であった。「位置座標変位」の重要度が高い理由として、「集中力の低下」や「不眠」等のうつ状態を捉えた特徴量であることが考えられる。また「会話」に関する特徴量も上位で選択された。「会話」に関する特徴量について分析すると、非うつ状態群に比べてうつ状態群で会話の減少が確認された。この傾向は「会話が減る」といううつ状態の傾向を捉えており、判別に寄与したと考えられる。今後性能向上を目指すうえで、うつ状態を捉えた特徴量抽出が重要である。

表 2: 関連研究との比較結果

評価指標	G-mean	感度	特異度
本研究	0.71	0.71	0.71
Wahle ら [3]	0.59	0.73	0.47
山本ら [4]	0.68	0.61	0.77

表 3: 上位 5 次元の特徴量とその重要度

特徴量	特徴重要度
位置座標変位の分散	14
会話回数の分散	10
位置座標変位	10
位置座標変位 (3-6 時)	7
会話時間	6

4 まとめと今後の展望

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によるうつ状態の判別の結果、G-mean、感度、特異度で 0.71 を確認した。今後の展望として、うつ状態の傾向を捉えた新たな特徴量の抽出による性能向上を目指す。また、ハイパーパラメータのグリッドサーチにおいて Optuna を用いることで、より効率的に探索する。そして、多くの人々が携帯するスマートフォンで得られる情報からうつ状態を判別し、精神科や心療内科への受診誘導に貢献するシステムを目指す。

謝辞

本研究は、一部、文科省科研費 (JP19H01137, および, JP20H04018), ならびに、NICT 委託研究の助成により行われた。

参考文献

- [1] Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R. and Walters, E. E.: Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication, *Archives of general psychiatry*, Vol. 62, No. 6, pp. 593–602 (2005).
- [2] Sakaguchi, K., Sakuma, T. and Kato, S.: Underbagging-Based Ensemble Learning approach to Detection of Depression from Smartphone Logs, in *2023 IEEE 12th Global Conference on Consumer Electronics (GCCE 2023)*, pp. 1030–1031 (2023).
- [3] Wahle, F., Kowatsch, T., Fleisch, E., Rufer, M., Weidt, S., et al.: Mobile sensing and support for people with depression: a pilot trial in the wild, *JMIR mHealth and uHealth*, Vol. 4, No. 3, p. e5960 (2016).
- [4] 山本直樹, 落合桂一, 稲垣章弥, 深澤佑介, 木本勝敏, 霧生和樹, 上西康平, 太田順, 沖村宰, 寺澤悠理他: 生体情報を教師データとしたスマートフォンログを用いたストレス状態推定, 研究報告モバイルコンピューティングとパーベイシブシステム (MBL), Vol. 2018, No. 33, pp. 1–7 (2018).
- [5] Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D. and Campbell, A. T.: StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones, in *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, pp. 3–14 (2014).