

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

阪口 航太[†] 佐久間 拓人[‡] 加藤 昇平[‡][†]名古屋工業大学工学部情報工学科[‡]名古屋工業大学大学院工学研究科工学専攻

1 はじめに

近年、うつ病患者の増加が問題となっている。Goodwinら[1]によると、アメリカ合衆国における過去1年のうつ病罹患者の人口比率の推移は2015年から2020年にかけて7.3%から9.2%に増加している。年代別に見ると、18~25歳では最も高い17.2%であることから、若い世代にとってうつ病は特に大きな問題である。自身の心理状態を正確に判断するには精神科や心療内科を備えた医療機関での診断が必要となる。しかしながら、こうした機関の受診に対する抵抗感から、受診すべき状態の人々が受診に至っていない。

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。PHQ-9は、アメリカ精神医学会が出版している精神疾患の診断基準 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition: DSM-5) による9つの診断基準に対して、それぞれの項目ごとにスコア化した自記式質問票である。しかしPHQ-9は回答者自身が質問票に回答する意思が必要であるため、うつ状態の自覚に乏しい者にとっては有効な手段ではない。

そこでうつ状態の自覚がない者を評価するため、スマートフォンで取得したデータからうつ状態を推定する研究がある。スマートフォンはユーザにより使用傾向が大きく異なるため、スマートフォンのライフログはうつ状態の推定に有用であると考えられる。したがって、本研究では若年層を対象にスマートフォンで収集したライフログデータを用いて、うつ状態を判別する。

2 関連研究

DaSilvaら[2]は、Wangらによって独自開発されたStudentLife app[3]によりセンシングデータを収集した。またセンシングデータから抽出した特徴量をペナルティ付き一般化推定方程式を用いて分析した結果、会話時間や会話回数、移動距離を含む6次元の特徴がストレスレベルと有意に関連すると結論づけている。またWangら[4]はセンシングデータから得られる特徴量とPHQ-8の全体のスコアの相関関係を評価した結果、スマートフォンのロック解除の回数や会話時間、訪問件数などの特徴に相関が見られた。これらの関連研究からスマートフォンのロックや会話、および位置情報

Depression Discrimination Model from Smartphone Logs Using Ensemble Learning

Kota SAKAGUCHI[†], Takuto SAKUMA[‡], Shohei KATO[‡][†]Dept. of Computer Science and Engineering, Faculty of Engineering, Nagoya Institute of Technology[‡]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

がメンタルヘルスと関連すると推測できる。

3 方法

3.1 データ

本研究では、ダートマス大学におけるスマートフォンを用いてメンタルヘルス、学業成績、行動傾向を評価した研究において公開されたStudentLife dataset[3]を使用した。この研究は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名を対象とした。スマートフォンログの収集にはNexus4sの端末を使用し、10週間に及びデータを計測した。スマートフォン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に関わらず自動的に行われた。PHQ-9の回答は10週間の計測期間の前後の2回でのみ実施された。本研究ではこのデータセットのうち、計測期間終了時にPHQ-9に回答した38名における最後の14日間を解析対象とする。解析対象から会話、GPS、スマートフォンのロックに関する9次元の特徴量と計測期間終了後のPHQ-9の回答データを抽出する。

3.2 本研究で用いるデータセット

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、上記9次元の特徴量を抽出した。各サンプルのラベルはうつ状態/非うつ状態とし、PHQ-9において一般的に用いられる基準により、スコアが10以上か否かで判定した。PHQ-9のデータあるいは9次元の特徴量すべてが欠損しているサンプルは解析対象から除外した。なお9次元の特徴量が部分的に欠損していた場合は0で補完した。こうして、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。

3.3 特徴量

本節ではスマートフォンログから抽出した9特徴量について詳述する。

• 会話に関する特徴量

会話は発生時に開始時刻と終了時刻が記録されており、これをもとに実験協力者ごとに1日あたりの会話回数、総会話時間、平均会話時間の3次元の特徴量を抽出した。

• 位置情報に関する特徴量

位置情報はGPSをもとに20分間隔で収集時刻、プロバイダ、緯度、経度が記録されている。本研究では、緯度と経度をもとに実験協力者ごとに1日あたりの訪問件数、総滞在時間、平均滞在時間、移動距離の4次元の特徴量を抽出した。訪問件数は緯度・経度から算出された位置座標のユニーク

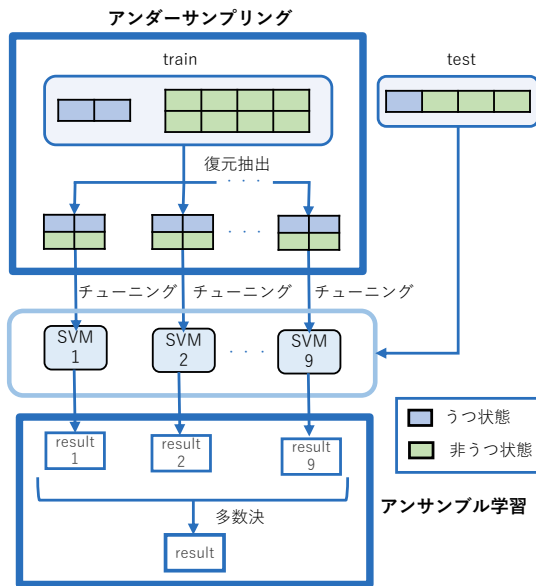


図 1: 判別モデルの概要図

な数とし、滞在時間は位置座標が不変で継続している時間、移動距離は前後の位置座標における距離の和とした。

・スマートフォンロックに関する特徴量

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開始時刻と終了時刻が記録されており、これをもとに実験協力者ごとに1日あたりのロック回数と総ロック時間の2次元の特徴量を抽出した。

3.4 判別モデル

図1に判別モデルの概要を示す。本研究ではデータの不均衡性を考慮した復元抽出によるアンダーサンプリングと複数の弱学習器の予測値を多数決により統合するアンサンブル学習を組合せたモデルによりうつ状態を判別する。データセットを学習用とテスト用に分割した後、学習用データにおいてアンダーサンプリングを実施した。アンダーサンプリングにおいて、データ数の少ないうつ状態群と同数となる55サンプルのデータを非うつ状態群のデータからランダムに抽出し、データ数が110サンプルのデータセットを9つ作成した。アンサンブル学習における各弱学習器に、9つのデータセットを1つずつ入力し、アンサンブル学習した。弱学習器はRBFカーネル付きSupport Vector Machine (SVM)を用い、ハイパーパラメータは弱学習器ごとに、Cを $[10^{-1}, 1, \dots, 10^3]$ 、 γ を $[1, 10^{-1}, \dots, 10^{-4}]$ の範囲でチューニングした。モデルの性能評価は5-fold Cross Validation法により、G-mean, 感度, 特異度を算出した。G-meanは不均衡なデータを用いた判別性能評価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。

4 結果と考察

図2にテストデータの2群判別における混同行列を、表1に性能評価指標をそれぞれ示す。G-meanは0.691、感度は0.912、特異度は0.529であった。これらの結果から会話、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考えられる。また非うつ状態群の判別性能である特異度が、う

True label	depression	63	6
	non-depression	145	163
		depression	non-depression
		Predicted label	

図 2: 混同行列 (N=377)

表 1: 実験結果

評価指標	G-mean	感度	特異度
平均	0.691	0.912	0.529
標準偏差	0.046	0.116	0.065

つ状態群を指す感度と比較して低値を示している。メンタルヘルスの分野において、うつ状態の者に対する誤判別（偽陰性）は、非うつ状態の者に対する誤判別（偽陽性）に比べ重大である。そのためメンタルヘルス分野への応用を考慮すると、感度が特異度より優位である結果は有用である。しかしながら、特異度については改善の余地がある。そのため、判別に寄与する特徴量を検討し、判別性能を改善していく必要がある。

5 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、良好な感度を確認できた。今後の展望として、特徴量の抽出を改善し、うつ状態群の判別性能を維持しつつ、偽陽性を改善する。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。

参考文献

- [1] Goodwin, R. D., Dierker, L. C., Wu, M., Galea, S., Hoven, C. W., & Weinberger, A. H. (2022). Trends in US Depression Prevalence From 2015 to 2020: The Widening Treatment Gap. *American journal of preventive medicine*, 63(5), 726-733.
- [2] DaSilva, A. W., Huckins, J. F., Wang, R., Wang, W., Wagner, D. D., Campbell, A. T. (2019). Correlates of stress in the college environment uncovered by the application of penalized generalized estimating equations to mobile sensing data. *JMIR mHealth and uHealth*, 7(3), e12084
- [3] Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., ... & Campbell, A. T. (2014, September). 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).
- [4] Wang, R., Wang, W., DaSilva, A., Huckins, J. F., Kelley, W. M., Heatherton, T. F., Campbell, A. T. (2018). Tracking depression dynamics in college students using mobile phone and wearable sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1), 1-26.