BenchCouncil Transactions on Benchmarks, Standards and Evaluations 4 (2024) 100202

Contents lists available at ScienceDirect

BenchCouncil Transactions on Benchmarks, Standards and Evaluations

journal homepage: www.keaipublishing.com/en/journals/benchcouncil-transactions-onbenchmarks-standards-and-evaluations/

Full Length Article

AI-powered Mathematical Sentiment Model and graph theory for social media trends

M. VENKATACHALAM [a], R.VIKRAMA PRASAD [b,*]

[a] Research Scholar, Department of Mathematics, Government Arts College (Autonomous), Salem-7, Tamilnadu, India
[b] Assistant Professor, Department of Mathematics, Government Arts College (Autonomous), Salem-7, Tamilnadu, India

ARTICLE INFO

Keywords:
Graph theory
Markov decision process
Sentiment analysis
Monkeypox
Knowledge graph

ABSTRACT

Significant issues have arisen as a result of the global spread of monkeypox, such as the extensive transmission of false information, public fear, and stigmatization on social media. Increased fear, prejudice, stigmatization of minority groups, and opposition to public health initiatives are frequently the results of these problems. Furthermore, health authorities are unable to provide correct information and prompt actions due to a lack of efficient methods for analyzing the enormous amounts of unstructured social media data. This disparity weakens crisis management initiatives and increases public skepticism of health guidelines. In order to address these issues, this study looks into the attitude around monkeypox on social media in order to pinpoint public worries, counter false information, and enhance communication tactics. The study intends to improve public comprehension, offer practical insights, and help health authorities manage the outbreak by fusing graph theory with AI-driven sentiment analysis. In order to facilitate semantic analysis of tweets through structured information extraction, graph theory is used to organize unstructured or semi-structured data by creating meaningful links between entities. Furthermore, opinions on monkeypox infection in social media are analyzed and user sentiments are detected using a reinforcement Markov decision process. According to experimental results, the suggested model's accuracy on the Monkeypox tweet dataset was 98 %. These results help raise awareness of monkeypox among the general population and promote an educated and robust social response.

1. Introduction

In 1958, the monkeypox virus was initially identified in research-breed monkey colonies. In 1970, the first recorded human case of monkeypox virus occurred in the Democratic Republic of the Congo. Vaccines against the virus have since been created [1]. The monkeypox virus was deemed exterminated in 1980, and population immunization was discontinued. The fatality rate during a monkeypox outbreak has historically ranged from 1 % to 10 %, despite the fact that the majority of patients may recover. Originally affecting African nations, monkeypox is an infectious illness that has recently spread to almost every city on the planet. Although the world health organization does not recognize it as a pandemic, some experts believe it should be treated as such [2]. Many articles and comments on the symptoms, treatments, side effects, and other people's thoughts about the monkeypox virus have been made on social media sites like Reddit and Twitter. To find patterns and trends, it's critical to examine these user-generated materials [3]. The same tactics may be applied in the event of monkeypox. There are very limited early studies reported for understanding the general public's attitude toward monkeypox or general analysis, but a detailed analysis should be carried out to get a clear picture of the trends and facts [4]. Given the recent spread of monkeypox, associated digital information and opinions have also spread on different social media platforms, including Twitter. Determining the public trends and views about monkeypox is fundamental for governments, policymakers, healthcare providers, and researchers to use the available resources to control and mitigate the burden of the recent outbreak in an efficient and timely manner [5]. Opinion mining primarily deals with a person's concrete view of something, while sentiment refers to an attitude or thought prompted by a feeling [6]. Sentiment analysis and opinion mining [7] were initially used for product review applications but have recently shifted to other tasks, including: stock markets, elections, disasters, healthcare and software engineering [8]. The content shared across social media platforms provides a valuable source of knowledge about the physical environment and social phenomena [9]. As a result, the public security domain has become an important application domain in sentiment

* Corresponding author.
E-mail address: vikramprasad20@gmail.com (R. PRASAD).

https://doi.org/10.1016/j.tbench.2025.100202
Received 9 December 2024; Received in revised form 18 March 2025; Accepted 21 April 2025
Available online 17 May 2025

analysis and opinion mining [10].

In the sentiment analysis and opinion mining, graph theory plays a pivotal role by enabling the modeling of entities and their connections as networks [11]. In weighted graphs, edges carry sentiment scores enable the quantification and sentiment flow across the network, offering deeper insights into the dynamics of public opinion [12]. Knowledge graphs [13], a specialized application of graph theory, extend this concept by integrating semantic relationships between entities. Techniques such as spectral clustering, graph neural network (GNN) [14] and diffusion modeling can predict how sentiments evolve over time or simulate the impact of interventions, such as targeted awareness campaigns. By capturing both structural and contextual relationships, graph theory not only enhances the understanding of public sentiment but equips policymakers with actionable insights to address public concerns effectively during outbreaks like monkeypox [15]. Table 1 shows the detailed analysis of sentiment analysis uses AI, which provides creative ways to examine enormous volumes of unstructured social media data.

Table 1
Existing state-of-art works on sentimental analysis on healthcare social media trends.

Ref.	Method	Key features	Contributions	Performance
[16]	MWGCN	Local Context Weighted Graph (LCG), Multigrain Dot-Product Weighting (MGDW)	Reduces long-distance dependencies, emphasizes local context	Improves sentiment classification accuracy
[17]	GCN	Graph structure-based learning	Captures contextual relationships in data	Accuracy increased by 78 % over baseline
[18]	Combination Graph with KL-divergence	KL-divergence between likelihood models	Enhances in formativeness of nodes in graph-based models	Better structured learning performance
[19]	LSTM + *N*-gram Graph-Cut	Sequence modeling with LSTM and N-gram graph structure	Improves feature extraction and representation	9 % accuracy gain in three-way classification
[20]	Semantic-HGCN	Hierarchical semantic graph encoding	Models multi-level semantic relationships	Improved sentiment prediction
[21]	FGCN	Fuzzy logic integrated with GCN layers	Reduces ambiguity in sentiment detection	Enhances robustness of sentiment classification
[21]	BERT + BiLSTM + GCN	Contextual embedding with BERT-BiLSTM and fuzzy adjacency	Captures deep semantic and structural features	Improves interpretability and classification performance
[22]	Graph using Publication Attributes	Uses attributes like authors, keywords to build graphs	Organizes topics with Louvain community detection	Better thematic structure in sentiment analysis
[23]	Hierarchical Graph Contrastive Learning	Learns local and global representations	Captures complex relationships in utterances	Enhances multimodal sentiment extraction
[24]	MGMFN (GNN + MLP-Mixer)	Combines multiple GNN graphs and long-range MLP features	Strengthens spatial and semantic representation	83.72 % and 86.43 % accuracy in Chinese text classification
[25]	GNN + RF	Uses GNN for learning and Random Forest for classification	Analyzes user attitudes from social data (ChatGPT tweets)	Efficient multi-class sentiment categorization

From the review [11–26], we found the problems associated with using AI-powered sentiment analysis and graph theory for analyzing monkeypox social media trends. While techniques like multi-weight graph convolutional network (MWGCN) and fuzzy graph convolutional networks (FGCN) address local context and syntactic features, their adaptation to monkeypox-specific discussions remains limited [16]. Handling unstructured social media data is a persistent challenge due to its noisy and multimodal nature. Although graph-based methods like Semantic-HGCN and fuzzy logic [17,18]integration help structure data, their scalability for large-scale real-time analysis is insufficient. Existing models, such as hierarchical graph contrastive learning, excel in multimodal sentiment extraction but fail to explicitly tackle the dynamics of misinformation and its impact. Moreover, current sentiment analysis approaches often lack the capacity to analyze interconnected factors such as stigma, misinformation, and public health narratives within a single framework. While hybrid methods like N-gram Graph-Cut [18] combined with LSTM improve accuracy, their application to evolving social media trends and high-dimensional datasets faces scalability challenges. The inability to adequately represent complex social relationships [16–25] is another problem, as discussions about monkeypox often involve intricate interactions between users, groups, and topics. Methods like adjacency graphs and the Louvain algorithm provide structural insights but are limited in capturing nuanced interdependencies in this context [19]. Bias and stigmatization against minority groups also complicate sentiment analysis, requiring models to account for these ethical concerns while ensuring interpretability [21,22]. Despite achieving high accuracy, such as 98 % in monkeypox sentiment classification, AI models lack transparency, hindering their adoption by public health authorities. The limited integration of domain-specific health knowledge and the challenges in adapting to real-time trends restrict the effectiveness of current approaches, making it crucial to develop models that are both adaptable and context-aware.

This study introduces an innovative approach that leverages AI-powered sentiment analysis and graph theory to gain valuable insights into social media trends surrounding monkeypox. By integrating these advanced techniques, the work aims to enhance public understanding, provide actionable insights, and support health authorities in effectively managing the outbreak. The primary contributions of this research are summarized as follows:

1. To address the challenges of analyzing unstructured or semi-structured social media data, graph theory is employed to establish meaningful connections between various entities such as keywords, hash-tags, and user interactions. This facilitates semantic analysis of tweets, transforming chaotic data into an organized framework that can be effectively analyzed for sentiment and thematic trends.
2. The study utilizes a reinforcement Markov decision process (RMDP) to analyze opinions and detect user sentiments regarding monkeypox-related discussions on social media. It enables a nuanced understanding of public perception, including the identification of misinformation, stigmatization, and emotional responses, which are critical for devising targeted communication strategies.
3. The proposed methodology is validated using a comprehensive Monkeypox tweet dataset comprising 61,379 tweets collected from Twitter between May 7 and June 11, 2022. The results showed the model's high accuracy in sentiment analysis and potential to uncover meaningful insights, contributing to resilient public response to the monkeypox outbreak.

2. Material and methods

Fig. 1 illustrates the workflow of the AI-driven sentiment analysis model designed for detecting monkeypox infection sentiment using game theory. The process begins with the Monkeypox tweet dataset, comprising 61,379 tweets published on Twitter between May 7 and June 11, 2022.

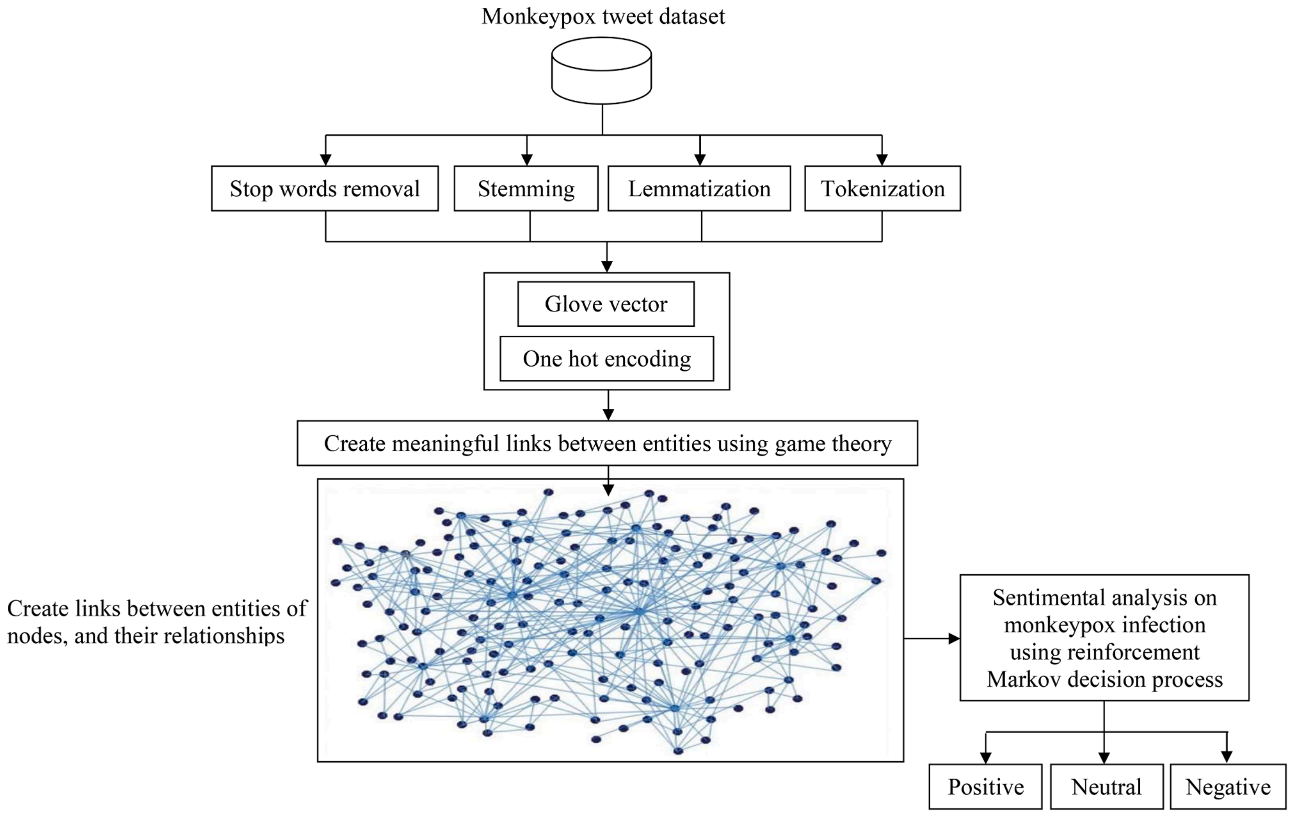

Fig. 1. AI-driven sentiment analysis for monkeypox infection detection using game theory.

The raw textual data undergoes preprocessing steps to ensure its readiness for analysis. Tokenization then splits the text into smaller units, such as words or phrases, enabling efficient processing. After preprocessing, feature extraction techniques are applied to represent the textual data numerically. GloVe vector representation is used to embed words into high-dimensional vector spaces, capturing their semantic and contextual relationships. Additionally, one-hot encoding convert's text into binary vectors, ensuring each unique word is represented distinctly. Next, game theory is employed to create a graph that establishes meaningful links between entities. Nodes in the graph represent entities like keywords, hash tags, user mentions, or topics, while edges are formed based on semantic or contextual similarities. Game-theoretic principles evaluate the importance of these relationships, ensuring that the graph highlights significant patterns and connections among the data. The sentiment classification is performed using a reinforcement Markov decision process (RMDP), which categorizes tweets into positive, neutral, or negative sentiments. RMDP operates by iteratively learning optimal policies through a state-action framework, where states represent the current sentiment context and actions assign appropriate sentiment labels based on extracted features. This iterative learning ensures precise classification by continuously refining predictions. By integrating graph theory for semantic analysis and leveraging RMDP for classification, the model effectively captures public sentiment trends on monkeypox.

2.1. Data source and description

The open-source website GitHub provided the Monkeypox Twitter dataset used in this study, which consists of an extensive collection of tweets on the disease. 61,379 tweets that were posted on Twitter between May 7 and June 11, 2022, make up the dataset [26].These tweets show a variety of public sentiments on monkeypox, including neutral, negative, and favorable views. All tweets are deemed pertinent to debates about monkeypox in this study, offering a wide range of user viewpoints. To make sure the dataset is clean, balanced, and organized for analysis, pre-processing is an essential step before beginning any classification or prediction activities. The gathered raw tweets are naturally disorganized and include superfluous parts like stop words, duplicate records, and non-standardized content. Because Twitter is an unstructured medium with multilingual support, careful data pre-processing is necessary to get relevant results. A two-step procedure is used to eliminate duplicate records at the start of the pre-processing pipeline. Initially, the "is retweets" characteristic that Twitter gave was used to find duplicates. Then, based on their distinct tweet IDs and content, repetitive tweets were removed. After that, tweets were cleaned up to eliminate unnecessary content:

- Elements including URLs, email addresses, hash tags, mentions, and numerical data were removed using TextBlob analyzers.
- To ensure uniformity, all text was changed to lowercase, and stop words and punctuation were eliminated to concentrate on the important information.
- To maintain consistency in the linguistic research, only tweets written in English were kept.

The dataset is now well-structured and prepared for semantic analysis after being cleaned to remove noise and unnecessary information. The dependability of ensuing categorization and prediction tasks is guaranteed by this thorough pre-processing. Furthermore, the dataset utilized in this study is openly accessible, which encourages transparency and makes it possible for the research findings to be replicated.

2.2. Create meaningful links between entities

Creating meaningful links between entities is a crucial step in semantic analysis, particularly when dealing with unstructured or semi-structured data such as social media tweets. This process transforms disorganized datasets into structured knowledge, enabling deeper insights and efficient data interpretation. The primary objective of linking entities is to uncover relationships, organize fragmented information, and support analytical applications such as sentiment analysis and trend monitoring. For instance, tweets about monkeypox may mention various entities like "Human Monkeypox," "Zoonotic Disease," and "Monkeypox Virus." Establishing meaningful connections among these entities allows for better understanding and representation of the data, making it more accessible for analysis. Graph theory plays a central role in this process by structuring data as a network of nodes and edges, where nodes represent entities and edges denote their relationships. Knowledge Graphs (KGs) [27], a practical application of graph theory, facilitate this transformation by encoding information as semantic triplets. These triplets, such as ("Human Monkeypox", "is a", "Zoonotic Disease") or ("Zoonotic Disease", "caused by", "Monkeypox Virus"), represent real-world knowledge in a machine-readable format. KGs are particularly useful for organizing complex datasets, enabling the discovery of hidden relationships and improving the scalability of data analysis. In the context of monkeypox-related tweets, creating meaningful links helps structure the data, enabling accurate sentiment analysis, misinformation detection, and trend identification. Techniques like entity extraction, relationship identification, and graph construction form the foundation of this process. By linking entities based on their semantic relationships, the data becomes more coherent, allowing AI models to perform more effective reasoning and prediction. This structured representation is not only vital for understanding public sentiment but also aids in healthcare analytics by improving data representation and enabling knowledge inference. Ultimately, the integration of graph theory into entity linking enhances the ability to extract meaningful insights from unstructured data, supporting data-driven decision-making in public health and other domains.

When v represents the set of B nodes, $|v| = B$; εrepresents the set of edges linking these nodes, and Z is the adjacency matrix, a graph may be summed up as follows: $J = (v,\ \varepsilon,\ z)$. The networks between any two nodes in v are designated by the adjacency matrix, where the entry of Z in the h-th row and g-th column indicates the significance of the link between the h-th and g-th nodes, and is represented as z_{hg}. The convolution action for spectral-based KG is defined in the Fourier domain by calculating the graph Laplacian Eigen decomposition. The graph that has been normalized Laplacian is definite as (D is the graph's degree matrix and A is its adjacency matrix), where Λ is a diagonal matrix containing its eigenvalues and the columns of U are its eigenvector matrix.

$$j_\theta{}^{*}p = u_{j_\theta}(\wedge)u^S p \tag{1}$$

A Chebyshev polynomial $S_a(p)$of instruction m appraised at $\widetilde{l}$is recycled, and the action is definite as

$$j_\theta{}^{*}p \approx \sum_{a=0}^{A-1} \theta_a S_a(\widetilde{l})p \tag{2}$$

where$\widetilde{l}$ is the diagonal scale matrix. The graph fusion layer in KG combines information from multiple vertices into a single vertex, which reduces the size of the graph and expands the acceptance field of graph filters [28]. To alleviate the problem of overestimating the local neighborhood structure of maps with a very wide node size distribution, the convolutional filter is reduced in size to $K = 1$ and approximated by λ ≈ 2,

$$j_\theta{}^{*}p \approx \theta'_0 p + \theta'_1 p(l - H_B)p = \theta'_0 p + \theta'_1 C^{-1/2} M C^{-1/2} p \tag{3}$$

Here,θ'_0, $\theta'_1$are two unimpeded variables. To restrain the number of limitations and avoid over fitting, KG further assume that$\theta\ = \theta'_0\ - \theta'_1$, leading to the subsequent description of a graph convolution as follows.

$$j_\theta{}^{*}p \approx \theta\left(H_B + C^{-1/2} M C^{-1/2}\right)p \tag{4}$$

The definition of a signal$P \in r^{BPf}$ with C input networks and F filters for functional mapping is as surveys:

$$W = \widetilde{C}^{-1/2}\widetilde{MC}^{-1/2} P\Theta \tag{5}$$

where$\Theta \in r^{DPf}$the matrix is generated by the filter bank parameters, and$W \in r^{BPf}$ is the signal matrix attained by difficulty. GraphSAGE is spatial-GCN that uses node implanting with maximum union combination. In order to save memory while sacrificing time performance, the authors propose a block training algorithm for GCNs. The GraphSAGE framework builds embeddings by selecting and combining features from the local neighborhood of a node.

$$\begin{aligned} i^s_{B_v} &= aggregate_s\left(\left\{i^{s-1}_U,\ \forall U\ \in\ B_v\right\}\right), \\ i^s_V &= \sigma\left(Z^s \cdot \left[i^{s-1}_V \| i^s_{B_v}\right]\right) \end{aligned} \tag{6}$$

where$B_v$the area set of node is$i^s_V$is the concealed state of node V at time step S, and Z^s is the heaviness matrix at layers. Finally, K represents vector concatenation and σ is the logistic sigmoid function. The following is a formulation of the focus mechanism:

$$U_s = \tan i(Zi_s + n) \tag{7}$$

$$\alpha_s = \frac{Exp\left(U_s^S U_z\right)}{\sum_{g=1}^{b} Exp\left(U_s^S U_z\right)} \tag{8}$$

$$T_s = \sum_s \alpha_s i_s \tag{9}$$

wherei_s is the production of every deposit; Z, U_z, and n are trainable masses and bias. The position of every component in i_s is unhurried by appraising the compilation between U_s and i_s, which is randomly prepared. αt is a SoftMax function. A graph courtesy network by stacking a single graph devotion deposit, a, which is a single-layer feed forward neural network, parameterized by a weight vector$\overrightarrow{m} \in r^{2fh}$. The layer figures the constants in the consideration devices of the node pair (h, g) by

$$\alpha_{h,g} = \frac{Exp\left(leakyrelu\left(\overrightarrow{m}^S\left[Z\overrightarrow{i}_h \| Z\overrightarrow{i}_g\right]\right)\right)}{\sum_{K\in B_h B} Exp\left(leakyrelu\left(\overrightarrow{m}^S\left[Z\overrightarrow{i}_h \| Z\overrightarrow{i}_g\right]\right)\right)} \tag{10}$$

where || represents the chain operation. The courtesy layer takes as input a set of node features$i = \left\{\overrightarrow{i_1}, \overrightarrow{i_2}, \ldots, \overrightarrow{i_B}\right\}$, $\overrightarrow{i_1} \in r^f$, where B is the number of protuberances of the input graph and f the number of structures for every node, and foodstuffs set of node features $i' = \left\{\overrightarrow{i_1}', \overrightarrow{i_2}', \ldots, \overrightarrow{i_B}'\right\}$, $\overrightarrow{i_1}' \in r^f$as its output. The first stage in creating higher-level features is to apply a common linear transformation to each node, which is parameterized by a weight matrix $Z \in r^{f'*f}$. Each node may then be subjected to a masked attention mechanism, which yields the following scores:

$$E_{hg} = m\left(X\overrightarrow{i}_h, X\overrightarrow{i}_g\right) \tag{11}$$

Which specifies the position of node g's features to node i. A nonlinearity, σ, can be applied to each node to obtain its final output feature.

$$i'_h = \sigma\left(\sum_{g\in B_h} \alpha_{hg} Z i_g\right) \tag{12}$$

In order to stabilize the learning process, the layer additionally employs numerous attentions. The following representation is produced by combining the individual characteristics that are computed in parallel by k distinct nodes:

$$\acute{i}_h = \|_{k=1}^{k} \sigma\left(\sum_{g\in B_h} \alpha_{hg}^K Z^K \overrightarrow{i}_g\right) \tag{13}$$

By retaining be an average of and delay applying the final nonlinearity.

$$\acute{i}_h = \sigma\left(\frac{1}{K}\sum_{K=1}^{K}\sum_{j\in B_h} \alpha_{hg}^K Z^K \overrightarrow{i}_g\right) \tag{14}$$

where$\alpha_{hg}^K$is the standardized consideration coefficient compute by the k-th attention mechanism.

2.3. Sentiment analysis to detect opinions on monkeypox infection

Once meaningful links between entities are created using Knowledge Graphs, the process of sentiment classification begins. Each tweet is analyzed to determine whether the sentiment expressed is positive, negative, or neutral. This study employs a Reinforcement Markov Decision Process (RMDP) [29] to enhance sentiment analysis by treating it as a sequential decision-making problem. The process begins with pre-processing and feature extraction using techniques like GloVe embeddings and one-hot encoding, which capture the syntactic and semantic nuances of the text. These features are then passed through the RMDP framework, where sentiment analysis is treated as a series of actions within a structured decision-making environment. The Markov decision process (MDP) provides the foundational framework for reinforcement learning by modeling the problem as a set of states, actions, rewards, and transitions. In this context, the states represent tweet features, the actions correspond to sentiment classifications (positive, negative, or neutral), and the rewards signify the accuracy of classifications based on ground truth. RMDP-based sentiment analysis not only improves classification accuracy but also provides a robust and adaptive methodology for analyzing public sentiment around monkeypox infection in social media, offering valuable insights to guide health communication strategies. Then, depending on the state changeover probabilities$X_{t_s,t_{s+1}}(m_s)$, the situation evolves to a new state. For this development, the agent is immediately rewarded withR_s. The agent's objective is to maximize the expected cumulative advertising reward over time by obtaining a policy $\pi(m_s|t_s)$ that associates each state t_s with an accomplishmentm_s.

$$Y^{\pi}(t_s, m_s) = e^{\pi}[R_s|t_s, m_s] \tag{15}$$

For any state action pairings (t, m), a policy π* is considered the best if and only if its projected payoff is greater than or equal to π.

$$Y^{\pi^*}(t, m) \geq Y^{\pi}(t, m) \tag{16}$$

The ultimate objective of an MDP is the Y^{π^*} function, which describes the highest anticipated reward achievable by carrying out a certain action in a specific condition and then following the best course of action.

$$Y^*(t_s, m_s) = e_{t_{s+1}\sim X_{t_s,t_{s+1}}(m_s)}\left[r(t_s, m_s) + \gamma \underset{m'\in M}{Max} Y^*(t_{s+1}, m')\right] \tag{17}$$

where $r(t_s, m_s)$is the instantaneous reward gotten after executing action m_s in state$t_s$at time s, $t_{s+1}$is the next state, $m'$is any achievement that can be busy at t_{s+1}, and γ is a markdown factor that regulates the weight agreed to coming prizes. Conferring to the value of active software design can be written as follows:

$$Y^*(t_s, m_s) = Max\{r(t_s, m_s) + \gamma \sum_{m'\in M} X_{t_s,t_{s+1}}(m_s) Y^*(t_{s+1}, m')\} \tag{18}$$

Since the transition prospects of states ($X_{t_s,t_{s+1}}(m_s)$) and the optimal payoffs of posterior sub-processes ($Max_{m'\in M}Y^*(t_{s+1}, m')$)are often unknown. When given some experience following a policy π, the sentiment analysis updates the estimate $Y(t_s, m_s)$ for the non-terminal states t_s occurring in that experience. The simplest bring up-to-date rule for the opinion detection is as follows:

$$Y(t_s, m_s) \leftarrow Y(t_s, m_s) + \alpha[R_{s+1} + \gamma Y(t_{s+1}, m_{s+1}) - Y(t_s, m_s)] \tag{19}$$

where α is the erudition rate. Signifies an error measures the difference among the value of $Y(t_s, m_s)$ and the better estimation$R_{s+1} + \gamma Y(t_{s+1}, m_{s+1})$. This number is frequently referred to as the false alarm, and arises in various forms throughout reinforcement learning.

$$\delta(s) \doteq R_{s+1} + \gamma Y(t_{s+1}, m_{s+1}) - Y(t_s, m_s) \tag{20}$$

The RMDP consists of two networks [30]: the current network with parameters ω and the target network with parametersω'. On the other hand, the target network is used to estimate the target Y value which guides the training process. The present network's parameters are regularly transferred to the target network at intervals off epochs.

$$l(\omega) = e_{(t,m,r,t')\sim u(C)}[(Y_{tar}(t', m', \omega') - Y_{tar}(t, m, \omega))]^2 \tag{21}$$

$$\nabla_{\omega} l(\omega) = e_{(t,m,r,t')\sim u(C)}[l(\omega)] \tag{22}$$

where$(t, m, r, t') \sim u(C)$specifies that data $(t, m, r, t')$is sampled from involvement replay pool C, and define the objective functions as follows.

$$l'(\omega) = (Y_{tar}(t', m', \omega') - Y_{tar}(t, m, \omega))\nabla_{\omega} Y(t, m, \omega) \tag{23}$$

Instead of randomly selecting the optimal action based on the Y function, the agent randomly selects actions with a fixed probability at each step. If the random number ξ generated by the algorithm is less than the search probability calculated in the power greedy algorithm, then the current optimal solution is based on the objective function (Y) is $\frac{1}{|M|}$selected by probability where |m| indicates the size of the working space [30]. Otherwise, the current optimal action based on the Y function is selected with probability.

$$x(m^*|t) = \begin{cases} \dfrac{1}{|M|} & if\ \xi <\in \\ 1 & if\ \xi <\in \end{cases} \tag{24}$$

wherem^*, Y, |M| the current ideal function of the purpose is the size of the functional space and 0 ≤ ξ ≤ 1 is a even random number. The Softmax system uses the Boltzmann distribution to estimate the Y values of actions and determine the probability of selecting an action.

$$x(m_h|t) = \frac{E^{\frac{Y(t, m_h)}{\tau}}}{\sum_{g=1}^{|M|} E^{\frac{Y(t, m_g)}{\tau}}} \tag{25}$$

where $h = 1, 2, 3, \ldots, |M|$,$x(m_h|t)$ is the probability of choosing an action in state t$Y(t, m_h)$, the expected value Y of the action $m_h$in state t, τ is the temperature parameter, and |M| Size of working space.

$$F(p, q) = \frac{1 - E^{\frac{-|Y_{tar}(t, m, \omega') - Y(t, m, \omega)|}{\sigma}}}{1 + E^{\frac{-|Y_{tar}(t, m, \omega') - Y(t, m, \omega)|}{\sigma}}} \tag{26}$$

where$Y_{tar}(t, m, \omega')$is the Y charge for the mark net trim state act pair (t, m) and y(t, m, ω) is the Y worth for the network close-fitting state action pair (t, m). ω'And ω are the limitations of the goal and existing systems, individually, and σ is a optimistic persistent called the inverse compassion factor. As stated in the earlier section, the false alarm is usually articulated as δ(s).

$$j(\delta(s), \sigma) = \frac{1 - E^{\frac{-|\delta(s)|}{\sigma}}}{1 + E^{\frac{-|\delta(s)|}{\sigma}}} \tag{27}$$

The value of σ the function$j(\delta(s), \sigma)$ is not monotonically decreasing with respect to δ(s) and its value lies in the interval [0, 1). The probet_{s-1} returns the probability ∈ $(t_{s-1})$of state s - 1. where β ∈ (0, 1), typically 1/|M|, which defines the effect of TD error on detection probability, where |M| Indicates the size of the working space.

$$\in (t_s) = (1-\beta)\cdot j(\delta(s), \sigma) + \beta\cdot \in (t_{s-1}) \tag{28}$$

It is designed to optimize an agent's search strategy in an RL system. The search probability is constantly modified depending on the agent's present environmental information by including TD error into the softmax algorithm. The difference between the Y values derived from the target network and the present network is measured by the TD error.

3. Results and discussion

This section presents the results and comparative analysis of sentiment analysis for opinion detection in monkeypox. Python programming is used to do experiments on an X86–64 Ubuntu 18.04.4 LTS computer. With 16 GB of RAM, the CPU is an Intel(R) Core(TM) i7–8550 U running at 1.80 GHz. Using the suggested model, we examine sentiment emotions using this configuration. The Monkeypox Twitter dataset used in this study was sourced from the open-source platform GitHub. It includes a comprehensive collection of tweets related to the disease, comprising 61,379 tweets posted on Twitter between May 7 and June 11, 2022 [31]. These tweets capture a range of public sentiments regarding monkeypox, including neutral, negative, and positive opinions. All the tweets are considered relevant to the discussions surrounding monkeypox, providing diverse perspectives from users. The results of proposed Graph+RMDP model is compared with the existing models such as Navie bayes (NB), support vector machine (SVM), logistic regression (LR), random forest (RF), decision tree (DT), convolutional neural network (CNN), long short-term memory (LSTM) and CNN+LSTM [32]. The performance can validated through different metrics such as accuracy, precision, recall, F-measure and area under curve (AUC).

3.1. Impact of graph theory in sentiment analysis for opinion detection in monkeypox

The integration of graph theory in sentiment analysis offers a powerful method for understanding complex relationships within large datasets, such as social media discussions surrounding events like the monkeypox outbreak. This approach enhances the ability to detect public opinion, categorize sentiments accurately, and uncover deeper insights into societal reactions. Graph theory provides a structured way to organize unstructured social media data. Tweets about monkeypox are often disjointed and varied in content, making it difficult to extract meaningful relationships between entities (e.g., symptoms, prevention, public health policies). By using graphs, each tweet or piece of information can be treated as a node, and relationships between these nodes—such as co-occurrence of terms or sentiment-related associations—can be established as edges. This allows the creation of a network (Fig. 2) that illustrates how different concepts (e.g., fear, misinformation, and vaccination) are interconnected, facilitating deeper semantic analysis.

Traditional sentiment analysis techniques typically classify text into predefined sentiment categories (positive, negative, or neutral). However, this often oversimplifies complex opinions that may contain mixed emotions or contradictory viewpoints. Graph theory helps address this by identifying patterns of sentiment propagation across the network. For example, the sentiment of an individual tweet about monkeypox may influence subsequent tweets, either amplifying or modifying the general public's perception. In the context of monkeypox, misinformation and rumors can spread rapidly, exacerbating panic and confusion. Graph theory enhances misinformation detection by analyzing the flow of information through the network and identifying unusual patterns that may indicate false narratives. For instance, a rapid increase in the spread

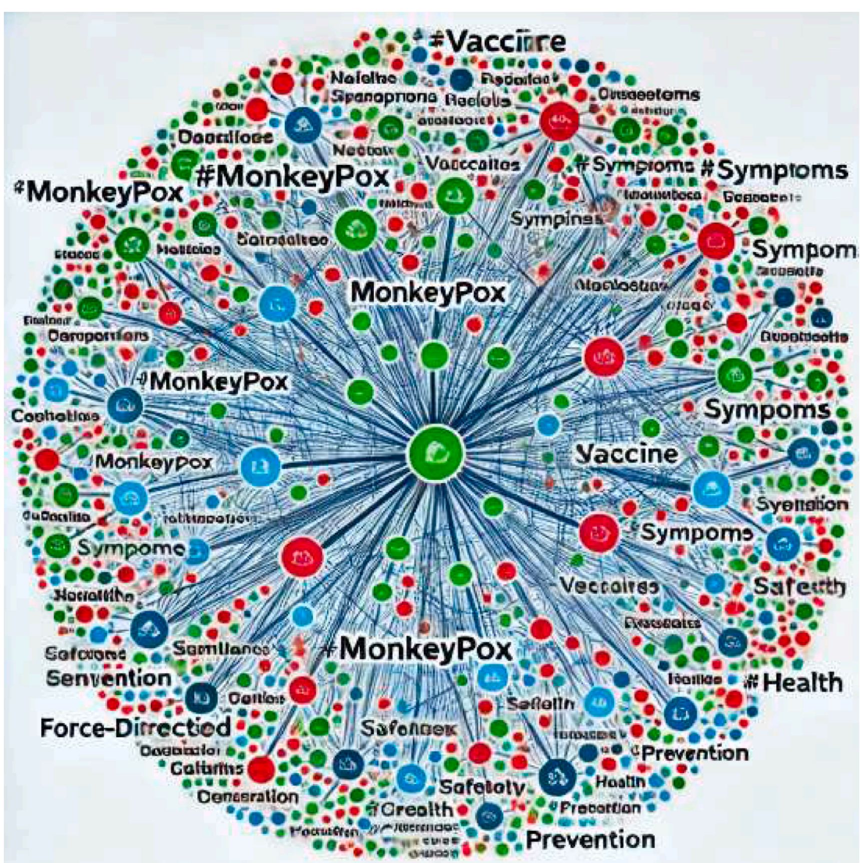

Fig. 2. Knowledge graph for limited tweets from Monkeypox tweet dataset, the plot highlights positive (green), negative (red), and neutral (blue) sentiments. The graph shows the relationships between keywords, hashtags, and user interactions, such as "monkeypox", "vaccine", "symptoms", "health", and "prevention".

of a specific hashtags associated with misinformation could be detected by examining the graph's connectivity and temporal patterns. Sentiment propagation models using graph theory take into account the influence of neighboring nodes in a network. When a tweet expressing a strong sentiment spreads through a network, the sentiment of connected nodes may also shift. Fig. 3visually represents the relationships between key entities and sentiment in the monkeypox tweet dataset. By applying graph-based algorithms, sentiment analysis can account for this dynamic process, leading to a more accurate and nuanced understanding of how sentiment evolves over time and across different demographic groups. This allows for timely interventions and better-targeted public health campaigns regarding monkeypox. By combining sentiment analysis with graph theory, researchers can contextualize public opinion in ways that are more meaningful for decision-makers. For example, graph-based sentiment analysis help identify which areas or regions are most concerned about monkeypox, the type of concerns and how these concerns evolve over time. It enables health authorities to tailor their communication strategies, target high-risk groups, and address misconceptions in a more informed and efficient manner.

3.2. Results analysis of sentiment analysis for monkeypox tweets

This section presents a comparative analysis of the proposed Graph+RMDP model and several existing models, including NB, SVM, LR, RF, DT, CNN, LSTM and CNN+LSTM [32], for sentiment analysis to detect public opinion on monkeypox tweets. The performance of these models is evaluated using various metrics, including accuracy, precision, recall, F-measure, and AUC. Fig. 4 shows the results of the proposed Graph+RMDP model indicate effective learning and generalization as seen in both the loss and accuracy curves across the epochs. The train loss starts at 0.56 in epoch 0 and consistently decreases, reaching a minimal value of 0.0000002 by epoch 59. This steady reduction in train loss suggests that the model is effectively learning from the training data, improving its predictions over time. The test loss follows a similar downward trend, starting at 0.59 and decreasing to 0.17 by epoch 59, albeit with some fluctuations. The test loss is slightly higher than the train loss throughout, which is typical in machine learning models, indicating a minor degree of overfitting. The train accuracy shows rapid improvement, starting at 0.6 in epoch 0 and reaching 1.0 by epoch 20. After this point, it stabilizes at a perfect score, suggesting that the model is fitting the training data very well. In contrast, test accuracy starts at 0.55 and gradually increases, peaking at 0.99 by epoch 23. Although it fluctuates slightly between 0.94 and 0.99 after this point, the general upward trend in test accuracy indicates that the model is not just memorizing the training data but is also able to generalize well to new, unseen data. The fluctuations in test accuracy, especially between epochs 27 to 37, where it dips to 0.95, could be due to minor overfitting or noise in the validation set, but the overall high values demonstrate that the model performs reliably on test data. The Graph+RMDP model shows excellent performance, with a steady decline in both train and test loss and high, stable test accuracy. Although the model achieves perfect training accuracy, the test accuracy fluctuates slightly, indicating a small degree of overfitting. These results suggest that the model is effectively learning from the data, with good generalization ability, though further refinements or regularization techniques could be explored to reduce the minor fluctuations in test accuracy.

Table 2 summarizes the performance analysis of the proposed and existing sentiment analysis models for the Monkeypox tweet dataset. In the accuracy comparison for the Monkeypox tweet dataset (Fig. 5), all

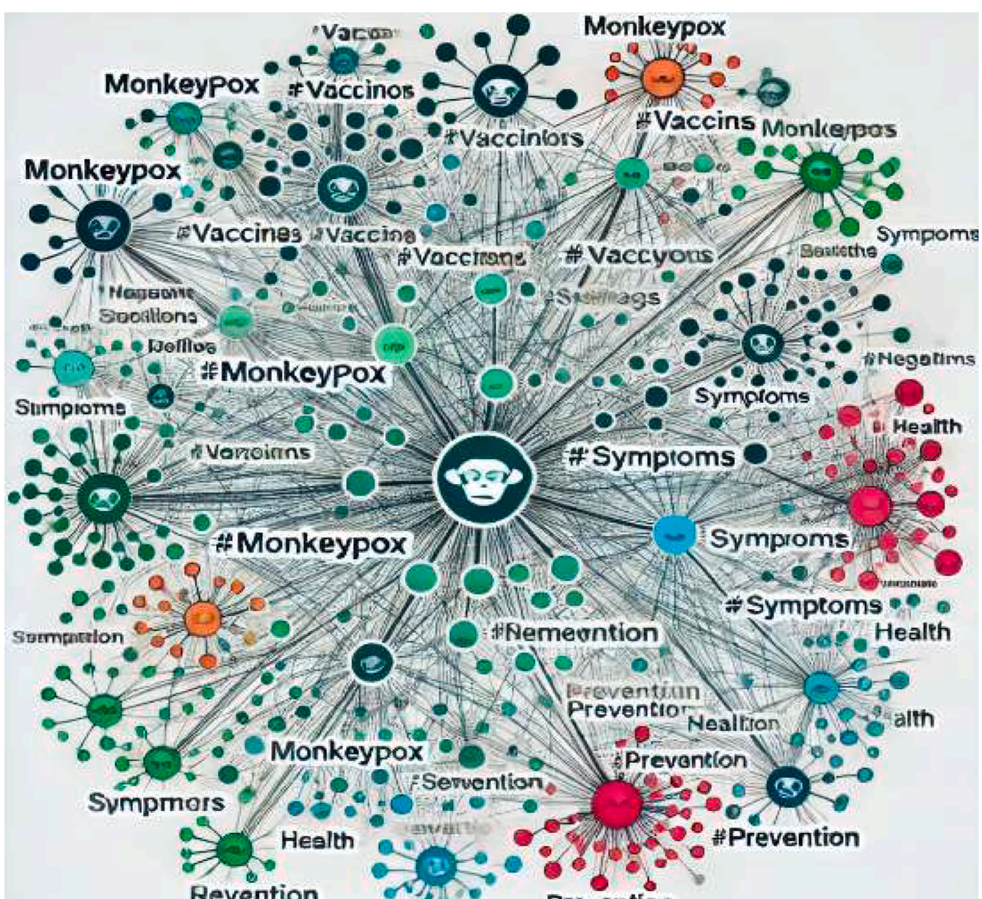

Fig. 3. Knowledge graph for entire Monkeypox tweet dataset.

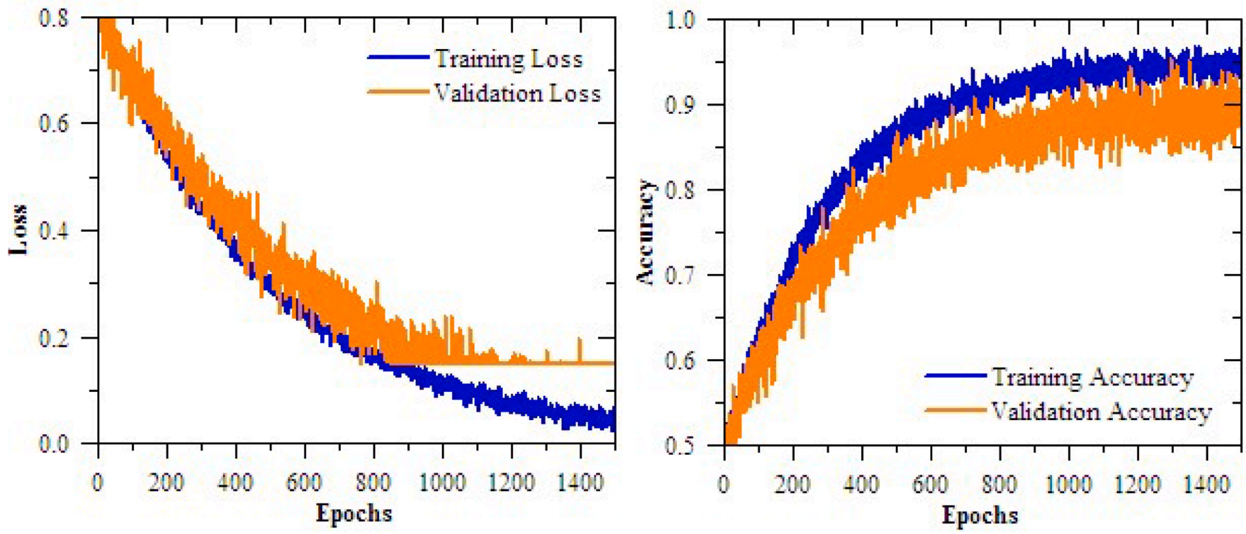

Fig. 4. Loss and accuracy of proposed Graph+RMDP model with varying epochs.

Table 2
Performance analysis of proposed and existing sentiment analysis models for Monkeypox tweet dataset.

Models	Accuracy (%)					Precision (%)				
	300	600	900	1200	1500	300	600	900	1200	1500
NB	75.236	75.858	78.958	81.256	84.578	72.526	73.112	76.436	77.104	80.302
SVM	78.528	79.158	81.254	82.366	85.546	75.307	77.348	79.451	80.903	82.739
LR	80.147	82.355	86.645	89.578	90.125	78.109	78.723	81.286	83.301	85.220
RF	84.153	85.158	86.985	90.124	90.857	80.414	81.752	84.240	85.234	87.518
DT	85.633	86.247	87.985	91.985	92.357	81.201	82.254	84.853	86.053	88.353
CNN	89.547	90.058	91.547	92.285	93.398	83.673	84.381	87.110	89.022	90.441
LSTM	90.258	91.086	92.357	93.325	94.475	85.420	86.539	88.211	89.441	91.012
CNN+LSTM	95.628	95.857	95.957	96.012	96.235	88.119	88.986	90.407	91.228	92.731
Graph+RMDP	98.564	98.618	98.855	98.958	99.245	93.212	94.451	95.640	96.125	97.076
	Recall (%)					F-measure (%)				
NB	70.345	71.850	74.491	75.693	78.204	71.419	72.476	75.451	76.392	79.239
SVM	73.270	74.634	77.122	78.479	81.017	74.275	75.967	78.269	79.673	81.869
LR	75.301	76.429	79.110	80.563	82.354	76.679	77.559	80.183	81.909	83.762
RF	77.028	78.024	80.136	81.543	82.953	78.685	79.845	82.137	83.348	85.174
DT	78.712	79.324	82.139	83.127	85.043	79.937	80.762	83.474	84.565	86.666
CNN	80.210	81.352	84.016	85.178	87.322	81.905	82.839	85.535	87.058	88.854
LSTM	82.145	83.509	85.431	86.450	88.360	83.750	84.997	86.799	87.920	89.666
CNN+LSTM	85.235	85.951	88.024	89.301	90.702	86.653	87.442	89.200	90.254	91.705
Graph+RMDP	90.124	91.016	92.410	93.047	94.310	91.642	92.702	93.997	94.561	95.673

models show improved performance as the dataset size increases from 300 to 1500 epochs. The NB model improves from 75.236 % to 84.578 %, SVM from 78.528 % to 85.546 %, LR from 80.147 % to 90.125 %, RF from 84.153 % to 90.857 %, and DT from 85.633 % to 92.357 %. Deep learning models like CNN and LSTM also perform better, with CNN improving from 89.547 % to 93.398 %, and LSTM from 90.258 % to 94.475 %. The CNN+LSTM hybrid model performs even better, increasing from 95.628 % to 96.235 %. The Graph+RMDP model consistently achieves the highest accuracy, improving from 98.564 % to 99.245 %. This superior performance is attributed to its ability to capture complex relationships between tweet features using graph-based learning, while the RMDP component enhances decision-making by dynamically optimizing classification outcomes based on long-term reward signals. Unlike traditional models that rely on static feature vectors, Graph+RMDP offer a more adaptive and context-aware sentiment analysis approach, making it highly effective for dynamic and noisy social media data.

Fig. 6 presents the precision comparison of the proposed and existing models on the Monkeypox tweet dataset. The Graph+RMDP model achieves the highest precision across all dataset sizes, starting at 93.212 % and reaching 97.076 %, reflecting a 3.86 % improvement. The CNN+LSTM model also shows strong performance, improving from 88.119 % to 92.731 % (4.61 % increase), followed by LSTM and CNN with gains of 5.59 % and 6.77 %, respectively. Among traditional models, DT improves from 81.201 % to 88.353 % and RF from 80.414 % to 87.518 %, showing steady gains. NB starts at 72.526 % and improves to 80.302 % (7.78 % increase), while SVM and LR record improvements of 7.43 % and 7.11 %, respectively. The results highlight the effectiveness of advanced models—particularly Graph+RMDP and CNN+LSTM—in delivering superior precision in sentiment analysis of Monkeypox-related tweets, outperforming traditional approaches by a considerable margin. Fig. 7 illustrates the recall comparison of the proposed and existing models on the Monkeypox tweet dataset. The Graph+RMDP model consistently achieves the highest recall, increasing from 90.124 % at 300 epochs to 94.31 % at 1500 epochs, marking 4.19 % improvement. The CNN+LSTM model follows, improving from 85.235 % to 90.702 % (5.47 % increases). Similarly, LSTM and CNN show notable gains of 6.22 % and 7.11 %, respectively, across the dataset sizes. Among traditional models, DT and RF show steady improvements, with DT rising by 6.33 % and RF by 5.92 %. LR and SVM

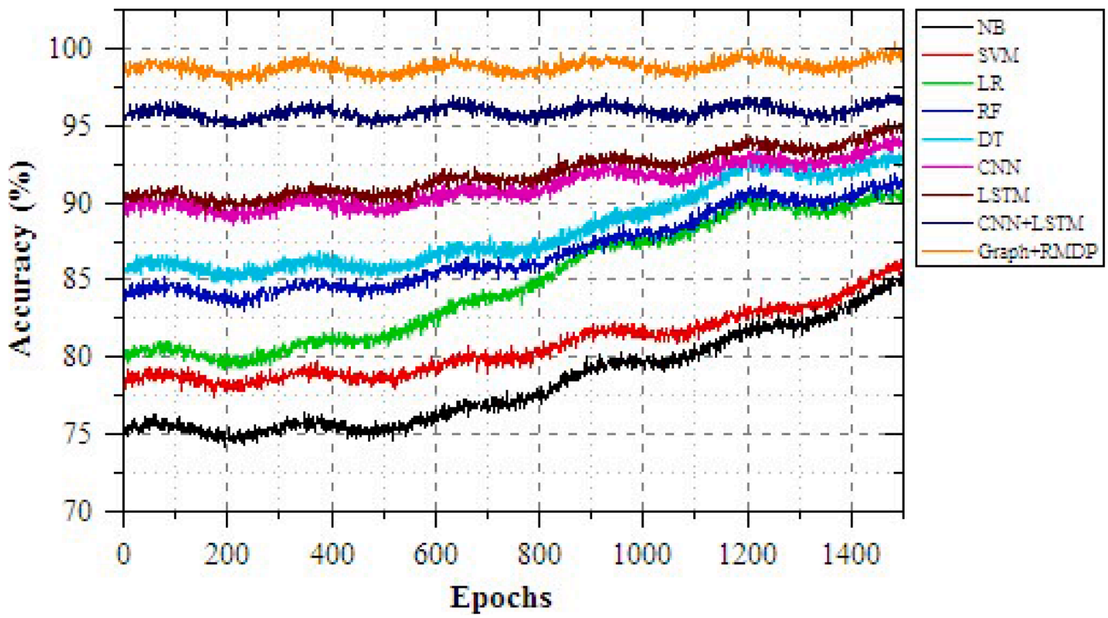

Fig. 5. Accuracy comparison for proposed and existing models for Monkeypox tweet dataset.

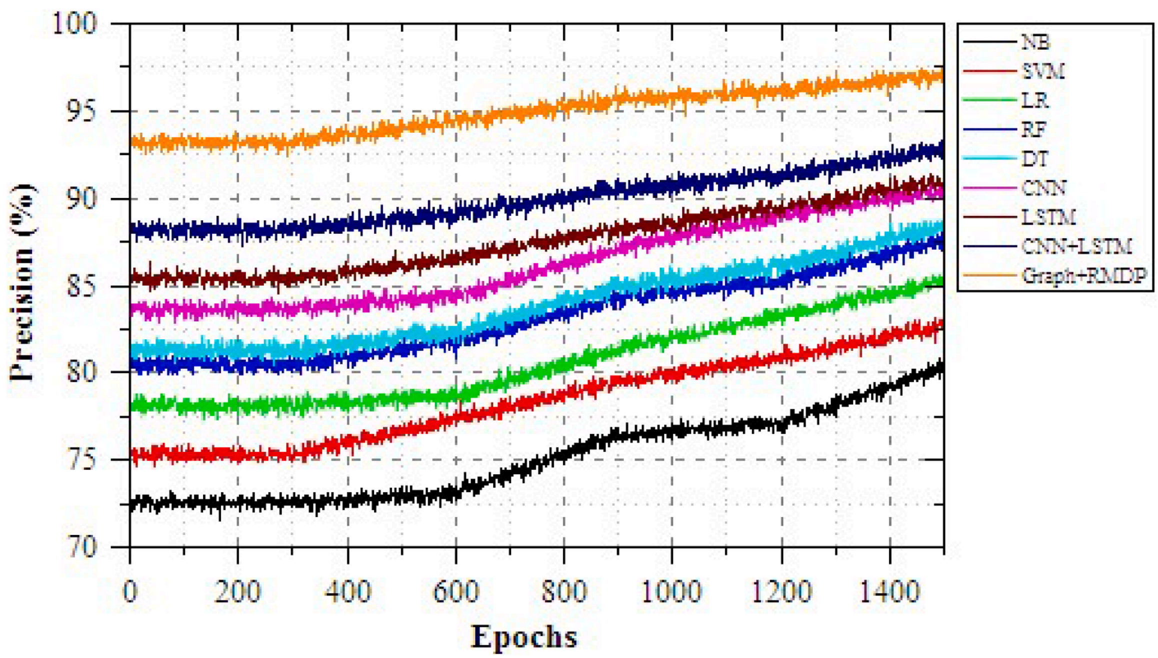

Fig. 6. Precision comparison for proposed and existing models for Monkeypox tweet dataset.

demonstrate moderate growth, with recall increasing by 7.05 % and 7.75 %, respectively. Although NB shows the smallest improvement (7.86 %), it performs relatively well at lower dataset sizes. The results highlight the superior recall performance of advanced models like Graph+RMDP and CNN+LSTM, while traditional models also exhibit consistent gains with increased dataset size.

Fig. 8 presents the F-measure comparison of the models on the Monkeypox tweet dataset. The Graph+RMDP model achieves the highest F-measure, improving from 91.642 % at 300 to 95.673 % at 1500 epochs, reflecting a 4.03 % increase. The CNN+LSTM model follows closely, rising from 86.653 % to 91.705 %, marking 5.05 % improvement. Similarly, LSTM and CNN show notable gains of 5.92 % and 6.95 %, respectively. Among the traditional models, DT and RF exhibit steady performance, with DT improving by 6.73 % and RF by 6.49 %. LR and SVM show moderate increases of 7.08 % and 7.59 %, respectively. The NB model shows the smallest gain, increasing by 7.82 %, from 71.419 % to 79.239 %. The results indicate that hybrid models like Graph+RMDP and CNN+LSTM outperform traditional approaches, offering improvements in F-measure, particularly as dataset size increases.

Fig. 9 shows the performance of various sentiment analysis models for opinion detection on the Monkeypox tweet dataset was assessed across five key metrics: accuracy, precision, recall, F-measure, and AUC. The Graph+RMDP model emerged as the top performer, achieving the highest accuracy of 98.848 %, which was 2.91 % improvement over CNN+LSTM. Compared to traditional models, Graph+RMDP showed a significant increase in accuracy, with models like NB lagging far behind by 19.671 %. The precision of Graph+RMDP was also the highest at

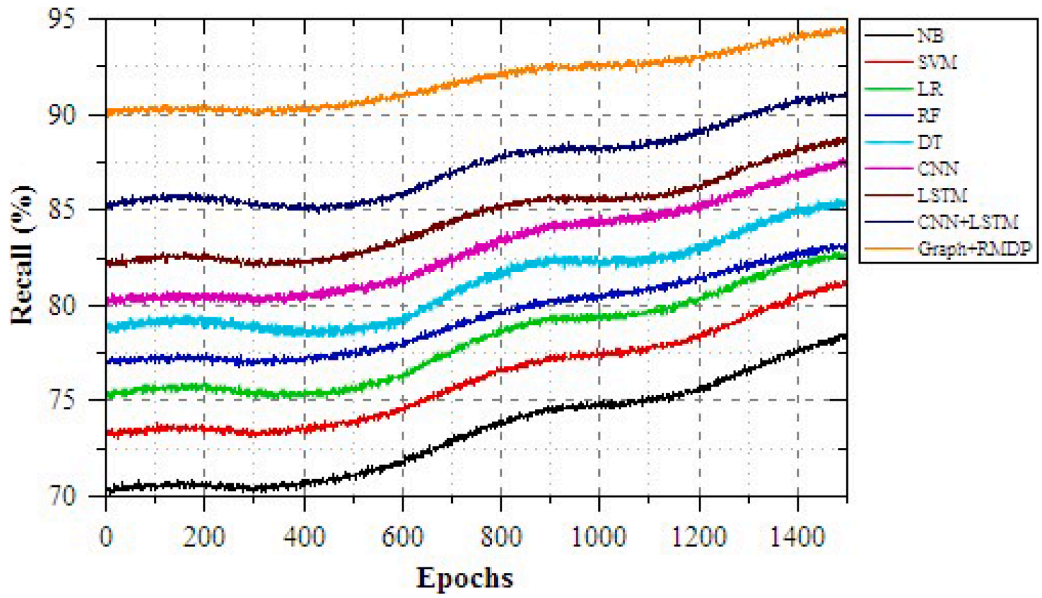

Fig. 7. Recall comparison for proposed and existing models for Monkeypox tweet dataset.

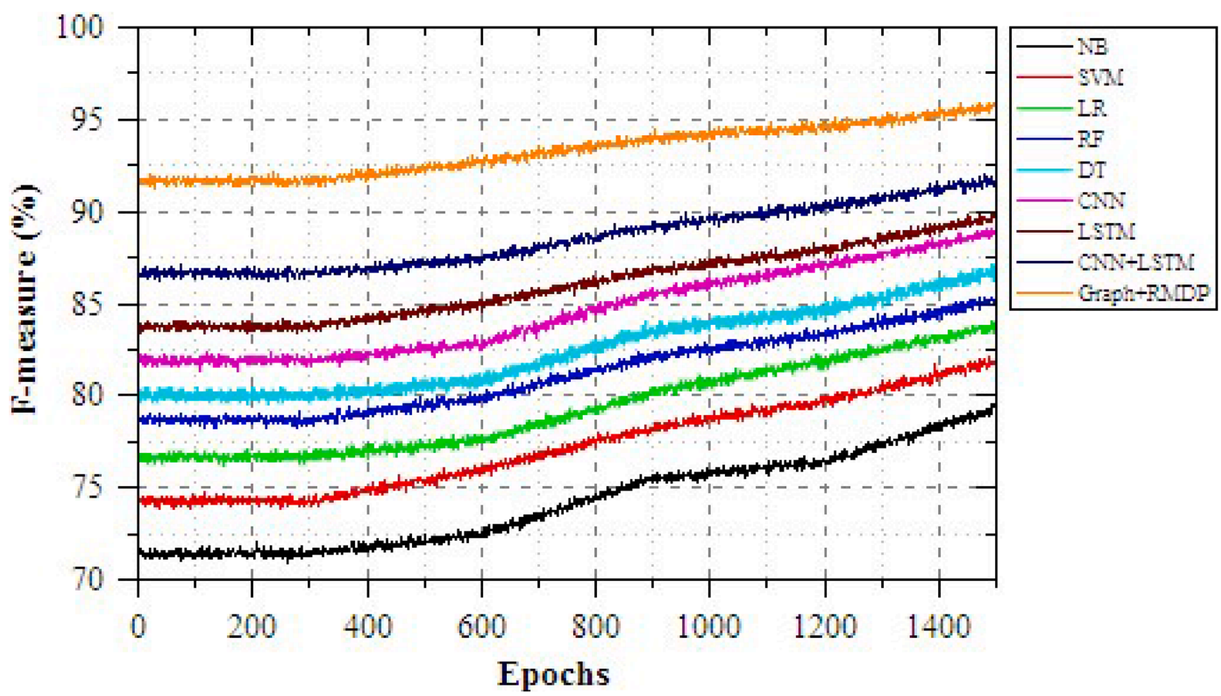

Fig. 8. F-measure comparison for proposed and existing models for Monkeypox tweet dataset.

95.301 %, exceeding CNN+LSTM by 5.007 %, with NB again showing the lowest precision at 75.896 %, a decrease of 19.405 % compared to the top model. In terms of recall, Graph+RMDP led with a score of 92.181 %, outperforming CNN+LSTM by 4.057 %, and traditional models like RF and DT showed lower performance. NB had the lowest recall at 74.117 %, trailing the best model by 18.064 %. For F-measure, Graph+RMDP also achieved the highest score of 93.715 %, surpassing CNN+LSTM by 4.664 %. This marked an impressive improvement over models like RF and DT with NB again showing the weakest performance with 74.995 %, falling behind by 18.72 %. The AUC scores mirrored these findings, with Graph+RMDP achieving 95.8 %, a 3.4 % improvement over CNN+LSTM. Traditional models like RF and DT exhibited lower AUC values, while NB again lagged behind with the lowest score of 78.45 %, a 17.35 % decrease compared to Graph+RMDP. Graph+RMDP consistently outperformed all other models across all metrics, showing significant improvements in accuracy, precision, recall, F-measure, and AUC. While CNN+LSTM also performed well, traditional ML models like NB, SVM, LR, RF, and DT demonstrated lower performance across the board.

4. Conclusion

The proposed AI-powered sentiment analysis, combined with graph theory, effectively addresses the challenges of analyzing unstructured or semi-structured social media data surrounding Monkeypox. By using graph theory to establish meaningful connections between keywords,

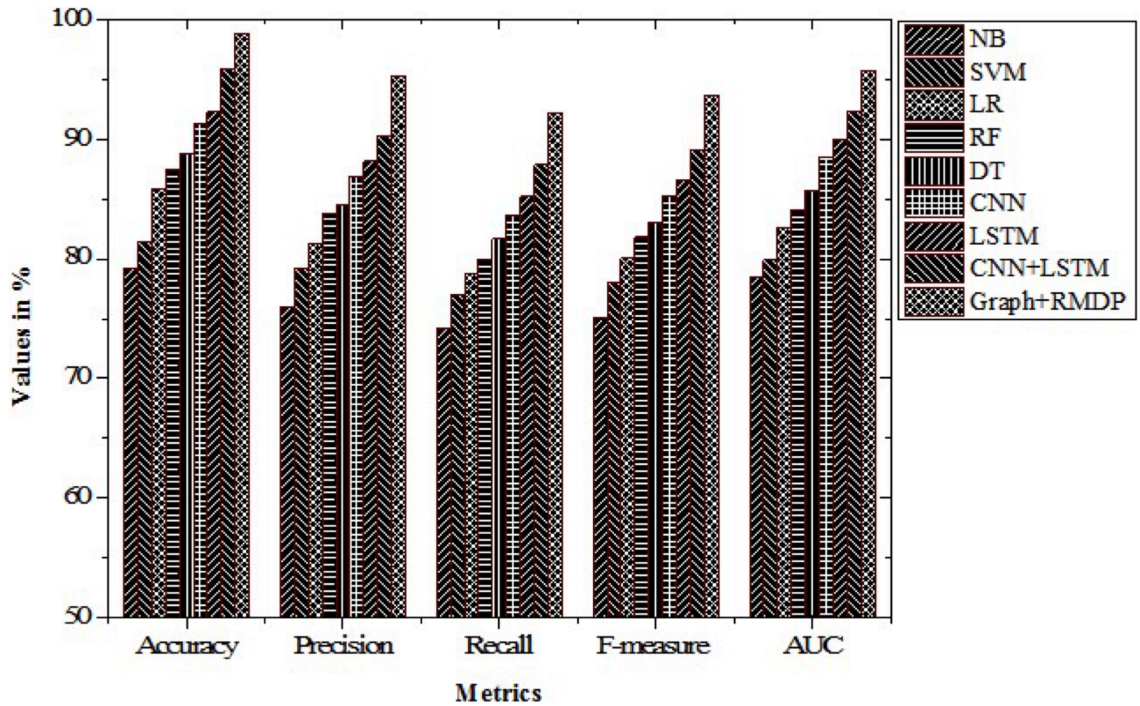

Fig. 9. Performance comparison of different models for sentiment analysis for opinion detection monkeypox.

hashtags, and user interactions, and using a reinforcement Markov decision process (RMDP) to analyze opinions and detect sentiment. The methodology was validated using a Monkeypox tweet dataset comprising 61,379 tweets collected from Twitter between May 7 and June 11, 2022.The results demonstrate that the Graph+RMDP model outperforms existing sentiment analysis models for the Monkeypox tweet dataset. It achieved the highest accuracy of 98.848 %, precision of 95.301 %, recall of 92.181 %, F-measure of 93.715 %, and AUC of 95.8 %, reflecting substantial improvements over the next best model, CNN+LSTM, with increases of 2.91 % in accuracy, 5 % in precision, 4.057 % in recall, 4.664 % in F-measure, and 3.4 % in AUC. When compared to traditional models such as Naïve Bayes, the Graph+RMDP model demonstrated a performance boost of up to 19.671 % in accuracy. The Graph+RMDP model is poised to provide valuable insights into public sentiment and trends related to public health crises like Monkeypox, thereby enabling more informed and data-driven decisions for policymakers and public health organizations.

Ethical approval

Not Applicable

Author contributions

All authors contributed equally to this work and discussed the results and implications and commented on the manuscript at all stages.

Funding

No funding were received from any research agency and organizations.

Availability of data and materials

Data available on reasonable request

CRediT authorship contribution statement

M. VENKATACHALAM: Writing – original draft, Visualization. **R. VIKRAMA PRASAD:** Writing – original draft, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] N. Comeau, A. Abdelnour, K. Ashack, 42223 Assessing public interest in monkeypox via social media platforms: google Trends, YouTube, and TikTok, J. Am. Acad. Dermatol. 89 (3) (2023) AB123.
[2] R.A. Farahat, M.A. Yassin, J.A. Al-Tawfiq, C.A. Bejan, B. Abdelazeem, Public perspectives of monkeypox in Twitter: a social media analysis using machine learning, New Microbes New Infect 49 (2022).
[3] M. Baroudi, I. Smouni, H. Gourram, A. Labzai, M. Belam, Optimizing control strategies for monkeypox through mathematical modeling, Partial Diff. Eq. Appl. Math. (2024) 100996.
[4] K. Soni, A.K. Sinha, Modeling and stability analysis of the transmission dynamics of Monkeypox with control intervention, Partial Diff. Eq. Appl. Math. 10 (2024) 100730.
[5] J. Chire-Saire, A. Pineda-Briseño, J. Oblitas-Cruz, Sentiment analysis of monkeypox tweets in Latin America, in: International Conference on Applied Machine Learning and Data Analytics, Springer Nature Switzerland, Cham, 2023, pp. 230–245.
[6] S. Li, J. Chen, Virtual human on social media: text mining and sentiment analysis, Technol. Soc. 78 (2024) 102666.
[7] J. Du, J. Xu, H. Song, X. Liu, C. Tao, Optimization on machine learning based approaches for sentiment analysis on HPV vaccines related tweets, J. Biomed. Semantics 8 (2017) 1–7.
[8] Z.A. Khan, Y. Xia, A. Khan, M. Sadiq, M. Alam, F.A. Awwad, E.A. Ismail, Developing lexicons for enhanced sentiment analysis in software engineering: an innovative multilingual approach for social Media reviews, Comput. Mater. Contin. 79 (2) (2024).
[9] J. Yang, Y. Xiong, Social media sentiment contagion and stock price jumps and crashes, Pacific-Basin Finance J. 88 (2024) 102520.
[10] M. Ashayeri, N. Abbasabadi, Unraveling energy justice in NYC urban buildings through social media sentiment analysis and transformer deep learning, Energy Build 306 (2024) 113914.
[11] W. An, F. Tian, P. Chen, Q. Zheng, Aspect-based sentiment analysis with heterogeneous graph neural network, IEEe Trans. Comput. Soc. Syst. 10 (1) (2022) 403–412.

[12] L. Yuan, J. Wang, L.C. Yu, X. Zhang, Syntactic graph attention network for aspect-level sentiment analysis, IEEe Trans. Artif. Intell. 5 (1) (2022) 140–153.

[13] K. Liang, L. Meng, M. Liu, Y. Liu, W. Tu, S. Wang, S. Zhou, X. Liu, F. Sun, K. He, A survey of knowledge graph reasoning on graph types: static, dynamic, and multi-modal, IEEE Trans. Pattern Anal. Mach. Intell. (2024).

[14] X. Luo, S. Zhang, J. Wu, H. Chen, H. Peng, C. Zhou, Z. Li, S. Xue, J. Yang, ReiPool: reinforced pooling graph neural networks for graph-level representation learning, IEEE Trans. Knowl. Data Eng. (2024).

[15] M. Kalaimathi, B.J. Balamurugan, Topological indices of molecular graphs of monkeypox drugs for QSPR analysis to predict physicochemical and ADMET properties, Int. J. Quantum Chem. 123 (22) (2023) e27210.

[16] B. Yu, S. Zhang, A novel weight-oriented graph convolutional network for aspect-based sentiment analysis, J. Supercomput. 79 (1) (2023) 947–972.

[17] A.M. Schoene, L. Bojanić, M.Q. Nghiem, I.M. Hunt, S. Ananiadou, Classifying suicide-related content and emotions on Twitter using Graph Convolutional Neural Networks, IEEE Trans. Affect. Comput. 14 (3) (2022) 1791–1802.

[18] V. Shumovskaia, M. Kayaalp, M. Cemri, A.H. Sayed, Discovering influencers in opinion formation over social graphs, IEEe Open. J. Signal. Process. 4 (2023) 188–207.

[19] B.P. Nandi, A. Jain, D.K. Tayal, Aspect based sentiment analysis using long-short term memory and weighted N-gram graph-cut, Cognit. Comput. 15 (3) (2023) 822–837.

[20] Y. Zeng, Z. Li, Z. Chen, H. Ma, Aspect-level sentiment analysis based on semantic heterogeneous graph convolutional network, Front. Comput. Sci. 17 (6) (2023) 176340.

[21] H.T. Phan, N.T. Nguyen, A fuzzy graph convolutional network model for sentence-level sentiment analysis, IEEE Trans. Fuzzy Syst. (2024).

[22] W. Fu, S. Akbar, Expert profile identification from community detection on author-publication-keyword graph with keyword extraction, IEEe Access. (2024).

[23] J. Du, J. Jin, J. Zhuang, C. Zhang, Hierarchical graph contrastive learning of local and global presentation for multimodal sentiment analysis, Sci. Rep. 14 (1) (2024) 5335.

[24] Y. Kang, X. Yang, L. Zhang, X. Luo, Y. Xu, H. Wang, J. Liu, MGMFN: multi-graph and MLP-mixer fusion network for Chinese social network sentiment classification, Multimed. Tools Appl. (2024) 1–22.

[25] V.S. Anoop, C.S. Krishna, U.H. Govindarajan, Graph embedding approaches for social media sentiment analysis with model explanation, Int. J. Inf. Manage. Data Insights 4 (1) (2024) 100221.

[26] R. Olusegun, T. Oladunni, H. Audu, Y.A.O. Houkpati, S. Bengesi, Text mining and emotion classification on monkeypox Twitter dataset: a deep learning-natural language processing (NLP) approach, IEEe Access. 11 (2023) 49882–49894.

[27] L. Yang, H. Chen, Z. Li, X. Ding, X. Wu, Give us the facts: enhancing large language models with knowledge graphs for fact-aware language modeling, IEEE Trans. Knowl. Data Eng. (2024).

[28] Chen, Z., Zhang, Y., Fang, Y., Geng, Y., Guo, L., Chen, X., Li, Q., Zhang, W., Chen, J., Zhu, Y. and Li, J., 2024. Knowledge graphs meet multi-modal learning: a comprehensive survey. arXiv preprint arXiv:2402.05391.

[29] A. Bennett, N. Kallus, Proximal reinforcement learning: efficient off-policy evaluation in partially observed markov decision processes, Oper Res 72 (3) (2024) 1071–1086.

[30] P. Adjei, N. Tasfi, S. Gomez-Rosero, M.A. Capretz, Safe reinforcement learning for arm manipulation with constrained Markov decision process, Robotics 13 (4) (2024) 63.

[31] N. Thakur, Monkeypox2022tweets: The fIrst Public Twitter dAtaset On the 2022 MonkeyPox oUtbreak, 2022 2022060172, https://doi.org/10.20944/preprints202206.0172.v1. Preprints.

[32] Gaurav Meena, Krishna Kumar Mohbey, Sunil Kumar, K. Lokesh, A hybrid deep learning approach for detecting sentiment polarities and knowledge graph representation on monkeypox tweets, Decision Anal. J. 7 (2023) 100243, https://doi.org/10.1016/j.dajour.2023.100243. ISSN 2772-6622.