Data-Driven Prediction of Self-Harm Trends Using Social Media Signals and Machine Learning Techniques

Y. Shakuntala[1], *Mohammed Zubair Ahmed*[2], Syed Raheem Ahmed[3], *Syed Eeshan*[4], Mudassir Shareef[5]

[1]Assistant Professor, Department of CSE (Data Science), Lords Institute of Engineering and Technology, Hyderabad, Telangana, India.
[2,3,4,5] UG Students, Department of CSE (Data Science), Lords Institute of Engineering and Technology, Hyderabad, Telangana, India.

***Abstract*— This study focuses on forecasting national-level self-harm trends by analyzing emotional patterns extracted from social media data. With the increasing influence of digital communication, individuals often express their mental states through online platforms, making them a valuable source for understanding psychological well-being. The proposed approach integrates sentiment analysis from user-generated content with official health statistics to build a predictive dataset. Various machine learning models, including ARIMA, Support Vector Machine, Bayesian Ridge, XGBoost, Random Forest, and Decision Tree, are applied to estimate injury and death trends related to self-harm. The results indicate that ensemble methods such as XGBoost and Decision Tree provide more accurate predictions with lower error rates. This framework can assist authorities in identifying at-risk populations and implementing timely intervention strategies, ultimately contributing to improved mental health awareness and prevention efforts.**

***Keywords*— Self-Harm Prediction, Social Media Analysis, Sentiment Analysis, Machine Learning, XGBoost, Decision Tree, Time Series Forecasting, Mental Health Analytics**

I. INTRODUCTION

Mental health has become a growing concern worldwide, with depression emerging as a significant factor influencing individual well-being and societal stability [12]. Rapid changes in lifestyle, increased competition in education and careers, and social pressures have contributed to emotional distress among people of different age groups [4]. In recent years, self-harm and suicide rates have shown an alarming increase, making it necessary to understand the underlying causes and patterns [1], [9]. Traditional methods of studying mental health trends often rely on surveys and clinical reports, which may not capture real-time emotional changes [11]. With the widespread use of social media platforms, individuals now frequently express their feelings, thoughts, and struggles online [3]. These digital expressions provide valuable insights into their mental state. By analyzing such data, researchers can gain a deeper understanding of emotional patterns and identify early warning signs of self-harm behavior at a larger scale [10].

Social networking platforms have transformed the way people communicate and share personal experiences [3]. Users often post messages that reflect their emotional conditions, including stress, anxiety, sadness, or hopelessness [5]. These platforms generate a vast amount of unstructured textual data, which can be analyzed using natural language processing and sentiment analysis techniques [6]. By converting text into measurable emotional indicators, it becomes possible to study trends in mental health across populations. Integrating this information with official health statistics, such as injury and mortality records, enables the creation of a comprehensive dataset for predictive analysis [9]. This combined approach helps bridge the gap between digital behavioral signals and real-world outcomes. As a result, researchers can develop models that not only identify patterns but also forecast future risks [2]. Such predictive capabilities are crucial for designing preventive measures and supporting mental health interventions effectively.

Machine learning techniques play a vital role in analyzing complex datasets and uncovering hidden relationships within data [7]. In the context of self-harm prediction, various algorithms can be employed to model patterns between emotional signals and recorded incidents of injury or death. Time series models like ARIMA are useful for capturing temporal trends, while algorithms such as Support Vector Machine, Bayesian Ridge, Random Forest, and XGBoost are effective in handling non-linear relationships and large datasets [13]. Each model has its strengths and limitations, making it important to evaluate their performance using standard metrics such as Mean Absolute Error, Root Mean Square Error, and Mean Absolute Percentage Error. By comparing these models, researchers can determine the most suitable approach for accurate forecasting. The use of multiple algorithms ensures robustness in analysis and provides a deeper understanding of how different techniques respond to the same dataset.

Accurate prediction of self-harm trends can significantly contribute to public health planning and preventive strategies [11]. Governments and healthcare organizations can use such predictions to identify high-risk groups and allocate resources more efficiently [14]. Early detection of mental health issues allows for timely counseling, awareness programs, and community support initiatives. In addition, predictive models can help monitor changes in emotional patterns over time, enabling authorities to respond proactively to emerging risks. The integration of technology and healthcare in this manner promotes a data-driven approach to mental health management. It also highlights the importance of interdisciplinary research combining data science, psychology, and public health [15]. By leveraging these advanced analytical techniques, societies can work towards reducing self-harm incidents and improving overall mental well-being in a structured and effective manner.

II. RELATED WORK

Adam-Troian et al., [2020] [2] Adam-Troian and Arciszewski explored the relationship between language patterns in online search data and suicide rates across different states in the United States. Their study focused on identifying “absolutist words,” which are often associated with rigid thinking and emotional distress. By analyzing search volume data, they demonstrated that such linguistic patterns could serve as indicators of mental health trends. The research highlights how digital behavioral data can be used to predict real-world outcomes. The findings suggest that online search activity reflects collective psychological states. This approach provides a cost-effective and scalable method for monitoring suicide risk. It also emphasizes the importance of language analysis in mental health research. The study contributes to early detection strategies by linking online expressions with offline consequences.

Aiello et al., [2020] [3] Aiello and colleagues examined the use of social media and internet-based platforms for disease surveillance in public health. Their work highlights how digital data sources can complement traditional monitoring systems. By analyzing user-generated content, researchers can track emerging health trends in real time. The study emphasizes the advantages of accessibility and timeliness in online data collection. It also discusses challenges such as data reliability and privacy concerns. The authors suggest that integrating multiple data sources improves accuracy. Their findings show that social media can act as an early warning system for health-related issues. This research supports the use of big data analytics in healthcare decision-making. It provides a foundation for using online platforms in mental health prediction.

Arunpongpaisal et al., [2022] [9] Arunpongpaisal and co-authors conducted a time-series analysis of suicide trends in Thailand between 2013 and 2019. The study examined both successful and attempted suicide cases along with influencing factors. Their findings revealed noticeable variations in suicide rates over time. The authors identified key predictors such as demographic and social variables. The use of time-series methods helped capture seasonal and long-term trends effectively. The study demonstrates the importance of analyzing historical data for understanding future risks. It also highlights the role of regional factors in shaping suicide patterns. The results provide valuable insights for policymakers and healthcare professionals. This work supports the development of predictive models for suicide prevention.

Belsher et al., [2019] [11] Belsher and colleagues presented a comprehensive review of predictive models for suicide attempts and deaths. Their research evaluated various statistical and machine learning approaches used in previous studies. The authors found that many models showed limited accuracy when applied in real-world settings. They emphasized the need for improving data quality and model validation techniques. The study also discussed challenges such as overfitting and lack of generalization. Despite limitations, predictive models were found useful for identifying high-risk individuals. The authors suggested combining multiple data sources to enhance performance. Their work highlights the complexity of suicide prediction. It provides guidance for future research in developing reliable prediction systems.

Chen et al., [2015] [18] Chen and co-authors introduced XGBoost, a powerful gradient boosting algorithm designed for high performance and efficiency. The model is known for its ability to handle large-scale datasets and complex relationships. It incorporates regularization techniques to prevent overfitting and improve accuracy. The study demonstrated its effectiveness in various machine learning tasks. XGBoost supports parallel processing, making it faster than many traditional algorithms. Its flexibility allows it to work well with structured and tabular data. The algorithm has been widely adopted in predictive analytics applications. The authors showed that it consistently outperforms other models in many scenarios. This makes it highly suitable for forecasting problems, including self-harm trend prediction.

III. DATASET DETAILS

The dataset used in this study combines emotional signals derived from social media activity with recorded statistics related to self-harm outcomes. It contains multiple attributes representing different sentiment categories such as positive, negative, neutral, and other emotion-based indicators extracted from user-generated text. These values are obtained by processing posts from online platforms using text analysis techniques. In addition to emotional features, the dataset also includes corresponding numerical values representing injury cases and death counts collected from official health records. Each row in the dataset represents observations for a specific time period, allowing the data to be treated as a time series. The first row defines the column headers, while the remaining rows consist of actual data entries. This structured format enables efficient analysis, as it aligns emotional patterns with real-world outcomes, making it suitable for training predictive models.

Before applying machine learning algorithms, the dataset undergoes several preprocessing steps to ensure quality and consistency. These steps include separating input features and target variables, handling missing or inconsistent values, and organizing the data into a format suitable for model training. The emotional attributes serve as input features, while injury and death counts act as output labels for prediction. The dataset is then shuffled to remove any bias and divided into training and testing subsets, typically using an 80:20 ratio. This split allows models to learn patterns from historical data and evaluate performance on unseen samples. Performance metrics such as Mean Absolute Error, Root Mean Square Error, and Mean Absolute Percentage Error are used to measure prediction accuracy. Overall, the dataset provides a meaningful combination of behavioral and statistical information, enabling reliable forecasting of self-harm trends.

IV. PROPOSED METHODOLOGY

The proposed system is designed to forecast self-harm trends by integrating emotional insights from social media data with historical records of injury and death. Initially, the dataset is collected from publicly available sources, where emotional features are extracted using text processing techniques such as tokenization and sentiment analysis. These features represent various emotional states like positivity, negativity, neutrality, and other psychological indicators. The dataset is then cleaned to handle missing values, remove noise, and ensure consistency across all attributes. After preprocessing, the input features (emotions) and output variables (injury and death counts) are clearly separated. The dataset is shuffled to eliminate any ordering bias and then divided into training and testing sets, typically using an 80:20 ratio, ensuring proper evaluation of model performance.

In the next phase, multiple machine learning models are implemented to analyze and predict self-harm trends. Time-series forecasting is performed using ARIMA to capture temporal dependencies, while regression and ensemble techniques such as Support Vector Machine, Bayesian Ridge, Random Forest, XGBoost, and Decision Tree are applied to learn complex relationships between emotional signals and outcomes. Each model is trained using the same dataset and evaluated using standard performance metrics including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). Graphical analysis is also performed to compare predicted and actual values, helping to visualize model accuracy. Based on evaluation results, models with lower error values are selected as optimal predictors. Finally, the trained model is used to forecast injury and death values from new social media inputs, enabling early detection and preventive action.

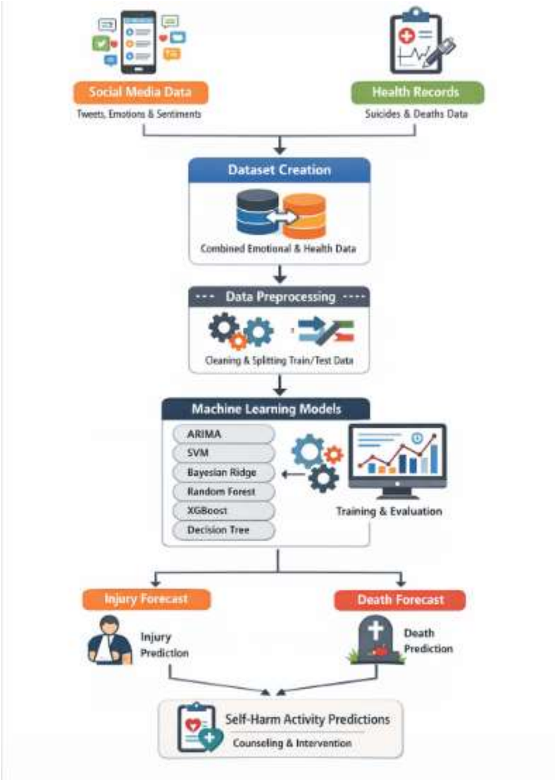

Figure [1]: System Architecture for Forecasting National-Level Self-Harm Trends Using Social Network Data

Figure [1] This architecture illustrates the complete workflow for forecasting national-level self-harm trends using social network data. The process begins with data collection from two main sources: social media platforms, where user posts are analyzed to extract emotional and sentiment features, and official health records containing injury and death statistics. These data sources are combined to form a unified dataset, which then undergoes preprocessing steps such as cleaning, feature selection, and splitting into training and testing sets. The processed data is input into multiple machine learning models including ARIMA, Support Vector Machine, Bayesian Ridge, Random Forest, XGBoost, and Decision Tree for training and evaluation. Each model's performance is assessed using error metrics, and the best-performing models are selected. Finally, the system generates predictions for injury and death trends, enabling early identification of self-harm risks and supporting preventive interventions such as counseling and awareness programs.

V.RESULT AND DISCUSSION

The experimental results demonstrate the effectiveness of different machine learning models in forecasting self-harm trends based on emotional and historical data. Each algorithm was trained and evaluated using standard performance metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). Among all the models, ARIMA showed comparatively higher error values, indicating its limitations in capturing complex relationships within the dataset. In contrast, models like Bayesian Ridge and Support Vector Machine produced moderate performance with reasonably lower error rates. Ensemble methods, particularly Random Forest and XGBoost, achieved significantly better accuracy due to their ability to handle non-linear patterns and feature interactions effectively. The extension using the Decision Tree algorithm also showed competitive performance, in some cases producing even lower MAE values. Graphical comparisons between actual and predicted values revealed that models with better performance had closely overlapping lines, indicating higher prediction accuracy. Overall, the results highlight that advanced machine learning techniques, especially XGBoost and Decision Tree, are more suitable for predicting injury and death trends related to self-harm.

Figure [2]: Self-Harm Prediction System Interface

Figure [2] The interface represents the implementation of the self-harm prediction system using a Jupyter Notebook environment. It provides a structured workspace where users can execute and monitor each step of the workflow. The screen displays the import of required Python libraries and packages necessary for data analysis, visualization, and model building. It also shows the process of loading the dataset, handling missing values, and displaying initial data samples for verification. The interface allows users to perform preprocessing tasks such as feature extraction, data splitting, and model training in an interactive manner. Additionally, it supports execution of various machine learning algorithms and evaluation metrics within the same environment. This setup enables users to easily test different models, observe outputs, and analyze prediction results efficiently, making the system both flexible and user-friendly for experimentation and development.

Figure [3]: ARIMA Model Training and Prediction Output

Figure [3] The interface displays the implementation of the ARIMA model for predicting injury values within the self-harm forecasting system. It shows the process of training the model using the training dataset and generating predictions on the test data. The screen includes the calculated performance metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), which are used to evaluate the model's accuracy. The output section also lists the comparison between actual injury values and predicted values, providing a clear understanding of prediction differences. Additionally, a graphical representation is included where the true values and predicted values are plotted for visual analysis. The noticeable gap between the two lines indicates that the ARIMA model does not perform well for this dataset. This interface helps users analyze model behavior and identify its limitations in capturing complex patterns.

Figure [4]: ARIMA Model Training and Prediction on Death Data

Figure [4] The interface illustrates the application of the ARIMA model for forecasting death values in the self-harm prediction system. It shows the training process using historical death data and the generation of predictions on the test dataset. The output includes performance evaluation metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), which indicate the accuracy of the model. The screen also presents a comparison between actual death values and predicted values, allowing users to assess prediction quality. Additionally, a graph is displayed where true and predicted values are plotted for visual interpretation. The visible differences between the two lines suggest that the ARIMA model has limited accuracy in capturing complex patterns in the dataset. This interface helps in understanding model performance and highlights the need for more advanced algorithms for better prediction results.

Figure [5]: Comparative Performance Analysis of Machine Learning Models

Figure [5] The interface presents a comparative analysis of different machine learning algorithms used for predicting death values in the self-harm forecasting system. It displays a tabular representation of performance metrics including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) for each algorithm. The models compared include ARIMA, Bayesian Ridge, Linear SVR, XGBoost, Random Forest, CatBoost, and the extension Decision Tree. This comparison helps in identifying the most accurate model based on lower error values. From the table, it is evident that XGBoost and the extension Decision Tree achieve better performance with relatively lower error rates compared to other models. The interface allows users to easily evaluate and compare multiple algorithms in a structured manner, supporting the selection of the best model for forecasting. This step plays a crucial role in improving prediction accuracy and overall system effectiveness.

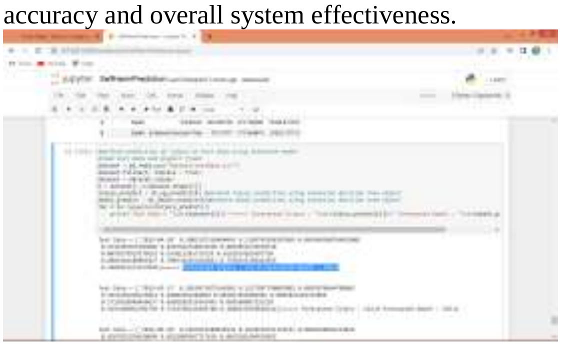

Figure [6]: Forecasting Output Using Extension Decision Tree Model

Figure [6] The interface shows the final prediction stage of the self-harm forecasting system using the extension Decision Tree model. It illustrates the process of loading test dataset values, performing preprocessing, and applying the trained model to generate predictions. The screen displays both the input test data and the corresponding forecasted results for injury and death values. Each output line clearly shows the original data followed by predicted injury and death values, allowing easy comparison. This step demonstrates how the trained model can be applied to new unseen data for real-time forecasting. The results indicate that the Decision Tree model produces consistent and reliable predictions. The interface provides a clear and organized view of prediction outputs, helping users understand how the system transforms input data into meaningful forecasts. This stage confirms the practical applicability of the model in predicting self-harm trends.

DISCUSSION

The findings of this study emphasize the importance of selecting appropriate machine learning models for forecasting sensitive outcomes such as self-harm trends. The poor performance of ARIMA suggests that traditional time-series models may not be sufficient when dealing with complex and multi-dimensional datasets that include emotional features. On the other hand, ensemble and tree-based models demonstrated better adaptability in learning patterns from combined social and health data. The success of XGBoost can be attributed to its boosting mechanism, which improves prediction accuracy by correcting errors iteratively. Similarly, the Decision Tree model performed well due to its simplicity and ability to capture clear decision boundaries within the data. The study also highlights the value of integrating social media sentiment with real-world statistics, as it provides a more comprehensive understanding of mental health trends. However, the approach may face challenges related to data quality, privacy concerns, and generalization across different populations. Despite these limitations, the proposed system shows strong potential in supporting early detection and preventive strategies. It can assist policymakers and healthcare professionals in designing targeted interventions to reduce self-harm incidents.

VI. CONCLUSION

This study presents an effective approach for forecasting national-level self-harm trends by combining emotional patterns from social media with historical health data. The integration of sentiment-based features with real-world injury and death statistics provides a meaningful foundation for predictive analysis. Various machine learning models were applied to evaluate their ability to capture patterns within the dataset. The results showed that traditional models like ARIMA were less effective, while advanced models such as XGBoost and Decision Tree achieved better accuracy with lower error values. This highlights the importance of selecting appropriate algorithms for handling complex and multi-dimensional data.

The proposed system demonstrates how data-driven techniques can support early identification of self-harm risks. By analyzing user-generated content and forecasting potential outcomes, the model can assist authorities and healthcare professionals in planning timely interventions such as counseling and awareness programs. Although challenges like data quality and generalization exist, the approach provides a strong foundation for future improvements. Overall, this work contributes to the development of intelligent systems that can enhance mental health monitoring and help reduce self-harm incidents in society.

REFERENCES

[1] Suicide rates among girls are increasing: Are smartphones responsible? The Economist, May 2023. [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, “Prediction of state-level suicide rates in the U.S. using absolutist language trends from search data,” *Clinical Psychological Science*, vol. 8, no. 4, pp. 788–793, Jul. 2020.
[3] A. E. Aiello, A. Renson, and P. Zivich, “Public health surveillance using social media and internet-based data sources,” *Annual Review of Public Health*, vol. 41, pp. 101–118, Apr. 2020.
[4] M. Akyuz and C. Karul, “Analyzing the influence of economic conditions on suicide in developing nations,” *International Journal of Human Rights in Healthcare*, Jul. 2022.
[5] S. Z. Alavijeh et al., “Insights into psychological disorders through users’ musical preferences on Twitter,” *Information Processing & Management*, vol. 60, no. 3, May 2023.
[6] A. Aldayel and W. Magdy, “Trends and developments in stance detection on social media platforms,” *Information Processing & Management*, vol. 58, no. 4, Jul. 2021.

[7] P. Angelov and A. Sperduti, "Key challenges in deep learning research," in *Proc. ESANN*, 2016, pp. 489–496.
[8] A. Apisarnthanarak et al., "Effects of COVID-19-related anxiety on infection control practices among healthcare workers in Thailand," *Infection Control & Hospital Epidemiology*, vol. 41, no. 9, pp. 1093–1094, Sep. 2020.
[9] S. Arunpongpaisal et al., "Time-series study of suicide trends and predictors in Thailand (2013–2019)," *BMC Psychiatry*, vol. 22, no. 1, Aug. 2022.
[10] J. M. Barros et al., "Evaluating Google Trends data for forecasting national suicide rates in Ireland," *International Journal of Environmental Research and Public Health*, vol. 16, no. 17, p. 3201, Sep. 2019.
[11] B. E. Belsher et al., "A systematic review of predictive models for suicide attempts and deaths," *JAMA Psychiatry*, vol. 76, no. 6, pp. 642–651, 2019.
[12] L. Braghieri, R. Levy, and A. Makarin, "Impact of social media usage on mental health outcomes," *American Economic Review*, vol. 112, no. 11, pp. 3660–3693, 2022.
[13] R. G. Brereton and G. R. Lloyd, "Applications of support vector machines in classification and regression tasks," *Analyst*, vol. 135, no. 2, pp. 230–267, 2010.
[14] J. A. Bridge et al., "National survey on emergency department management of deliberate self-harm cases," *JAMA Psychiatry*, vol. 76, no. 6, pp. 652–654, 2019.
[15] L. Cao, H. Zhang, and L. Feng, "Enhancing suicidal ideation detection using personal knowledge graphs," *IEEE Transactions on Multimedia*, vol. 24, pp. 87–102, 2022.
Babburi, S. Lightweight Distributed Provenance Framework for Edge and IoT Data Systems.
Gaddam, S. From Fixed Specifications to Self-Adapting Systems: A Machine Learning Perspective on Software Engineering.
Immadi, S. K. (2025). Optimizing ERP for Human Capital Management. Applied Research for Growth, Innovation and Sustainable Impact, 377–384. https://doi.org/10.1201/9781003684657-63
Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.
Poojari, R. Frameworks for Data Management and Lineage in Large-Scale Healthcare Data Systems.
Mahimalur, R. K., Vasgam, M., & Manoharan, D. Devops Lifecycle Management And Cloud Migration Assessments: A Security-Driven CICD Perspective.
Viswanathan, V. Generative AI for Smarter Workforce Planning and Enterprise Resource Decisions.
Gajula, S. (2026, March). Two Pillars of Banking Intelligence: A Comparative Analysis of AI Techniques for Fraud Prevention and Churn Mitigation. In 2026 14th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE.
Maturi, S. Y. (2023). Crowdsourced frontier: Unveiling autonomous adversarial cybercapabilities via open AI competition. International Journal of Intelligent Systems and Applications in Engineering, 11(1s), 275–284.
Venkata Ramana, P. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. International Journal of Research in Information Technology and Computing, 8(4).
Kavuri, S. (Ed.). (2024). Shift-left and shift-right testing approaches: A practical roadmap for continuous quality in agile and DevOps. Journal of Information Systems Engineering and Management, 9(4). https://doi.org/10.52783/jisem.v9i4.127
Sreenivasulu Gajula. (2024). Adaptive Zero Trust Architecture for Securing Financial Microservices. Computer Fraud and Security, 643–655. https://doi.org/10.52710/cfs.845
Akinapalli, S. (2024). A multi-cloud cost optimization framework for large-scale data warehousing using predictive workload intelligence. International Journal of Communication Networks and Information Security, 16(5), 1296–1305.
Kandula, S. T. R., Susarla, R. S., & Boyapati, P. K. (2025, July). Enhanced Cyber Security Using Global Local Artificial Neural Network Based Intrusion Detection in Big Data Environment. In 2025 IEEE 4th World Conference on Applied Intelligence and Computing (AIC) (pp. 426-431). IEEE.
Boyapati, P. K. Building a centralized data operations hub for healthcare enterprise integration. IJSAT-Int. J. Sci. Technol. 16 (2). https://doi.org/10.71097/IJSAT.v16.i2.3219