

SNS 上のテキストデータを用いたうつ病兆候の検出に関する研究 A Study on the Detection of Depressive Symptoms Using Text Data from Social Media

漆原 匠希[†] 康鑫[†] 松本 和幸[†] 吉田 稔[†]
Shoki Urushihara Kang Xin Kazuyuki Matsumoto Minoru Yoshida

1. はじめに

近年、SNS 投稿からユーザーの心理状態を推定する研究が進んでおり、うつ病の早期兆候の検出は医療や社会的支援の観点から重要である。従来は単一投稿を対象とする分析が主であったが、本研究ではユーザー単位の投稿履歴を時系列で捉えることで、より高精度な分類を目指す。

本研究では、時間情報を考慮した事前学習済み言語モデル ChronoBERT[1]を用い、SNS 投稿からうつ傾向の有無を分類するモデルを構築した。

2. 研究目的

本研究の目的は、SNS 上の投稿履歴に基づき、うつ傾向のあるユーザーを高精度に分類することである。近年、SNS 投稿から精神状態を推定する研究が増えているが、多くは単一投稿に基づいており、長期的な状態変化や行動パターンを十分に捉えられていない。

本手法では、複数投稿における意味的・時間的变化に着目し、ユーザー単位の投稿を系列データとして扱うことで、分類精度の向上を図る。投稿内容の変化傾向や時間帯の偏りといった行動的特徴もモデルに反映される。

各投稿からは ChronoBERT を用いて文脈特徴を抽出し、LSTM-Attention モデル[2]に時系列で入力する構成とした。さらに、投稿時間や単語数、曜日、深夜投稿フラグなどのメタデータも統合し、言語的・行動的情報を併用してユーザー分類を行う。

本研究は、実用性の高い精神状態推定モデルの構築を目指し、将来的なメンタルヘルス支援への応用も視野に入れている。

3. 関連研究

SNS 投稿を用いたうつ傾向検出の研究は近年活発に行われており、特に TUA1 による eRisk タスクへの取り組みでは、TF-IDF や CNN+LSTM といったモデルを用いたアプローチ[3]が提案されている。これらの手法は投稿文の表層的特徴や単語の出現頻度に基づくものであり、時間的情報の活用は限定的であった。

一方、本研究では時間情報を文脈に取り入れる ChronoBERT を用いることで、ユーザーの投稿時刻や深夜活動などの時系列的特徴を含む埋め込みベクトルを生成し、より精度の高い分類を実現している。また、文脈特徴とメタデータを統合した統合ベクトルによる分類手法は、従来の単一特徴に基づく手法と比較して、Precision および Recall の両面で安定したパフォーマンスを示した。

4. 提案手法

本手法では、SNS 上の投稿からユーザーのうつ傾向を分類するために、時間情報を反映した文脈特徴抽出モデルである ChronoBERT を用いるとともに、ユーザーの行動的特徴を表すメタデータ（投稿時刻や投稿頻度など）を併用し、

時系列性を保持した投稿系列に対して LSTM-Attention モデルによる分類を行う。これにより、単一投稿の内容のみならず、ユーザーの投稿履歴に内在する時間的パターンや行動的傾向を捉えることで、より高精度なユーザー単位のうつ傾向分類を実現することを目的とする。

4.1 データセット

2017 年に公開された XML 形式の SNS 投稿データ[4]を使用した。対象は、医療機関の診断結果に基づくうつ病患者の投稿 (pos) および一般ユーザーの投稿 (neg) である。各ユーザーの投稿には、タイトル、本文、投稿日時、およびユーザーID の情報が付与されており、ラベル付きの分類タスクとして扱える構成となっている。

また、全体の投稿数およびユーザー数のバランスにも配慮し、分類モデルの学習において偏りのないサンプリングを行った。特にユーザー単位での均衡を保つため、pos・neg クラスのユーザー数を揃えたうえで分析を進めた。

4.2 前処理

投稿日時は datetime 型に変換し、投稿文はタイトルと本文を連結して一つのテキストとした。ユーザーごとに投稿をまとめ、投稿日時順にソートを行った。また、各ユーザーの投稿数にはばらつきがあるため、系列長を揃えるためのゼロパディングを実施した。

4.3 特徴量抽出

本研究では、各投稿から文脈特徴を抽出するために ChronoBERT を使用した。具体的には、モデルの出力である最終隠れ層 (last_hidden_state) の平均を取り、各投稿の文脈ベクトル (sentence embedding) とした。比較実験では、ChronoBERT と同様に RoBERTa も文脈特徴抽出器として使用し、同様の方法で文ベクトルを取得した。

いずれのモデルでも、抽出した投稿ごとの文脈ベクトルを時系列として構成し、LSTM-Attention モデルに入力することで、ユーザー単位の分類を行った。この手法により、各投稿の意味的内容と、その時間的背景の変化を捉えることが可能になる。

加えて、各投稿について投稿時間（時間帯、曜日、深夜投稿フラグ）や投稿長（文字数、単語数）といったメタデータも抽出した。これらはユーザーの行動傾向や生活リズムを反映する特徴量であり、うつ傾向を捉える補助的な情報として機能することが先行研究により示されている。

各投稿に対して、ChronoBERT によって得られた文脈特徴ベクトルと、投稿時間・曜日・文字数・単語数・深夜フラグなどのメタデータを数値化・正規化の上で結合し、1 投稿あたりの統合特徴ベクトルを構成した。これらをユーザーごとの投稿系列としてまとめ、LSTM-Attention モデルに入力することで分類を行った。

[†] 徳島大学大学院 Tokushima University

4.4 分類モデル

ユーザーごとの投稿系列に対して、双方向 LSTM 層と Attention 機構を備えた LSTM-Attention モデルを構築した。LSTM により投稿の時系列的な変化を捉え、Attention により系列中の重要な投稿に対して重み付けを行うことで、より情報価値の高い特徴量に焦点を当てた分類を実現した。

このモデルにより、単に投稿数が多いユーザーが過学習されることなく、意味的に重要な投稿を適切に捉えられる。また、損失関数にはバイナリクロスエントロピーを用い、最適化には Adam オプティマイザを採用することで、学習の安定性と収束性を確保した。ハイパーパラメータとしては、隠れ層のサイズや LSTM の層数、バッチサイズなどを変えて精度の最適化を図った。

5. 実験と考察

ユーザー単位で投稿系列を分類し、データを 8:2 で学習・テストに分割して評価を行った。

使用モデル：ChronoBERT (文脈特徴抽出) + メタデータ + LSTM-Attention (系列分類)

評価指標：Accuracy, Precision, Recall, F1-score, ROC-AUC
結果:

Accuracy: 84.7%

Precision: 84.0%

Recall: 84.0%

F1-score: 83.6%

ROC-AUC: 91.5%

さらに、ChronoBERT の有効性およびメタデータの有用性を検証するために、以下の比較実験を実施した。表 1、2 に結果を示す。

表 1 異なる特徴量による分類性能の比較

特徴量構成	Precision	Recall	F1-score
文脈特徴	0.72	0.85	0.73
メタデータ	0.67	0.89	0.76
文脈+ メタデータ	0.72	0.85	0.78

表 2 異なる BERT モデルによる分類性能比較

モデル	Precision	Recall	F1-score
RoBERTa	0.71	0.81	0.76
RoBERTa 大型モデル	0.72	0.85	0.78
ChronoBERT	0.72	0.85	0.78

比較対象となる RoBERTa は、事前学習済みモデルを用いて各投稿から文脈特徴ベクトルを抽出するために利用し、分類処理自体は ChronoBERT と同様に LSTM-Attention モデルで行った。RoBERTa における入力データや特徴量抽出の手順、LSTM-Attention の学習条件 (エポック数、バッチサイズ、最適化アルゴリズム) は ChronoBERT と同一に揃えており、公平な比較ができるよう配慮した。また、メタデータと文脈特徴の両方を組み合わせることで、単独よりも

バランスの取れた高い精度が得られることが明らかになった。

次に既存研究との比較実験を実施した。表 3 に結果を示す。

表 3 既存研究と分類性能の比較

モデル	Precision	Recall	F1-score
SVM	0.35	0.71	0.47
CNN+LSTM	0.42	0.76	0.36
本モデル	0.72	0.85	0.78

既存研究と比較しても本研究で提案した ChronoBERT ベースのモデルは優れた F1-score を示しており、文脈と時間情報を融合させた表現の有用性が示された。また、メタデータまたは文脈特徴のどちらか一方だけを使用するよりも、両者を統合することで、より安定した分類性能が得られる点も特筆すべきである。

6. 考察

ChronoBERT を活用して投稿文の文脈特徴を抽出し、それらを LSTM-Attention モデルに時系列として入力することで、時間に依存した心理状態の変化を捉える分類が可能となった。特に、ユーザー単位の時系列分類においては、投稿の間隔や深夜の投稿傾向などがモデル学習に寄与していると考えられる。

また、比較実験からは ChronoBERT が既存のモデルと同等、あるいはそれ以上の性能を持つことが示された。さらに、メタデータと文脈特徴の双方を組み合わせることで、より一貫性のある性能が得られ、単独の特徴量よりも安定した分類が可能となった。

7. 今後の課題

SNS 投稿の多様性や時代背景の変化を踏まえ、今後のモデルにはさらなる汎化能力が求められる。Twitter や Instagram、Threads など複数のプラットフォームに対応し、異なる投稿形式や利用層に対しても安定した性能を示す必要がある。また、うつ傾向は投稿内容だけでなく、投稿頻度や感情の変化、ユーザー間の相互作用にも現れるため、これらの行動的特徴を取り入れることが今後の重要な方向性である。さらに、医療現場での実用化を視野に入れ、誤分類による心理的影響や倫理的リスクに配慮した、安全で信頼性の高いシステム設計も課題となる。

謝辞

本研究は JSPS 科研費 JP24K15193, JST PSI・GAP ファンド 支援プログラム (ステップ 1) の助成を受けたものです。

参考文献

- [1] Songrun He, Linying Lv, Asaf Manela, Jimmy Wu, "Chronologically Consistent Large Language Models", 2025, arXiv:2502.21206.
- [2] Zhiheng Huang, Wei Xu, Kai Yu, "Bidirectional LSTM-CRF Models for Sequence Tagging", 2015, arXiv:1508.01991.
- [3] Ning Liu, Zheng Zhou, Xin Kang, and Fuji Ren, "TUA1 at eRisk", 2018.
- [4] J. Parapar, P. Martín-Rodilla, D. E. Losada, F. Crestani, "Evaluation report of erisk 2022: Early risk prediction on the internet", CLEF (Working Notes), 2022.
- [5] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, Veselin Stoyanov, "RoBERTa: A Robustly Optimized BERT Pretraining Approach", 2019, arXiv:1907.11692.