Indonesian Journal of Data and Science

Volume 6 Issue 2
ISSN 2715-9936
https://doi.org/10.56705/ijodas.v6i2.265

Research Article

Hybrid CNN-LSTM and Cox Model for Bipolar Risk Analysis Using Social Media Data

Rizki Amanda [1,*], Jasmine Aulia [2]

[1] Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia, 22083010045@student.upnjatim.ac.id
[2] Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia, 22083010074@student.upnjatim.ac.id

Correspondence should be addressed to Rizki Amanda; 22083010045@student.upnjatim.ac.id

Received 10 February 2025; Accepted 11 May 2025; Published 31 July 2025

Abstract:

Introduction: Mental disorders such as bipolar disorder are becoming increasingly prominent, particularly with the rise of emotional expression through social media. Early detection remains a significant challenge due to the lack of non-invasive, real-time assessment methods. Methods: This study proposes a hybrid deep learning approach combining Convolutional Neural Network–Long Short-Term Memory (CNN-LSTM) and the Cox Proportional Hazards (Cox PH) model to analyze the risk and timing of bipolar disorder onset. A dataset of 3,511 tweets from 517 Twitter users was collected. The CNN-LSTM model classified bipolar risk levels based on text data, while the Cox PH model estimated the time-to-event for high-risk conditions using behavioral features and predicted risk labels. Results: The hybrid model demonstrated strong predictive performance. The risk label significantly influenced the time to high-risk condition (hazard ratio = 5.39, $p < 0.005$). The model achieved a concordance index (C-index) of 0.816, indicating high reliability in survival prediction. Conclusions: This case study highlights the potential of integrating deep learning and survival analysis for early bipolar disorder detection using social media data. The proposed non-invasive method can support mental health monitoring while raising awareness of ethical and privacy considerations.

Keywords: Risk Analysis, Mental Health, Bipolar, CNN-LSTM, Cox Proportional Hazard.
Dataset: https://drive.google.com/file/d/1xViUfmNcaihLs46z9HcDUFZPvmCOdf85/view?usp=sharing

1. Introduction

Mental health is a serious issue in Indonesia, including East Java, one of the provinces on the island of Java. According to the Basic Health Research, the prevalence of severe mental illnesses such as schizophrenia and other psychotic disorders in East Java is 2.2 people per mille, which is higher than Indonesia's national prevalence rate of 1.7 [1]. Mental disorders affect a person's feelings, mood, thinking, or behavior and include illnesses such as bipolar disorder, anxiety, depression, and schizophrenia [2]. Bipolar disorder is a chronic mood disorder that causes intense changes in mood, energy levels, and behavior. Manic and hypomanic episodes are an important feature of the illness, and many people with bipolar disorder also experience depressive episodes. The illness can be managed with medication, talk therapy, lifestyle changes, and other treatments [3]. However, there are still many individuals who do not realize they have bipolar disorder until it reaches the acute phase, due to lack of awareness and the strong social stigma towards mental health. Early detection and estimation of the time of onset of mental disorders is crucial to ensure appropriate interventions and prevent more severe impacts on individuals who experience them.

The development of digital technology and social media opens up new opportunities for early detection of mental disorders. Platforms such as Twitter and Facebook are often used by individuals to openly express feelings, opinions, and personal experiences that can be analyzed to detect early symptoms of mental disorders. This research provides a roadmap for analysis, in which the detection of mental disorders can be based on machine learning techniques. We describe a general approach to predict and identify these disorders using social network data [4]. Therefore, this study aims to conduct risk analysis and estimation of the time of occurrence of mental disorders in bipolar disorder using a hybrid approach of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) for classification, and Cox Proportional Hazard (Cox PH) for estimation of time of occurrence based on identified risk factors.

In a study conducted by Bunga and Yuliant (2023), the CNN-LSTM model was applied to the classification of hate speech on Twitter, where CNN was used to extract textual features and LSTM was used to understand the word order in tweets. This approach allows for a more in- depth analysis of the risk of mental disorders, including bipolar disorder, based on the text posted by individuals on social media [5]. The effectiveness of CNN-LSTM in capturing both spatial and temporal aspects of sequential data has also been demonstrated in various mental health studies, such as the classification of mental stress using ECG signals [6]. Meanwhile, the Cox Proportional Hazard model is widely used in survival analysis to estimate the time until the occurrence of specific events, including health outcomes. Katzman et al. introduced DeepSurv, a deep learning-based Cox model that improves predictive accuracy in clinical risk assessment [7]. The integration of these models enables a comprehensive approach for identifying individuals at higher risk and estimating the time to onset of mental disorders, which can support early intervention strategies.

In addition, Cox Proportional Hazard was used to estimate the time to onset of mental disorders based on the analyzed risk factors. This model allows researchers to understand the factors that influence the duration until a person experiences an episode of mental disorder, as well as identify individuals at higher risk. The use of Cox models in the context of mental disorders allows for more accurate estimation of time to onset and is useful for early treatment. Several recent studies support this application. For instance, Huang et al. (2024) used the Cox model to assess the risk of psychiatric disorders following infertility diagnosis, considering socioeconomic and medical comorbidity factors [8]. Similarly, Son et al. (2023) applied the Cox Proportional Hazards model to evaluate dementia-free survival and its risk factors in older adults in Korea [9]. Furthermore, Yuan et al. (2021) employed Cox regression to predict major depressive disorder using electronic health records and temporal features [10]. Another study by Yang et al. (2020) analyzed suicide attempts among psychiatric patients using Cox models to identify key timing and risk elements [11]. These studies demonstrate the growing relevance of Cox models in psychiatric epidemiology and time-to-event analysis related to mental health.

However, most existing studies either focus solely on classification without estimating the time to onset, or apply survival analysis independently without leveraging the sequential nature of behavioral data derived from social media. Moreover, few works have explored the integration of deep learning models such as CNN-LSTM with survival models like Cox PH to perform a joint analysis of risk detection and time prediction in the context of bipolar disorder. This creates a research gap that this study seeks to address: combining text-based risk classification and time-to-event estimation to provide a comprehensive early detection framework using social media data.

2. Method:

Research methods are the techniques, tools, and procedures used to collect, analyze, and interpret data for a study. They provide a systematic approach to answering research questions and ensuring that results are accurate, reliable, and relevant. Choosing the right research methodology is crucial to the success of a project, as it determines how data is collected and interpreted. In contemporary studies, research often involves statistical modeling and machine learning techniques, such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Cox Proportional Hazard models, especially for time-dependent or sequential data. These methods require structured and well-prepared datasets and typically produce results in the form of numerical values or quantitative predictions. For instance, deep learning has been integrated with traditional statistical models in biomedical contexts to enhance interpretability and prediction performance [12]. Hybrid models such as LSTM-Cox have also been developed to handle survival data with sequential characteristics [13], while CNN and LSTM architectures are frequently applied for classification and temporal analysis tasks in mental health and clinical studies [14]. Moreover, models like DeepSurv demonstrate how deep neural networks can be combined with Cox regression to provide personalized risk estimations [15]. Furthermore, comparative frameworks have shown the practical integration of machine learning and statistical survival models for hospital-based outcomes, with DeepSurv exhibiting superior performance in predicting patient mortality post-admission [16].

A research design is a systematic plan that describes how a study will be conducted, including data collection methods, procedures, and analytical tools. It ensures that the research remains focused, feasible, and ethical by matching the research question with appropriate methods. A well-chosen research design enhances the accuracy, reliability, and relevance of findings, particularly in studies involving structured methodologies like statistical modeling and machine learning [17]. Recent studies have emphasized the growing importance of aligning research questions with methodological rigor, especially in interdisciplinary and applied fields. For instance, an article highlights the necessity of selecting research designs that effectively answer research questions to improve validity and reliability [18]. Similarly, a chapter on research design discusses the importance of collecting relevant evidence to evaluate problems, test theories, or describe phenomena, reinforcing the role of systematic planning in social sciences research [19].

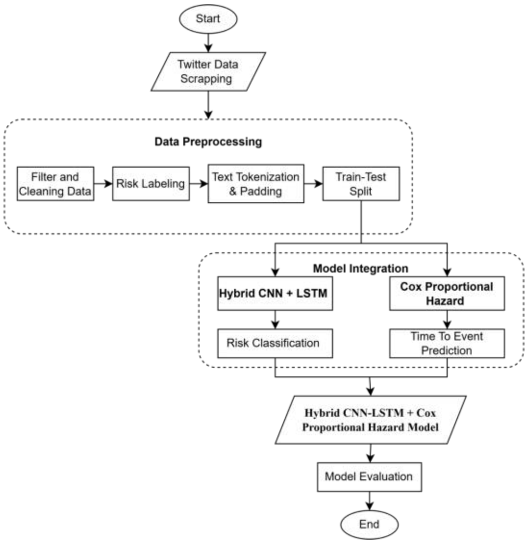

Figure 1. Mental Health Risk Analysis Research Design

Data Collection

Data collection involves gathering and analyzing information from various sources to address research questions, assess outcomes, and predict trends and probabilities [9]. In this study, the data was extracted from Twitter using web scraping, a method that retrieves content and data from the internet. The collected data is typically stored in a local file for further processing, manipulation, and analysis as required. One of the reasons behind using web scraping for social media data collection is sentiment analysis. By extracting data from social media, companies or specific parties can analyze user sentiment toward a brand, product, or service, providing valuable insights for decision- making and strategic planning. Thus, web scraping serves as an effective tool for gathering social media data, helping businesses obtain crucial insights for informed strategic decisions [20], [21], [22].

The data was collected using specific keywords representing potential symptoms of mental health disorders, particularly bipolar disorder. These keywords were selected through both observational insights and literature-based approaches, focusing on common linguistic expressions indicative of psychological distress. Examples of keywords used in the data collection process include “*capek hidup*” (tired of life), “*pengen mati*” (want to die), “self- harm”, and “*kesepian*” (loneliness). To retrieve relevant tweets, we employed TweetHarvest, an open-source Twitter scraping tool that interacts with the Twitter API v2. Prior to data collection, Twitter developer credentials (Bearer Token) were obtained to authenticate and authorize access to the Twitter API. The scraping process involved defining search queries based on the chosen keywords and setting language filters to Bahasa Indonesia. The tweets were collected using a programmatic query that retrieved recent tweets in real time, ensuring that each result was publicly available and within Twitter’s terms of service.

To maintain relevance and data freshness, the data collection period was limited to April 21–23, 2025. Only tweets containing at least one of the selected keywords were retained. After initial retrieval, the dataset underwent preprocessing to remove retweets, duplicate entries, advertisements, and tweets that lacked contextual relevance to mental health discourse. The final filtered dataset is summarized in the **Table 1**.

Table 1. Mental Health Risk Analysis Dataset

created_at	full_text	username
Mon Apr 21 19:20:15 +0000 2025	*emang ga cape apa ya*	_cningning
Mon Apr 21 19:17:49 +0000 2025	*Cape bgt sih yaa*	_cningning
…	...	…
Tue Apr 22 06:03:34 +0000 2025	*Bener emang pilih prioritas itu kesehatan mental*	zeokatt
Sat Apr 19 00:53:36 +0000 2025	*Jangan pernah jatuhkan mental seseorang asli*	zeokatt

Although keyword selection followed a structured approach, heuristic filtering may introduce bias by excluding subtle expressions of psychological distress or including irrelevant content due to ambiguous terms. Since data came from public tweets only, it may not represent users with private accounts or low activity. Linguistic nuances like sarcasm or slang can also be missed. A manual validation step was conducted to improve relevance, but this remains subjective. These limitations suggest the need for future research using human- in-the-loop validation, diverse data sources, and culturally aware language models for improved accuracy.

Pre-processing

After the data crawling process, the next step is data pre-processing and labeling. At this stage, feature engineering is conducted by assigning a bipolar risk level to each tweet based on keyword analysis in the cleaned text. The labeling process follows a rule-based classification approach, in which each text is categorized into one of three risk levels: low, medium, or high. This categorization is guided by a manually curated set of keywords derived from observations and literature, as follows:

- High risk (*tinggi*): Tweets containing severe psychological distress indicators, such as "self-harm", "*bunuh diri*" (suicide), or "*capek hidup*" (tired of life).
- Medium risk (*sedang*): Tweets referencing mental health concerns in a more general or clinical context, such as "mental health", "burnout", "*depresi*" (depression), "*gangguan jiwa*" (mental illness).
- Low risk (*rendah*): Tweets expressing common emotional states or everyday academic stressors, such as "*kuliah*" (college), "overthinking", "*stres kuliah*" (academic stress), or "*capek*" (tired).

The labeling process was automated using a Python function, which was applied to the entire cleaned dataset to generate risk labels

Table 2. Data Pre-processing Results

created_at	full_text	username	risk_label
Mon Apr 21 19:20:15 +0000 2025	*emang ga cape apa ya*	_cningning	*rendah*
Mon Apr 21 19:17:49 +0000 2025	*Cape bgt sih yaa*	_cningning	*rendah*
…	...	…	

created_at	full_text	username	risk_label
Tue Apr 22 06:03:34 +0000 2025	*Bener emang pilih prioritas itu kesehatan mental*	zeokatt	*sedang*
Sat Apr 19 00:53:36 +0000 2025	*Jangan pernah jatuhkan mental seseorang asli*	zeokatt	*sedang*

The resulting risk_label serves as an initial annotation for further analysis, such as model training, risk pattern identification, and temporal analysis using survival models. However, it is important to note that this heuristic labeling method introduces potential bias. Tweets with nuanced emotional expression may be mislabeled if they do not explicitly contain the predefined keywords. Additionally, sarcasm, slang, or cultural idioms may lead to false positives or negatives.

3. Results and Discussion

Results

Research results are the findings, conclusions, or outcomes obtained from a study or investigation.

CNN-LSTM

The Convolutional Neural Network (CNN) represents a foundational deep learning architecture particularly effective for processing structured grid-like data. Mathematically, the core convolution operation can be expressed as:

$$(f \times g)(t) = \sum_{T-\infty}^{\infty} f(\tau)g(t-\tau) \tag{1}$$

Where f represents the input feature map and g denotes the learnable kernel. This operation enables automatic feature extraction through hierarchical pattern recognition. overcoming the manual feature engineering limitations of traditional machine learning approaches [23]. The pooling operation, typically max-pooling, further enhances position invariance:

$$y_{i,j} \overset{max}{p, q \in R_{i,j}} x_{p,q} \tag{2}$$

For sequential data processing, we employ Long Short-Term Memory (LSTM) networks, which address the vanishing gradient problem through gated mechanisms. The key LSTM equations are:

$$\begin{gathered}
f_t = \sigma(Wf \cdot [h_{t-1}, x_t] + b_f) \\
f_t = \sigma(Wf \cdot [h_{t-1}, x_t] + b_f) \\
i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\
\tilde{c}\, t = tanh(W_i \cdot [h_{t-1}, x_t] + b_C) \\
C_t = f_t{}^\circ C_{t-1} + i_t{}^\circ \tilde{c}\mathrm{t} \\
O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\
h_t = O_t{}^\circ \tanh\,(C_t)
\end{gathered} \tag{3}$$

where f_t, i_t, and O_t represent forget, input, and output gates respectively, C_t is the cell state, and h_t denotes the hidden state. This architecture enables effective learning of long-range dependencies in sequential data [24]. After the bipolar risk labeling process, the text data is transformed into a numerical format using tokenization and padding. Tokenization is performed using the Keras Tokenizer (vocabulary size: 10,000), followed by padding to ensure all inputs have a uniform length (maximum of 50 tokens). The labels are encoded using LabelEncoder and converted into one-hot encoding for multi-class classification. The data is then split into training and test sets with stratification.

Table 3. CNN-LSTM Model Architecture

Layer (type)	Output Shape	Param #
Embedding	(None, 50, 128)	1,280,000
Conv1D	(None, 46, 64)	41,024
MaxPooling1D	(None, 23, 64)	0
LSTM	(None, 64)	33,024
Dropout	(None, 64)	0
Dense	(None, 64)	4,160
Dense_1	(None, 3)	195

Total params : 1,358,403 (5.18 MB)
Trainable params : 1,358,403 (5.18 MB)
Non-trainable params : 0 (0.00 B)

The CNN-LSTM model developed in this study consists of seven main layers with a total of 1,358,403 trainable parameters. The first layer is an Embedding layer that converts text into 128-dimensional vectors using a vocabulary size of 10,000, resulting in 1,280,000 parameters. Next, a Conv1D layer with 64 filters extracts local features and reduces the sequence length from 50 to 46, adding 41,024 parameters. The MaxPooling1D layer then downsamples the sequence from 46 to 23 without introducing new parameters. An LSTM layer with 64 units processes the sequence and produces the final output, contributing 33,024 parameters. A Dropout layer is used to prevent overfitting by randomly deactivating neurons during training. The following two Dense layers consist of 64 and 3 units, respectively, generating 4,160 and 195 parameters, with the final layer designed for 3-class bipolar risk classification

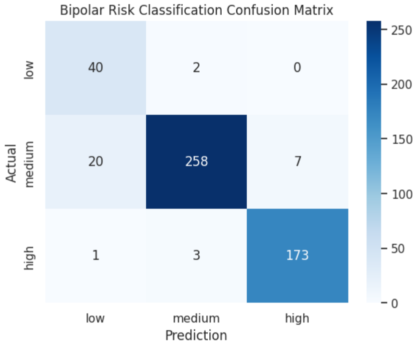

Figure 2. Confusion Matrix Bipolar Risk Classification

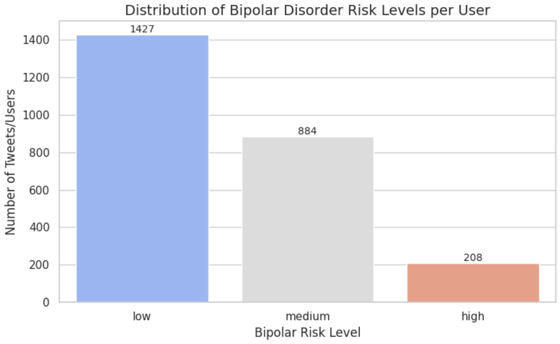

Figure 3. Distribution of Risk of Bipolar Mental Disorder for Each User

Based on the confusion matrix and the distribution of bipolar disorder risk among users, the classification model demonstrates reasonably good performance in identifying bipolar risk levels (low, medium, and high). Most data points were correctly classified, particularly in the low-risk class (273 data points), medium-risk (174 data points), and high-risk (37 data points) categories, with relatively low misclassification rates that primarily occurred between adjacent risk categories. Furthermore, the risk distribution chart reveals that the majority of users fall into the low-risk category (1,427 users), followed by medium-risk (884 users), with only a small proportion in the high-risk category (208 users). These findings indicate that while the prevalence of high bipolar risk is relatively low, the model maintains satisfactory accuracy in identifying this group, making it a potentially valuable tool for early detection of bipolar disorder through social media data analysis.

Cox Proportional Hazard (Cox PH)

The Cox Proportional Hazards (Cox PH) model was implemented to evaluate the influence of bipolar risk predictions generated by the CNN-LSTM model on the time- to-high-risk occurrence for each user. The Cox PH model is a semi-parametric regression model commonly used in survival analysis to estimate the hazard (risk) of an event occurring at time ttt given a set of covariates. The hazard function is expressed as:

$$\lambda(t \mid X_i) = \lambda_0(t) \cdot \boldsymbol{exp}(X_i \cdot \beta) \tag{4}$$

Where $\lambda(t \mid Xi)$ is the hazard function for individual $i, \lambda_0(t)$ is the baseline hazard function, X_i is the vector of covariates (in this case, bipolar risk predictions), and $\beta$is the vector of regression coefficients. This model allows for assessing how the predicted risk influences the timing until the high-risk event occurs without assuming a specific baseline hazard shape.

Several recent studies have successfully combined deep learning models such as CNN and LSTM with Cox PH for time-to-event analysis. Similarly, Kvamme et al. [25] introduced neural network architectures for survival prediction based on Cox models, enabling flexible modeling of complex covariate interactions. The estimation results from applying the Cox PH model on CNN-LSTM bipolar risk predictions can be seen in the **Table 4**.

Table 4. Variable Effects of Cox Proportional Hazard Model Results

	coef	Exp (coef)	Se (coef)	coef lower 95%	coef upper 95%	Exp (coef) lower 95%	Exp (coef) upper 95%	cmp to	z	p	-log2(p)
username_encoded	-0.00	1.00	0.00	-0.00	0.00	1.00	1.00	0.00	-1.88	0.06	4.04
Risk_label_encoded	1.68	5.39	0.07	1.55	1.81	4.73	6.14	0.00	25.36	<0.005	468.76

Table 5. Model Fit Metrics

Concordance	0.82
Partial AIC	6331.82
Log-likelihood ratio test	650.55 on 2 df
-log2(p) of II-ratio rest	469.27

Based on the estimation results, the variable risk_label_encoded indicates that higher predicted risk significantly increases the likelihood of developing high-risk bipolar disorder in a shorter time. Conversely, the variable username_encoded shows no significant effect on the time-to-event, suggesting that differences between users are not directly related to the acceleration of high-risk onset.

In the Cox Proportional Hazards model used in this study, the C-index value of 0.82 demonstrates that the model has reasonably good discriminative ability, with an 82% capability to distinguish individuals based on their risk of developing bipolar disorder. This means the model is effective in predicting the sequence of event times, thereby providing valid predictions about who is more likely to experience bipolar disorder earlier. The high C-index indicates that the CoxPH model successfully leverages the available data to estimate accurate time-to-event predictions based on risk factors, despite limitations in the duration distribution. With this adequate C-index value, the model proves to be a sufficiently reliable tool for risk analysis and mental health event prediction, particularly for bipolar disorder. However, to further improve accuracy, future research could explore more complex models or the inclusion of additional diverse features.

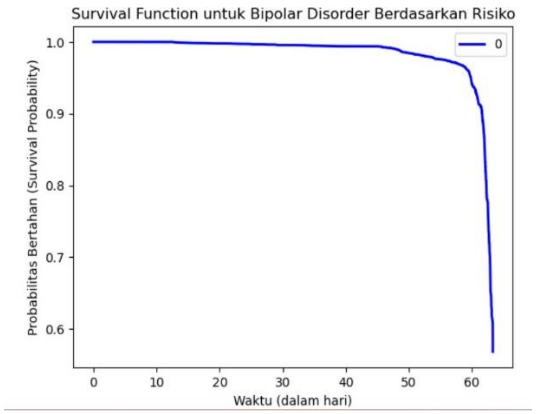

Figure 4. Survival Function for Bipolar Based on Risk

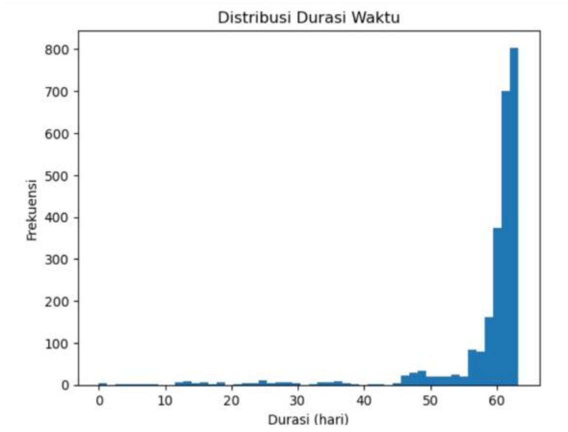

Figure 5. Distribution of Time Duration

Discussion

This study employs a hybrid CNN-LSTM and Cox Proportional Hazards (CoxPH) model to analyze bipolar disorder risk and estimate time-to-event occurrence. The CNN-LSTM model effectively extracts textual and sequential features from Twitter posts, while the CoxPH model enables survival analysis by estimating the time to high-risk states using behavioral and label-based covariates.

The Confusion Matrix **(Figure 2**) indicates that the classification model performs well in differentiating between low-, medium-, and high-risk categories. However, most misclassifications occurred between adjacent classes, such as medium being predicted as low or vice versa. This suggests a semantic overlap in the language used by users expressing moderate psychological distress and those experiencing milder symptoms. Additionally, expressions of distress may vary depending on linguistic style, cultural nuance, or sarcasm, which can lead to further inaccuracies. Future improvements could include the integration of sentiment intensity scores, transformer-based models, or contextual embeddings to reduce misclassification.

Figure 3, which shows the distribution of bipolar disorder risk among users, reveals that most users fall into the low-risk category. While this aligns with general population trends, it raises the challenge of class imbalance, which can affect model learning. Techniques such as resampling or class weighting may improve model robustness.

From a clinical perspective, the findings have strong practical implications. The model demonstrates the ability to distinguish users at high risk with reasonable accuracy, and the CoxPH results indicate that individuals with high-risk labels have a hazard ratio of 5.39, meaning they are over five times more likely to progress rapidly to a high-risk mental state. The survival function (**Figure 4**) and duration distribution (**Figure 5**) further illustrate the time dynamics associated with each risk group. These temporal patterns can be invaluable for mental health professionals aiming to prioritize early interventions and resource allocation.

Importantly, the model achieved a Concordance Index (C-index) of 0.82, signifying excellent predictive performance in estimating time-to-event. This level of accuracy indicates that the hybrid model is not only useful for detecting risk, but also for anticipating the timing of deterioration, which is crucial in mental health interventions where timing determines outcome efficacy.

Regarding generalizability, while the model is trained on Indonesian Twitter data, the framework can be extended to other languages and platforms (e.g., Facebook, Reddit) with proper adaptation of the keyword set and retraining using local linguistic features. However, the dependence on keyword-based initial labeling introduces limitations due to biases in language use and cultural interpretation. Broader application should be supported by manual annotation, diverse demographic sampling, and possibly cross-lingual model transfer to increase applicability across populations.

Overall, this study shows that the combination of CNN-LSTM for classification and CoxPH for survival estimation provides a powerful and interpretable tool for non-invasive mental health monitoring. These models, when integrated with digital behavioral data, can support real-time mental health surveillance and clinical decision-making, especially in resource- constrained settings where traditional psychiatric services are limited.

4. Conclusion

In summary, this study successfully demonstrates that the hybrid CNN-LSTM + Cox Proportional Hazards (CoxPH) model is a highly effective approach for analyzing bipolar disorder risk and predicting time-to-onset. The model provides valuable insights into how risk factors influence the timing of bipolar disorder manifestation, with high-risk individuals developing the disorder more rapidly. The CoxPH model achieved a C-index of 0.82, indicating strong discriminative ability in stratifying individuals by both risk level and time-to-onset.

These results prove that hybrid modeling combining CNN-LSTM with CoxPH can generate accurate predictions of bipolar disorder onset timing, showing significant potential for clinical applications in early detection. This study establishes a robust foundation for future research that could incorporate more diverse variables and advanced modeling techniques to further improve prediction accuracy.

In addition, while leveraging social media data enables scalable and non-invasive mental health monitoring, it also raises important ethical and privacy considerations. Ensuring user anonymity, securing data storage, and obtaining appropriate consent are essential to safeguard individual rights. Future implementations of such models must adhere to ethical guidelines and data protection regulations to balance technological advancement with responsible data use.

Overall, these findings open new possibilities for developing more efficient, reliable, and ethically grounded data-driven diagnostic systems for bipolar disorder.

References:

[1] N. Hartini, N. A. Fardana, A. D. Ariana, and N. D. Wardana, "Stigma toward people with mental health problems in Indonesia," Psychol Res Behav Manag, vol. Volume 11, pp. 535–541, Oct. 2018, doi: 10.2147/PRBM.S175251.

[2] The National Academies Press, Selected Health Conditions and Likelihood of Improvement with Treatment. Washington, D.C.: National Academies Press, 2020. doi: 10.17226/25662.

[3] V. Vasu and M. Indiramma, "A Survey on Bipolar Disorder Classification Methodologies using Machine Learning," in 2020 International Conference on Smart Electronics and Communication (ICOSEC), IEEE, Sep. 2020, pp. 335–340. doi: 10.1109/ICOSEC49089.2020.9215334.

[4] R. Safa, S. A. Edalatpanah, and A. Sorourkhah, "Predicting mental health using social media: A roadmap for future development," in Deep Learning in Personalized Healthcare and Decision Support, Elsevier, 2023, pp. 285–303. doi: 10.1016/B978-0- 443-19413-9.00014-X.

[5] A. B. Syahputri and Y. Sibaroni, "Comparative Analysis of CNN and LSTM Performance for Hate Speech Detection on Twitter," in 2023 11th International Conference on Information and Communication Technology (ICoICT), IEEE, Aug. 2023, pp. 190–195. doi: 10.1109/ICoICT58202.2023.10262656.

[6] M. Kang, S. Shin, J. Jung, and Y. T. Kim, "Classification of Mental Stress Using CNN- LSTM Algorithms with Electrocardiogram Signals," J Healthc Eng, vol. 2021, pp. 1– 11, Jun. 2021, doi: 10.1155/2021/9951905.

[7] J. L. Katzman, U. Shaham, A. Cloninger, J. Bates, T. Jiang, and Y. Kluger, "DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network," BMC Med Res Methodol, vol. 18, no. 1, p. 24, Dec. 2018, doi: 10.1186/s12874-018-0482-1.

[8] C.-C. Shen, A. C. Yang, J.-H. Hung, L.-Y. Hu, Y.-Y. Chiang, and S.-J. Tsai, "Risk of psychiatric disorders following pelvic inflammatory disease: a nationwide population- based retrospective cohort study," Journal of Psychosomatic Obstetrics & Gynecology, vol. 37, no. 1, pp. 6–11, Jan. 2016, doi: 10.3109/0167482X.2015.1124852.

[9] S. D. Ross, T. Lachmann, S. Jaarsveld, S. G. Riedel-Heller, and F. S. Rodriguez, “Creativity across the lifespan: changes with age and with dementia,” BMC Geriatr, vol. 23, no. 1, p. 160, Mar. 2023, doi: 10.1186/s12877-023-03825-1.

[10] D. Nickson, C. Meyer, L. Walasek, and C. Toro, “Prediction and diagnosis of depression using machine learning with electronic health records data: a systematic review,” BMC Med Inform Decis Mak, vol. 23, no. 1, p. 271, Nov. 2023, doi: 10.1186/s12911-023- 02341-x.

[11] M. Irigoyen et al., “Predictors of re-attempt in a cohort of suicide attempters: A survival analysis,” J Affect Disord, vol. 247, pp. 20–28, Mar. 2019, doi: 10.1016/j.jad.2018.12.050.

[12] L. Lac, C. K. Leung, and P. Hu, “Computational frameworks integrating deep learning and statistical models in mining multimodal omics data,” J Biomed Inform, vol. 152, p. 104629, Apr. 2024, doi: 10.1016/j.jbi.2024.104629.

[13] R. Zhang, J. Jiang, and X. Shi, “Modeling Long Sequences in Bladder Cancer Recurrence: A Comparative Evaluation of LSTM,Transformer,and Mamba,” Jul. 2024, [Online]. Available: http://arxiv.org/abs/2405.18518

[14] J. S. L. Figuerêdo, A. L. L. M. Maia, and R. T. Calumby, “Early depression detection in social media based on deep learning and underlying emotions,” Online Soc Netw Media, vol. 31, p. 100225, Sep. 2022, doi: 10.1016/j.osnem.2022.100225.

[15] Z. Wang et al., “Survival modeling using deep learning, machine learning and statistical methods: A comparative analysis for predicting mortality after hospital admission,” Mar. 2024.

[16] Y. Huang, J. Li, M. Li, and R. R. Aparasu, “Application of machine learning in predicting survival outcomes involving real-world data: a scoping review,” BMC Med Res Methodol, vol. 23, no. 1, p. 268, Nov. 2023, doi: 10.1186/s12874-023-02078-1.

[17] P. McInerney, A. Ajith, D. Pretorius, and S. Mabizela, “Considerations for choosing a research design,” Wits J Clin Med, vol. 6, no. 3, 2024, doi: 10.18772/26180197.2024.v6n3a11.

[18] K. Wisenthige, “Research Design,” 2023, pp. 74–93. doi: 10.4018/978-1-6684-6859- 3.ch006.

[19] S. P. Wasti, P. Simkhada, E. van Teijlingen, B. Sathian, and I. Banerjee, “The Growing Importance of Mixed-Methods Research in Health,” Nepal J Epidemiol, vol. 12, no. 1, pp. 1175–1178, Mar. 2022, doi: 10.3126/nje.v12i1.43633.

[20] Y. Qi and Z. Shabrina, “Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach,” Soc Netw Anal Min, vol. 13, no. 1, p. 31, Feb. 2023, doi: 10.1007/s13278-023-01030-x.

[21] N. M. Diaa, S. S. Ahmed, H. M. Salman, and W. A. Sajid, “Statistical Challenges in Social Media Data Analysis Sentiment Tracking and Beyond,” Journal of Ecohumanism, vol. 3, no. 5, pp. 365–384, Sep. 2024, doi: 10.62754/joe.v3i5.3912.

[22] J. R. Jim, M. A. R. Talukder, P. Malakar, M. M. Kabir, K. Nur, and M. F. Mridha, “Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the- art review,” Natural Language Processing Journal, vol. 6, p. 100059, Mar. 2024, doi: 10.1016/j.nlp.2024.100059.

[23] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, “Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach,” Procedia Comput Sci, vol. 132, pp. 679–688, 2018, doi: 10.1016/j.procs.2018.05.069.

[24] B. Ghojogh and A. Ghodsi, “Recurrent Neural Networks and Long Short-Term Memory Networks: Tutorial and Survey,” Apr. 2023.

[25] H. Kvamme, Ø. Borgan, and I. Scheel, “Time-to-Event Prediction with Neural Networksand Cox Regression,” Sep. 2019.