Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Analyzing Social Network Data for National Self-Harm Trend Predictions

Shaik Dehtaj[1], Ganta Jacob Victor[2]
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur, Andhra Pradesh, India, 522302

ARTICLE INFO	ABSTRACT

Received: 29 Dec 2024
Revised: 15 Feb 2025
Accepted: 24 Feb 2025

Introduction: Self-harm means the conscious demonstration of self-harming or self-injury, regardless of the basic inspiration or the level of self-destructive aim, which might prompt injury or casualty. Self-mischief and self destruction are huge issues, especially in unfortunate countries. A new report shows that around 77% of self destruction occasions happen in low-and center pay countries. This pattern is connected to the reception of specialized advancements and the quick movement of urbanization here.

Objectives: The goal of this research is to create and test a framework called FAST (Forecasting self-harm using Aggregated Social media patterns), which uses extensive social media data to forecast and nowcast patterns in self-harm at the national level. The research assesses the predictive ability of machine learning regressors to improve the accuracy of self-harm trend predictions, particularly in areas with little historical statistical data, by extracting psychological signals using language-agnostic language models and turning them into time series data.

Methods: This exploration presents FAST, a framework expected to foresee self-harm drifts broadly by utilizing mental signs got from broad online entertainment information. These signs go about as a proxy for real populace psychological well-being, which may be used to work on the consistency of self-harm designs. Language-skeptic calculations are at first prepared to remove different mental signs from amassed virtual entertainment interchanges. Accordingly, these signs are united and handled into multivariate time series, whereupon the time-delay implanting strategy is utilized to change over them into worldly inserted examples. At last, different ML regressors are surveyed for their prescient ability. We have explored different avenues regarding two of the most conspicuous ML calculations: the Decision Tree method and the Voting Regressor. In contrast with elective techniques, it yields a decreased MAE error.

Results: This study divides as Mean Absolute Error (MAE): Showed better overall accuracy with a lower average prediction error as compared to baseline models. The model's ability to handle significant deviations was confirmed by the Root Mean Squared Error (RMSE), which displayed a decreased squared error size. Mean Absolute Percentage Error (MAPE): The model's average improvement over the ARIMA baseline was 43.56% in predicting deaths from self-harm and 36.48% in predicting injuries.

Conclusions: This study presents FAST, a unique approach that uses psychological signals from social media to forecast national trends in self-harm. FAST predicted self-harm deaths and injuries in Thailand with high accuracy using 12 mental health variables from tweets. Based on MAPE, FAST increased forecast accuracy by 43.56% for deaths and 36.48% for injuries when compared to the conventional ARIMA model. Although additional enhancements utilizing models like as Decision Tree and Voting Regressor are recommended for future research, the results demonstrate FAST's potential for prompt public health intervention.

Keywords: *Self-harm, nowcasting, forecasting, online social networks, cross-lingual text classification.*

INTRODUCTION

The expansion in self-harm events prompts individual distress and misfortune, while likewise causing enduring adverse consequences on the economy, mostly through lessened long haul work efficiency. Observing and estimating

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

populace level self-harm examples might be critical for public policymakers and general wellbeing partners to actually survey conditions and take on procedures to relieve or turn away such expected calamities. For example, after discovering that particular severe strategies intended to battle cross country scourges have brought about psychological wellness challenges among the general population and are expected to essentially heighten self-harm rates, policymakers might examine fundamental acclimations to the current approaches liable for these issues. Also, the execution of general wellbeing drives, like portable psychiatry units or hotlines, could be used to address populaces confronting unfortunate results. Presently, the strategies used to assemble data on self-harm patterns at the

public level depends on authoritative reports from medical care offices and clinics cross country. This strategy requires significant monetary, human, and worldly assets, bringing about inconsistent and deferred information openness. Coarse-grained and postponed insights might have confined utility for proactive approach making. This examination presents FAST, a framework for foreseeing public self-harm patterns through mental signals got from web-based entertainment information. Language-freethinker models assess messages, creating multivariate time series altered through time-delay inserting. ML regressors, for example, Decision Tree and Voting, outperform ordinary techniques in foreseeing fatalities and wounds from self-harm, giving huge experiences to policymakers. The expansion in self-harm events, related with specialized progress and metropolitan development in emerging nations, presents extensive challenges for brief expectation and understanding at the public scale. Conventional methodologies reliant upon verifiable information might demonstrate lacking. This work presents FAST, which utilizes mental signs created from online entertainment to anticipate self-harm designs, so tending to the need for improved expectation models in general wellbeing.

OBJECTIVES

This study presents FAST, a unique approach that uses psychological signals from social media to forecast national trends in self-harm. FAST predicted self-harm deaths and injuries in Thailand with high accuracy using 12 mental health variables from tweets. Based on MAPE, FAST increased forecast accuracy by 43.56% for deaths and 36.48% for injuries when compared to the conventional ARIMA model. Although additional enhancements utilizing models like as Decision Tree and Voting Regressor are recommended for future research, the results demonstrate FAST's potential for prompt public health intervention.

The main goal is to show that social media communications may be used as reliable stand-ins for population mental health indicators when they are analysed using language-agnostic models. To forecast deaths and injuries from self-harm, these indicators are converted into multivariate time series using time-delay embedding. Machine learning regressors, like Decision Tree and Voting Regressor, are then used for analysis.

Twelve indicators pertaining to mental health that were taken from tweets are used in a case study from Thailand to test the framework. With performance gains of 43.56% and 36.48% in predicting self-harm deaths and injuries, respectively, based on MAPE, the study seeks to demonstrate that FAST considerably increases prediction accuracy over conventional forecasting techniques like ARIMA.

This study examines the potential of social media as a timely and scalable data source in light of the drawbacks of traditional surveillance systems, including delayed reporting, resource limits, and a lack of real-time insight. The system uses powerful machine learning methods, such as Decision Tree and Voting Regressor, to improve forecasting accuracy by analysing user-generated material using language-independent models and transforming mental health signals into structured time-series data. A case study in Thailand is used to assess the framework, and it performs noticeably better than the ARIMA benchmark model in terms of forecasting self-harm injuries and fatalities.

The ultimate goal of this project is to give stakeholders and public health officials a timely, scalable, and economical tool for predicting and reducing the effects of self-harm, particularly in low- and middle-income nations where traditional data gathering techniques might not be enough.

METHODS

3.1 Dataset Collection:

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

This study dataset is fundamental for analyzing the connection between psychological signs got from web-based entertainment information and the expectation of self-harm occurrences at the public scale. The examination fixates on Thailand, using a contextual investigation strategy that consolidates virtual entertainment information with checked insights on self-harm occurrences. This examination utilizes two separate datasets. The underlying dataset comprises of over 4.9 million tweets arbitrarily accumulated utilizing the Twitter Programming interface from October 2017 to January 2021. To ensure generalizability across numerous web-based entertainment stages, simply the timestamp and text based content of each tweet are protected. The language-skeptic strategy for distinguishing mental signs empowers the proposed anticipating structure to apply to multivariate time series information, advancing relevance across phonetic and geological settings.

The second dataset contains verifiable data on month to month events of fatalities and wounds coming about because of self-harm in Thailand. This dataset is gotten from the Department of Mental Health, Ministry of Public Health of Thailand, including irrefutable data for model validation. The dataset involves month to month tweet counts (displayed as a bar outline on the right Y-pivot) and reported occurrences of fatalities and wounds coming about because of self-harm (addressed as a line diagram on the left Y-hub). The dataset shows a huge expansion in death and injury cases from September 2019 to October 2019, reasonable because of changes in the detailing framework matching with the financial year shift in Thailand, which runs from October to September.

	date	MS-Pos	MS-Neg	MS-Amb	MS-Neu	ME-Ang	ME-Dis	ME-Fea	ME-Joy	ME-Sad	ME-Sur	ME-Neu	M-NST	M-ST	GH-Death	GH-Injure
0	2017-10-31	0.124349	0.217099	0.002639	0.655914	0.060558	0.001121	0.010423	0.180025	0.042351	0.098291	0.607230	0.734283	0.265717	56.0	224.0
1	2017-11-30	0.122213	0.199027	0.002266	0.676494	0.041806	0.001326	0.016026	0.182476	0.033406	0.104943	0.620016	0.753447	0.246553	45.0	239.0
2	2017-12-31	0.103728	0.244845	0.002444	0.648983	0.057183	0.001756	0.011395	0.179296	0.040314	0.113988	0.596068	0.743003	0.256997	60.0	255.0
3	2018-01-31	0.096537	0.269589	0.002332	0.631543	0.055182	0.001676	0.012206	0.152939	0.024959	0.116959	0.636079	0.746279	0.253721	88.0	336.0
4	2018-02-28	0.093888	0.288119	0.001998	0.615995	0.063627	0.001289	0.011892	0.163508	0.035321	0.118380	0.605983	0.728080	0.271920	60.0	299.0

3.2 Pre-processing:

During the data processing step, the examination uses Python apparatuses, including pandas and NumPy, to effectively oversee and alter the datasets. The datasets are at first coordinated into pandas data frames, offering a clear and proficient system for dealing with the plain information. NumPy is utilized for reshaping tasks, working with proficient information control and planning.

Unnecessary sections are killed from the datasets, ensuring that main appropriate data is safeguarded for model preparation. This stage works on computational proficiency and underlines perspectives basic for anticipating self-harm designs got from mental signs via web-based entertainment.

Normalization is executed on the preparation of information, a fundamental preprocessing move toward standardize mathematical qualities, guaranteeing they stick to a uniform scale. This normalization keeps explicit attributes from eclipsing others during model training, subsequently upgrading assembly and execution.

Resulting to preprocessing, the accompanying stage involves the extraction of training elements and names from the dataset. Highlights indicate the information factors utilized by the model for prediction, though marks mean the objective variable, explicitly the count of revealed self-harm cases. This division ensures that the model obtains designs from the qualities to give exact conjectures in regards to the objective variable, upgrading the general viability of the forecasting system.

3.3 Training & Testing:

The dataset is partitioned into training and testing sets to evaluate the model's presentation on new information, a fundamental stage in approving the speculation of the determining model. This division empowers the model to get designs from the training information and hence assess its anticipated presentation on free, concealed test information.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The training set, as a rule comprising a critical section of the dataset, is the establishment for training the determining models. The calculations examine authentic examples and linkages inside this information test, permitting them to distinguish the fundamental patterns and elements related with self-harm cases.

The testing set, alternately, stays isolated during the training stage and is assigned for assessing the model's accuracy and execution. The model's ability to foresee obscure information is an essential mark of its generalizability and viability in expecting self-harm designs. This evaluation ensures that the model doesn't just hold the preparation information however understands the principal designs relevant to novel, unencountered events.

The division of training and testing sets is ordinarily executed haphazardly to ensure delegate information dispersion, and the results on the testing set yield huge bits of knowledge into the model's expectation viability and conceivable genuine pertinence.

3.4 Algorithms:

"ARIMA (Autoregressive Integrated Moving Average)" is a period series estimating framework that models fleeting relationship inside information. It incorporates autoregression, differencing, and moving midpoints. This study uses ARIMA to recognize fleeting examples in self-harm propensities, profiting by its ability to oversee time-subordinate variances for exact estimating.

"Bayesian Ridge" is a probabilistic regression method that utilizes Bayesian standards for regularization. This study utilizes a probabilistic methodology for expecting self-harm occurrences, tending to vulnerability and improving model strength against dataset commotion.

"Support Vector Regression (SVR)" is a regression approach got from help vector machines. This trial use SVR to demonstrate the non-straight relationships between psychological signals and examples of self-harm. Its flexibility in distinguishing convoluted designs renders it suitable for anticipating in circumstances with complex fundamental elements.

"XGBoost (eXtreme Gradient Boosting)" is an ensemble learning strategy that exhibits prevalent execution in regression issues. This task pick "XGBoost" for its capability in overseeing perplexing associations, highlight communications, and anomaly recognizable proof. The expanded tree structure works on forecast accuracy by coordinating numerous powerless models.

"Random Forest" is an ensemble learning technique that builds a few decision trees and totals their expectations. This study utilizes "Random Forest" for its ability to oversee non-straight associations, direct element importance examination, and keep away from overfitting, consequently upgrading the accuracy of self-harm expectations.

"CatBoost is a gradient boosting" strategy explicitly created to oblige classification highlights. This study utilizes "CatBoost" for its adequacy in overseeing downright information from online entertainment, consequently empowering the proper usage of data appropriate to mental signs for exact anticipating of self-harm episodes.

"Decision Tree" is a clear yet proficient methodology for regression errands. This study uses "Decision Trees" to show the dynamic interaction related with self-harm designs. Their interpretability and ability to recognize non-direct examples render them fundamental components of the forecasting system.

RESULTS

MAE:

"Mean Absolute Error (MAE)" evaluates the typical size of blunders in a bunch of forecasts, ignoring their directional predisposition. It is measured as the mean outright deviation between the anticipated and genuine qualities, effectively assessing the viability of a regression model.

"The MAE loss function formula:"

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\widehat{y_i} - y_i)^2$$

Where:
$\widehat{y_i}$ = Predicted value for the ith data point
y_i= Actual value for the ith data point
n = number of observations

RMSE:

"Root Mean Squared Error (RMSE) is now and again alluded to as Root Mean Squared Deviation (RMSD)". The "Root Mean Squared Error (RMSE)" is gotten from the "Mean Squared Error (MSE)" and fills in as a traditional measurement for evaluating the accuracy of a model in anticipating quantitative information. It is the square base of the "Mean Squared Error (MSE)" and measures the error's size in similar units as the result variable.

Still up in the air by taking the square base of the normal of the squared errors between the anticipated qualities and the genuine qualities, as communicated in the accompanying equation:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widehat{y_i} - y_i)^2}$$

Where:
$\widehat{y_i}$ = Predicted value for the ith data point
y_i= Actual value for the ith data point
n = number of observations

MAPE:

The "Mean Absolute Percentage Error (MAPE)" is a measurement that evaluates the exactness of an estimating model communicated as a rate. It indicates the mean outright rate deviation among expected and real qualities across all perceptions. MAPE is determined by averaging the absolute percentage errors between projected and actual values.

The equation for MAPE is as follows:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\widehat{y_i} - y_i}{y_i}\right| \times 100$$

Where:
$\widehat{y_i}$ = Predicted value for the ith data point
y_i= Actual value for the ith data point
n = number of observations

	Prediction Type	Algorithm Name	MAE	RMSE	MAPE
0	Death	ARIMA	289.312052	331.195047	109690.159107
1	Death	Bayesian Ridge	167.404834	222.865473	49669.019022
2	Death	Linear SVR	234.143838	270.735772	73297.857991
3	Death	XGBoost	128.403181	191.146958	36537.159582
4	Death	Random Forest	154.500000	230.716697	53230.194444
5	Death	Cat Boost	236.175301	268.920308	72318.131919
6	Death	Extension Decision Tree	14.555556	43.666667	1906.777778

"**Fig 1** Performance evaluation table for death prediction"

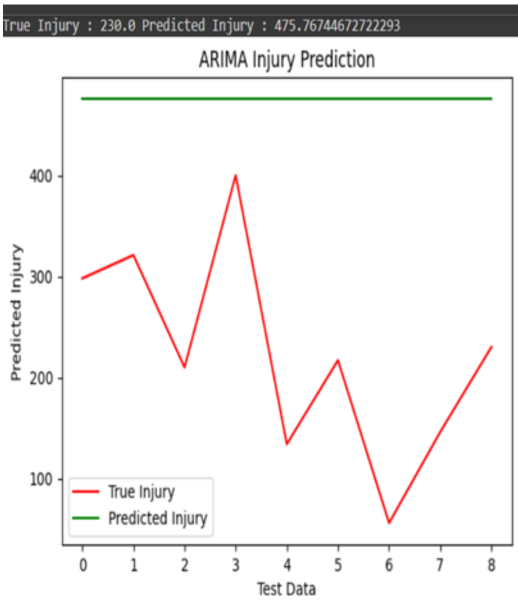
Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

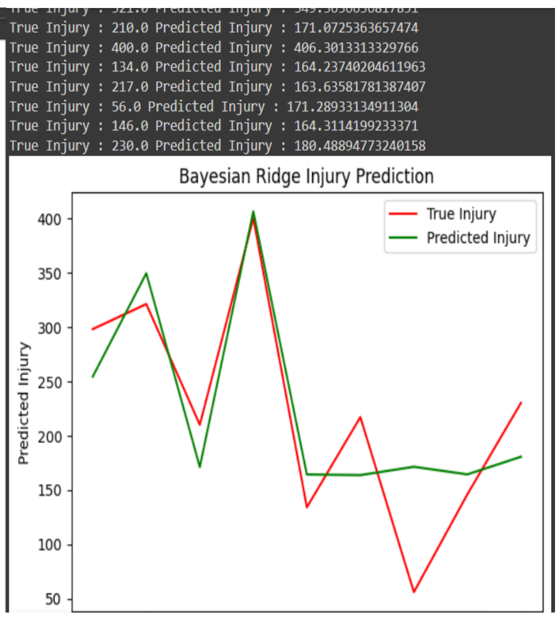
Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

	Prediction Type	Algorithm Name	MAE	RMSE	MAPE
0	Injury	ARIMA	145.395908	176.653211	31206.357057
1	Injury	Bayesian Ridge	50.849129	58.819382	3459.719713
2	Injury	Linear SVR	128.338791	137.697868	18960.702961
3	Injury	XGBoost	27.066800	30.373256	922.534677
4	Injury	Random Forest	41.777778	51.732753	2676.277778
5	Injury	Cat Boost	116.256856	118.311589	13997.632111
6	Injury	Extension Decision Tree	3.333333	8.246211	68.000000

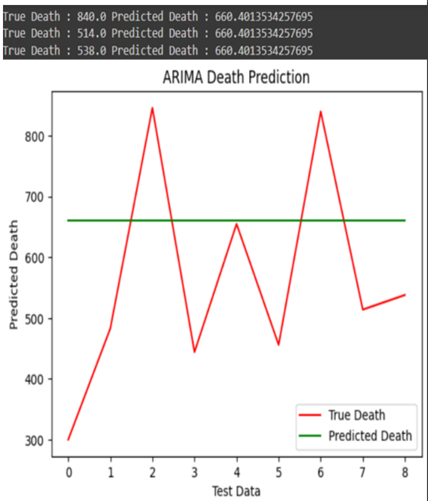
“**Fig 2** Performance evaluation table for injury prediction”

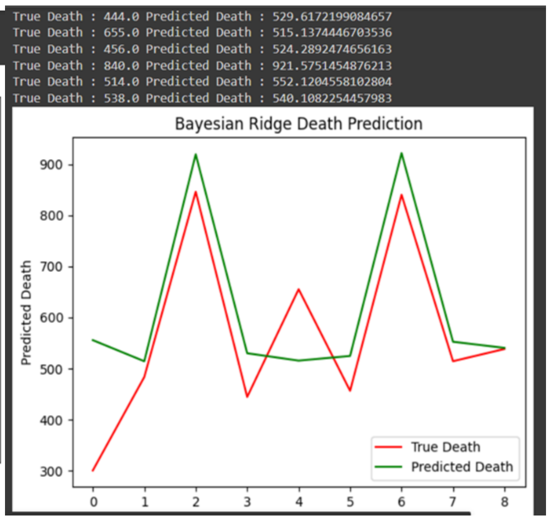




“**Fig 3** ARIMA injury prediction graph”

“**Fig 5** Bayesian Ridge injury prediction graph”





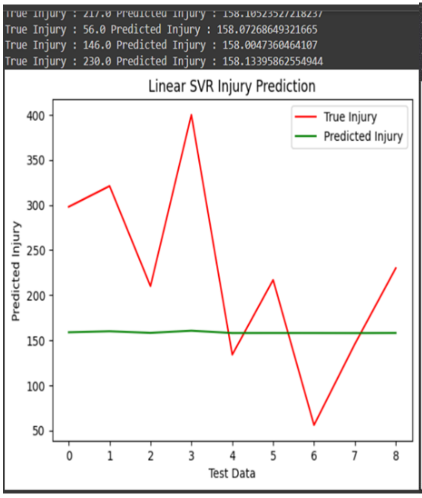
“**Fig 4** ARIMA death prediction graph”

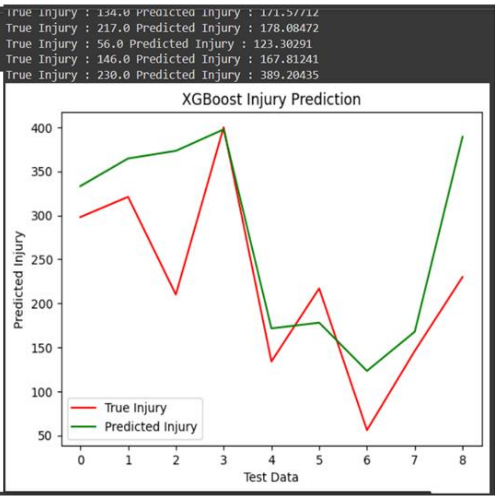
“**Fig 6** Bayesian ridge death prediction”

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

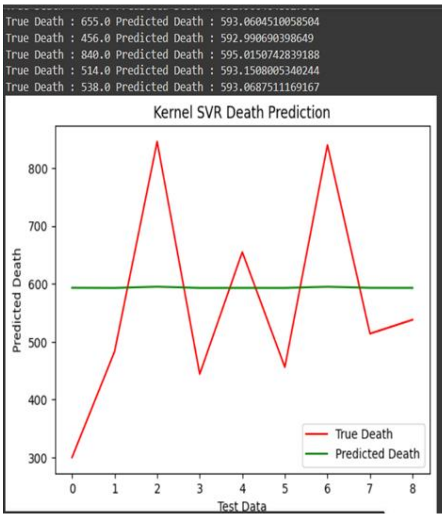
Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

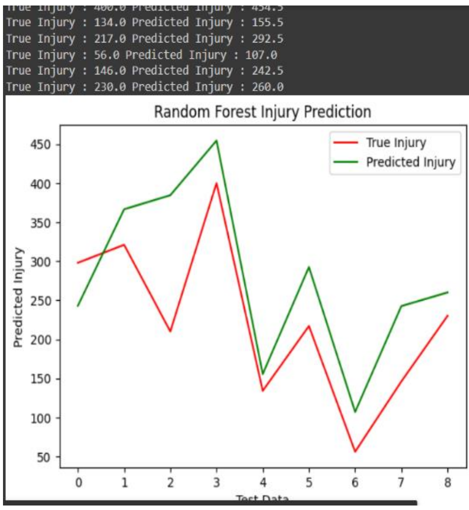


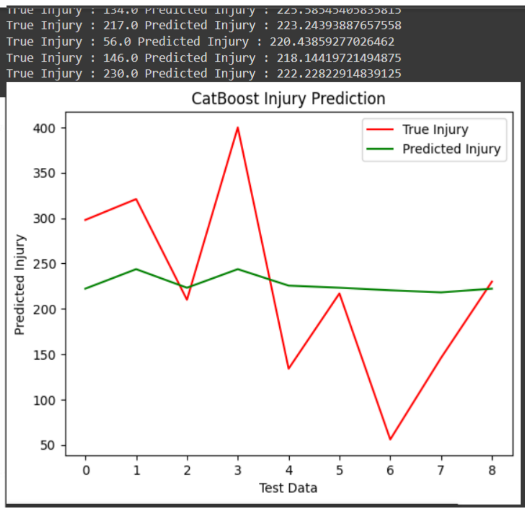


"**Fig 7** Linear SVR injury prediction graph" "**Fig 9** XGBoost injury prediction graph"



"**Fig 8** Linear SVR death prediction graph" "**Fig 10** XGBoost death prediction graph"



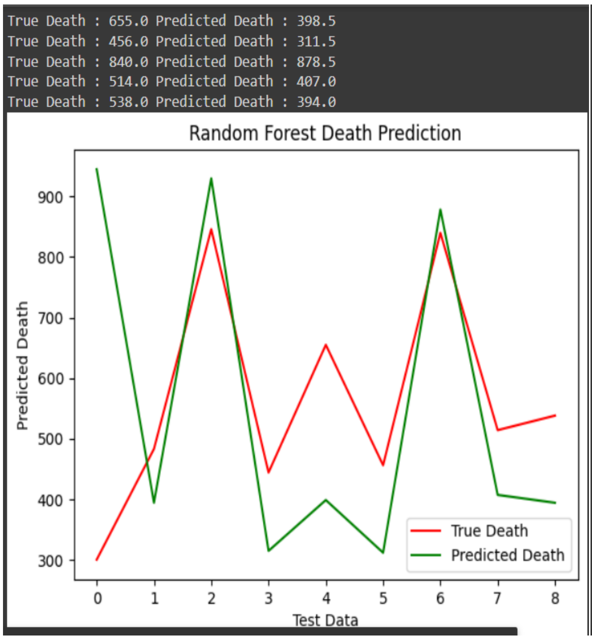


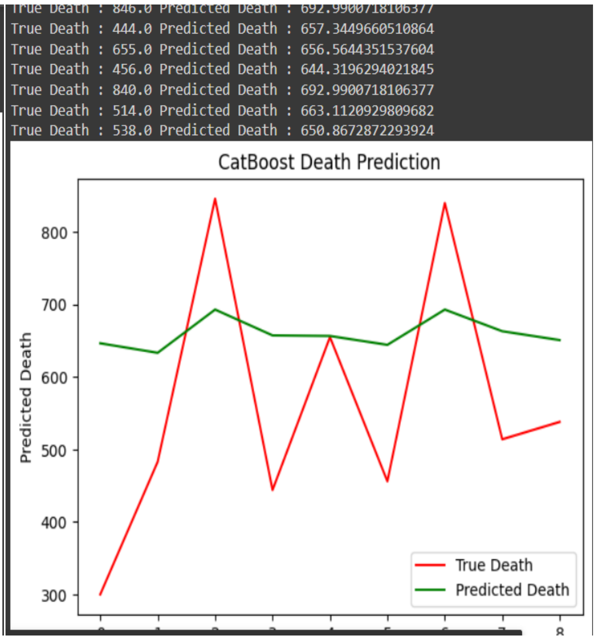
"**Fig 11** Random Forest injury prediction graph" "**Fig 13** CatBoost injury prediction graph"

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

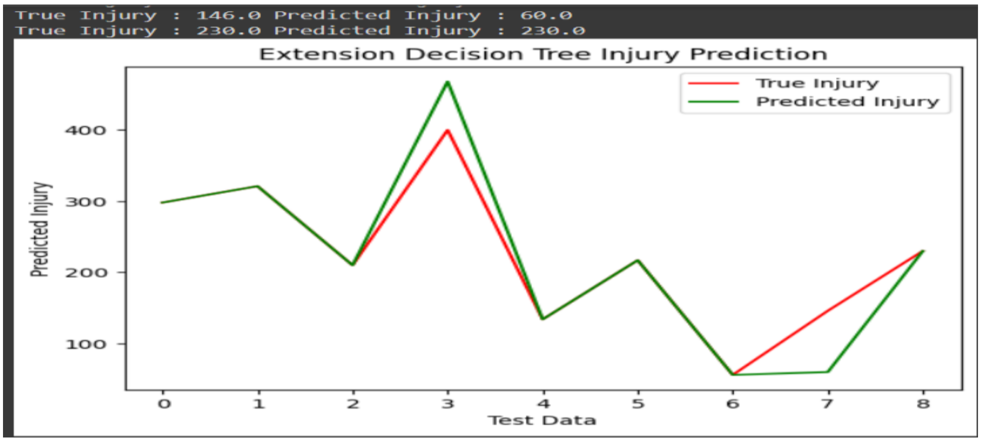
Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article





"**Fig 12** Random Forest death prediction graph" "**Fig 14** CatBoost death prediction graph"



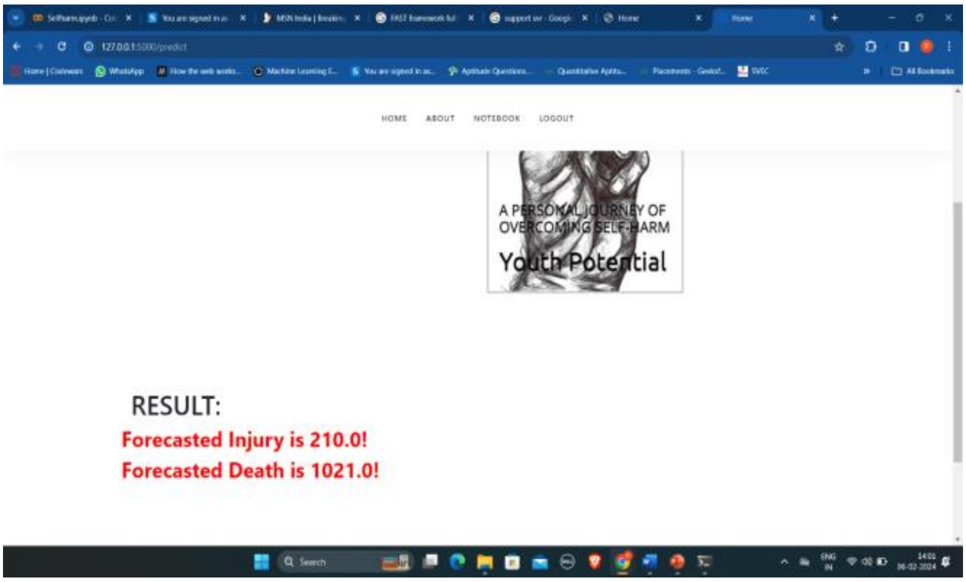
"**Fig 15** Extension Decision Tree injury prediction graph"

"**Fig 16** Extension Decision Tree death prediction graph"

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article



"**Fig 17** Predict result for forecasted injury is 210.0 and forecasted death is 1021.0"

DISCUSSION

Future review could expand its scope by exploring various internet based media sources, for example, news stories, a few online entertainment stages, and sight and sound data like recordings or pictures, to increase the dataset and further develop forecasting precision. Besides, the utilization of DL approaches might uncover more mind boggling designs inside the information. Looking at self-harm episodes at a point by point level, for example, territorial or segment explicit scales, offers potential for modifying the board procedures, working with the making of designated mediations fit to specific regions and socioeconomics, hence improving accuracy in preventive measures.

REFRENCES

[1] (May 2023). Suicide Rates for Girls Are Rising. Are Smartphones to Blame? [Online]. Available: https://www.economist.com/graphic-detail/ 2023/05/03/suicide-rates-for-girls-are-rising-are-smartphones-to-blame

[2] J. Adam-Troian and T. Arciszewski, "Absolutist words from search volume data predict state-level suicide rates in the United States," Clin. Psychol. Sci., vol. 8, no. 4, pp. 788–793, Jul. 2020.

[3] A. E. Aiello, A. Renson, and P. Zivich, "Social media-and Internet-based disease surveillance for public health," Annu. Rev. Public Health, vol. 41, p. 101, Apr. 2020.

[4] M. Akyuz and C. Karul, "The effect of economic factors on suicide: An analysis of a developing country," Int. J. Hum. Rights Healthcare, Jul. 2022.

[5] S. Z. Alavijeh, F. Zarrinkalam, Z. Noorian, A. Mehrpour, and K. Etminani, "What users' musical preference on Twitter reveals about psychological disorders," Inf. Process. Manage., vol. 60, no. 3, May 2023, Art. no. 103269.

[6] A. Aldayel and W. Magdy, "Stance detection on social media: State of the art and trends," Inf. Process. Manage., vol. 58, no. 4, Jul. 2021, Art. no. 102597.

[7] P. Angelov and A. Sperduti, "Challenges in deep learning," in Proc. 24th Eur. Symp. Artif. Neural Netw. (ESANN), 2016, pp. 489–496.

[8] A. Apisarnthanarak, P. Apisarnthanarak, C. Siripraparat, P. Saengaram, N. Leeprechanon, and D. J. Weber, "Impact of anxiety and fear for COVID-19 toward infection control practices among Thai healthcare workers," Infection Control Hospital Epidemiol., vol. 41, no. 9, pp. 1093–1094, Sep. 2020.

[9] S. Arunpongpaisal, S. Assanagkornchai, V. Chongsuvivatwong, and N. Jampathong, "Time-series analysis of trends in the incidence rates of successful and attempted suicides in Thailand in 2013–2019 and their predictors," BMC Psychiatry, vol. 22, no. 1, pp. 1–11, Aug. 2022.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[10] J. M. Barros, R. Melia, K. Francis, J. Bogue, M. O'Sullivan, K. Young, R. A. Bernert, D. Rebholz-Schuhmann, and J. Duggan, "The validity of Google trends search volumes for behavioral forecasting of national suicide rates in Ireland," Int. J. Environ. Res. Public Health, vol. 16, no. 17, p. 3201, Sep. 2019.
[11] B. E. Belsher, D. J. Smolenski, L. D. Pruitt, N. E. Bush, E. H. Beech, D. E. Workman, R. L. Morgan, D. P. Evatt, J. Tucker, and N. A. Skopp, "Prediction models for suicide attempts and deaths: A systematic review and simulation," JAMA Psychiatry, vol. 76, no. 6, pp. 642–651, 2019.
[12] L. Braghieri, R. Levy, and A. Makarin, "Social media and mental health," Amer. Econ. Rev., vol. 112, no. 11, pp. 3660–3693, 2022.
[13] R. G. Brereton and G. R. Lloyd, "Support vector machines for classification and regression," Analyst, vol. 135, no. 2, pp. 230–267, 2010.
[14] J. A. Bridge, M. Olfson, J. M. Caterino, S. W. Cullen, A. Diana, M. Frankel, and S. C. Marcus, "Emergency department management of deliberate selfharm: A national survey," JAMA Psychiatry, vol. 76, no. 6, pp. 652–654, 2019.
[15] L. Cao, H. Zhang, and L. Feng, "Building and using personal knowledge graph to improve suicidal ideation detection on social media," IEEE Trans. Multimedia, vol. 24, pp. 87–102, 2022.
[16] M. K. Y. Chan, H. Bhatti, N. Meader, S. Stockton, J. Evans, R. C. O'Connor, N. Kapur, and T. Kendall, "Predicting suicide following self-harm: Systematic review of risk factors and risk scales," Brit. J. Psychiatry, vol. 209, no. 4, pp. 277–283, Oct. 2016.
[17] Y. S. Chang and J. Lee, "Is forecasting future suicide rate possible?— Application of experience curve," Eng. Manag. Res., vol. 1, no. 1, p. 10, 2012.
[18] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, and H. Cho, "XGBoost: Extreme gradient boosting," R Package Version 0.4-2, Microsoft, Tech. Rep., 2015, pp. 1–4, vol. 1, no. 4.
[19] R. Cholakov and T. Kolev, "The GatedTabTransformer. An enhanced deep learning architecture for tabular modeling," 2022, arXiv:2201.00199.
[20] D. Demszky, D. Movshovitz-Attias, J. Ko, A. Cowen, G. Nemade, and S. Ravi, "GoEmotions: A dataset of fine-grained emotions," in Proc. 58th Annu. Meeting Assoc. Comput. Linguistics, Jul. 2020, pp. 4040–4054. [Online]. Available: https://aclanthology.org/2020.aclmain.372, doi: 10.18653/v1/2020.acl-main.372.
[21] J. B. Edgcomb, T. Shaddox, G. Hellemann, and J. O. Brooks, "Predicting suicidal behavior and self-harm after general hospitalization of adults with serious mental illness," J. Psychiatric Res., vol. 136, pp. 515–521, Apr. 2021.
[22] L. Favril, R. Yu, K. Hawton, and S. Fazel, "Risk factors for self-harm in prison: A systematic review and meta-analysis," Lancet Psychiatry, vol. 7, no. 8, pp. 682–691, Aug. 2020.
[23] F. Feng, Y. Yang, D. Cer, N. Arivazhagan, and W. Wang, "Languageagnostic BERT sentence embedding," 2020, arXiv:2007.01852.
[24] H. Fliege, J.-R. Lee, A. Grimm, and B. F. Klapp, "Risk factors and correlates of deliberate self-harm behavior: A systematic review," J. Psychosomatic Res., vol. 66, no. 6, pp. 477–493, Jun. 2009.
[25] A. Franzén, "Big data, big problems: Why scientists should refrain from using Google trends," Acta Sociologica, pp. 1–5, Jan. 2023.
[26] A. Froidevaux, J. Macalos, I. Khalfoun, M. Deffrasnes, S. d'Orsetti, N. Salez, and A. Sciberras, "Leveraging alternative data sources for socio-economic nowcasting," in Proc. Conf. Inf. Technol. Social Good, Sep. 2022, pp. 345–352.
[27] R. Gao, O. Duru, and K. F. Yuen, "High-dimensional lag structure optimization of fuzzy time series," Expert Syst. Appl., vol. 173, Jul. 2021, Art. no. 114698.
[28] M. George, "The importance of social media content for Teens' risks for self-harm," J. Adolescent Health, vol. 65, no. 1, pp. 9–10, Jul. 2019.
[29] N. A. Ghani, S. Hamid, I. A. T. Hashem, and E. Ahmed, "Social media big data analytics: A survey," Comput. Hum. Behav., vol. 101, pp. 417–428, Dec. 2018.
[30] A. Glazkova, M. Kadantsev, and M. Glazkov, "Fine-tuning of pre-trained transformers for hate, offensive, and profane content detection in English and Marathi," 2021, arXiv:2110.12687.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

[31] C. R. Glenn, E. M. Kleiman, J. Kellerman, O. Pollak, C. B. Cha, E. C. Esposito, A. C. Porter, P. A. Wyman, and A. E. Boatman, "Annual research review: A meta-analytic review of worldwide suicide rates in adolescents," J. Child Psychol. Psychiatry, vol. 61, no. 3, pp. 294–308, Mar. 2020.

[32] S. D. Gollapalli, G. A. Zagatti, and S.-K. Ng, "Suicide risk prediction by tracking self-harm aspects in tweets: NUS-IDS at the CLPsych 2021 shared task," in Proc. 7th Workshop Comput. Linguistics Clin. Psychol., Improving Access, 2021, pp. 93–98.

[33] K. L. Gratz, "Risk factors for and functions of deliberate self-harm: An empirical and conceptual review," Clin. Psychol., Sci. Pract., vol. 10, no. 2, pp. 192–205, 2003.

[34] [34] K. Hawton, D. Zahl, and R. Weatherall, "Suicide following deliberate self-harm: Long-term follow-up of patients who presented to a general hospital," Brit. J. Psychiatry, vol. 182, no. 6, pp. 537–542, Jun. 2003.

[35] K. Hawton, K. E. Saunders, and R. C. O'Connor, "Self-harm and suicide in adolescents," Lancet, vol. 379, no. 9834, pp. 2373–2382, 2012.

[36] E. Iob, A. Steptoe, and D. Fancourt, "Abuse, self-harm and suicidal ideation in the U.K. During the COVID-19 pandemic," Brit. J. Psychiatry, vol. 217, no. 4, pp. 543–546, Oct. 2020.

[37] M. I. Islam, R. Khanam, and E. Kabir, "Depression and anxiety have a larger impact on bullied girls than on boys to experience self-harm and suicidality: A mediation analysis," J. Affect. Disorders, vol. 297, pp. 250–258, Jan. 2022.

[38] S.-P. Jun, H. S. Yoo, and S. Choi, "Ten years of research change using Google trends: From the perspective of big data utilizations and applications," Technol. Forecasting Social Change, vol. 130, pp. 69–87, May 2018.

[39] S. Kandula, M. Olfson, M. S. Gould, K. M. Keyes, and J. Shaman, "Hindcasts and forecasts of suicide mortality in U.S.: A modeling study," PLOS Comput. Biol., vol. 19, no. 3, Mar. 2023, Art. no. e1010945.

[40] R. C. Kessler, R. M. Bossarte, A. Luedtke, A. M. Zaslavsky, and J. R. Zubizarreta, "Suicide prediction models: A critical review of recent research with recommendations for the way forward," Mol. Psychiatry, vol. 25, no. 1, pp. 168–179, Jan. 2020.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.