GraphFusion Sentiment Analyzer: Integrating Heterogeneous GNNs and Transformers for Enhanced Social Media Sentiment Analysis

NagaCharitavya Madala, Sai Durga Saradhi Pranu Deepak Tallapudi, Laikya Reddy Katta, Venkata Srikari Malladi, and Rajiv Senapati

SRM University, Andhra Pradesh, India
{nagacharitavya_m, deepak_tallapudi, laikyareddy_katta, venkatasrikari_m, rajiv.s}@srmap.edu.in

Abstract. Twitter's informal language and changing content present a challenge to established sentiment analysis techniques. This paper presents a unique graph-based method for improving sentiment analysis accuracy using Twitter data. We create graphs from tweets, with nodes representing words or phrases and edges showing relationships. Our approach combines a variety of techniques, including Heterogeneous Graph Neural Networks (Hete-GNNs), Gaussian Naive Bayes using Node2vec, a combined BERT-LSTM-CNN model(Transformer model), and hybrid approaches such as KNN+XGBoost and CNN+LSTM. These approaches produce rich vector representations that are fed into Artificial Neural Networks for sentiment prediction. Our approach outperforms established methods on different datasets, providing a reliable solution for real-time sentiment analysis in the ever-changing social media scene.

Keywords: Sentiment Analysis · Graph Embedding · Machine Learning · Neural Networks.

1 Introduction

The widespread use of social media platforms, particularly Twitter, has increased the importance of sentiment analysis in natural language processing (NLP).These platforms provide real-time insights into public emotions, but traditional methodologies are challenged by the content's informal and diverse nature, which includes slang, sarcasm, and multilingual phrases. To address these complexity, researchers are now using textual and graph-based representations to capture intricate sentiment patterns and increase accuracy, particularly in Twitter sentiment analysis.Recent studies have used creative ways to address the constraints of standard sentiment analysis. Some researchers see sentiment analysis as a link prediction problem that combines user connections and sentiment variables to improve graph network embeddings and classification accuracy. Heterogeneous graph convolutional networks (GCNs) are used to capture syntactic and semantic links, which improves the analysis of complicated text structures.

Multi-level graph neural networks (MLGNNs) improve accuracy by incorporating features at different levels using attention mechanisms.Multi-view learning frameworks handle issues such as data sparsity and multilingualism by combining text-based and graph-based representations. Deep graph convolutional neural networks (DGCNNs) and BERT classifiers are used to improve sentiment classification by combining data from different sources. Innovative ideas include the use of graphlet kernels for faster graph layout predictions and the integration of transformers with GCNs in the Sentiment Transformer Graph Convolutional Network (ST-GCN) to better handle diverse textual graphs.Integrating external information sources and increasing aspect-based sentiment analysis (ABSA) have improved performance in both English and multilingual environments. Attention mechanisms are used in conjunction with graph neural networks to combine text and user-connection data to improve sentiment prediction. Custom loss functions and multilayered structures improve data fusion and outperform conventional approaches. To summarize, integrating several data sources with strong machine learning models is critical for capturing the complexity of human mood on dynamic platforms such as Twitter. Our proposed approach makes use of graph-based embeddings and machine learning classifiers to make more accurate and robust sentiment predictions. The following are the contributions from this work.

- In this work we have proposed a Graph based Fusion Sentiment Analyzer using heterogeneous GNNs and Transformers.
- The proposed method may be helpful in enhancing social media sentiment analysis.

The rest of this paper is organized as follows. Sec. II presents the related works available in the literature. Sec. III presents the working methodology. Sec. IV presents the obtained results. Finally, Sec. V concludes this paper.

2 Literature Review

A technique for segmenting sentiment features by combining dependency relations into neural network models was presented in [1]. They tackled the problem of determining the sentiment polarity of particular aspect words in intricate phrases. In order to measure this relatedness, the authors devised a dependency weighting technique and suggested employing dependency relations to capture sentiment elements that are directly linked to aspect terms. The work reported in [2] addressed the drawbacks of conventional GNNs that only concentrate on nearby words by proposing a multi-level graph neural network (MLGNN) to capture both local and global variables in text sentiment analysis. The model combines word node features efficiently by providing a scaled dot-product attention mechanism and introducing node connection windows of different sizes. On public datasets, MLGNN scored better than other models in sentiment analysis tasks. In order to assess aesthetic measures and enhance graph layout, [3] presented a machine-learning strategy for big graph visualization utilizing newly created graphlet kernels. Compared to conventional methods, their methodology produced faster and more accurate findings, and user testing verified that computed topological similarity and human perception aligned.Twitter's importance as a top Online Social Network (OSN) was investigated by [4]. Because of its straightforward data model and API, Twitter is a great choice for social network research. Three main areas of Twitter research were laid out in this study: sentiment analysis, social graph structure, and risks such as hate speech, spam, bots, and fake news. It demonstrated machine learning, NLP, and graph sampling computing approaches in addition to evaluating deep neural networks with an accuracy of 82%.

The goal of this thorough survey was to provide guidance to scholars as they investigated Twitter's research landscape and methodology.NESA is a novel technique to sentiment analysis that treats it as a sign link prediction issue in social networks, as proposed in [5]. In order to enhance graph network embeddings, NESA incorporates user linkages, emotion polarity, and user-entity properties, emphasizing user interactions over grammatical analysis alone. By more accurately anticipating sentiment connection polarities, the system performs better than conventional techniques.A thorough survey of sentiment analysis (SA) research from 2002 to 2014 was given by [6] The research was divided into seven categories: opinion spam detection, lexicon building, subjectivity categorization, sentiment classification, review usefulness, and SA applications. In [7] suggested encoding heterogeneous relation information as a novel method for judging emotion tendency. They created a heterogeneous graph to preserve syntactic dependencies by using an interactive module to extract target aspect and context information and a GRU module to collect text sequence features. The efficacy of the model was established by experiments. Using the LDA model presented in [8] investigated sentiment analysis in social media to pinpoint important subjects from Scopus abstracts. They identified fifteen primary areas of focus, including big data, sentiment analysis in reviews, and machine learning, with applications in decision-making and transportation. Future research on clustering techniques and data sources expansion is recommended by the study. A context-dependent heterogeneous graph convolutional neural network for implicit sentiment classification was presented in [9]. The review provides direction for future opinion mining research by highlighting the trade-off between the precision of supervised approaches and their slowness and cost.In order to capture both syntactic and semantic data, [10] suggested a GNN-LSTM model for Weibo sentiment analysis by building semantic graphs with dependency syntax. With LSTM and a spatial domain graph filter, the model produced an accuracy of 95.25% and an F1 score of 95.22%. Future advancements focus on enhancing text relationships and introducing higher-order characteristics for richer semantic graph representation.An underutilized technique in neural network research is the combination of tweet-text and user-connection data for Twitter sentiment analysis. Word-embedding, user-embedding, and attentional graph network layers, along with customized loss functions, are the three layers that make up the

Attentional-Graph Neural Network-based Twitter Sentiment Analyzer (AGN-TSA) presented in [11]. Tests conducted using U.S. election data from 2016 reveal that AGN-TSA outperforms current techniques by more than 5%. AGN-TSA is different from Xing's ECN-LSTM in that it uses text and graph data instead of numerical data, focuses on sentiment analysis instead of asset allocation, and has a three-layered attentional network design instead of an RNN architecture. In information fusion tasks, it outperforms DSF by around 4% and also complies with privacy rules.Sentiment analysis is still difficult since internet content is so varied. By ranking terms worldwide and assigning domain-specific polarity, the Sentiment Analysis using Keyword Rank-based Polarity Assignment (SAKRPA) model, as introduced in [12] improves analysis results. To increase Node and Edge Rank (NE-Rank), SAKRPA uses a co-occurrence graph-based approach with a novel node weighting technique. Compared to four previous models, it exhibits increased accuracy and efficiency while addressing keyword bi-polarity and domain dependency. The goal of future work is to extend SAKRPA to support more review characteristics, emoticons, and abbreviations, with applications in political prediction and recommender systems.Deep neural networks (DNNs) have made significant strides in sentiment analysis in natural language processing (NLP); nonetheless, there are still issues with feature space dimensionality and the disregard for textual graph heterogeneity. The Sentiment Transformer Graph Convolutional Network (ST-GCN), which combines Transformer and graph convolutional network (GCN) approaches to manage heterogeneous graphs, is introduced in [13]. ST-GCN uses position encoding and message passing to handle sentiment analysis as a node classification task. On the SemEval, SST-B, IMDB, and Yelp 2014 datasets, it performs better than current models; next research will concentrate on Dynamic Neighborhood Aggregation and other network tasks. The AKM-IGCN approach is put forth in [14], who improve aspect-based sentiment analysis (ABSA) by using SenticNet's affective knowledge. The model works well in both Chinese and English datasets, enhancing performance through multi-head self-attention and an effective knowledge-augmented dependency tree. Word analysis and syntactic links will be improved in further studies.T-GCN for ABSA, which leverages attention mechanisms and dependency types in graph structures to capture contextual information, is introduced in [15]. On six English benchmark datasets, T-GCN achieves state-of-the-art accuracy and F1 scores.SciCheck, a technique for adding 300,000 new triples to scientific knowledge graphs (KGs), is presented in [16]. It performs better than current approaches, and further research will provide hypotheses and expand to other scientific KGs.The Heterogeneous Aspect Graph Neural Network (HAGNN) for ABSA is proposed in [17]. It leverages a heterogeneous graph to capture structural and semantic links. Significant gains in accuracy and F1 scores are demonstrated using HAGNN on a variety of datasets, including difficult ones. More than forty Twitter Sentiment Analysis (TSA) techniques are surveyed by [18], who group the techniques into four categories: machine learning, lexicon-based, hybrid-based, and graph-based. The study outlines current investigations and makes recommendations for future advancements using big data, cognitive science, and sophisticated visualization. Other approaches available in the literature [19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36] also can be used for such analysis.

3 Methodology

In this section, we have presented our proposed framework as depicted in Fig. 1. The dataset we have used is from Kaggle [37]. This dataset is a valuable resource for sentiment analysis related to the 2019 Indian General Lok Sabha elections. It contains English tweets from January 1 to May 23, 2019, focusing on tweets with relevant hashtags or references to political leaders. The dataset has been cleaned to include only unique entries, with primary columns being Date (the tweet's timestamp), Tweet (the tweet's full content), User (the person who posted the tweet), and Emotion (labeled as "pos" for positive or "neg" for negative sentiment). There are no missing values, and we analyzed the data distributions further. To ensure rigorous evaluation, the dataset was divided 80:20, with 80% used for training and 20% for testing.

The machine learning framework employs a variety of models, such as Heterogeneous Graph Neural Networks (Hete-GNNs), a combination of BERT, LSTM, and CNN (Transformer model), Gaussian Naive Bayes with Node2Vec, k-Nearest Neighbors (KNN) with XGBoost, Convolutional Neural Networks (CNN) combined with LSTM, and Artificial Neural Networks. These models were chosen because of their unique characteristics in dealing with a variety of data kinds, including graph-structured, text, and time-series data. The framework highlights the need of tailoring preprocessing strategies for each model and assesses their performance across a wide range of machine learning applications using measures such as accuracy, precision, recall, F1-score, and support.

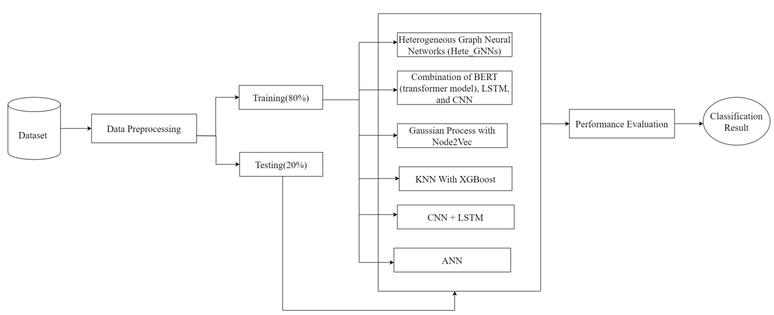

Fig. 1. The proposed framework.

3.1 Heterogeneous Graph Neural Networks

In our sentiment analysis project, we use Heterogeneous Graph Neural Networks (Hete-GNNs) to simulate complicated relationships in text data. Text, attributes, and user information are represented as independent nodes in a graph, with edges indicating how they interact. The MLHGAN (Multilayer Hybrid Graph Attention Network) model processes this graph by combining GCN, GAT, and Transformer layers, which capture both local and global patterns. This method efficiently combines textual information and contextual insights, improving sentiment prediction accuracy by accounting for intricate interactions between words, aspects, and users.

3.2 Combination of BERT with LSTM and CNN layer(Transformer Model)

In our sentiment analysis project, we employ a hybrid machine learning pipeline that integrates pre-trained BERT embeddings with LSTM and CNN layer. BERT tokenizes text, and a custom 'SentimentDataset' class prepares the data for training and validation. The HybridSentimentModel utilizes BERT for embeddings, LSTM for sequential analysis, and CNN for feature extraction. This architecture improves sentiment classifier accuracy. The Adam optimizer and CrossEntropyLoss are used during training, and performance is measured using accuracy, classification reports, and confusion matrices. This method combines the strengths of each component to ensure accurate and robust sentiment analysis.

3.3 Gaussian Naive Bayes with Node2Vec

In this work, we utilize Node2Vec to extract node embeddings from a graph. These embeddings are then reduced to 2D using PCA and trained using a Gaussian Naive Bayes classifier. The model's performance is evaluated using accuracy scores, a confusion matrix, and a classification report, with the confusion matrix visualized using a heatmap. This technique effectively captures network relationships, allowing for strong node classification.

3.4 CNN-LSTM hybrid model

In this work, we build a CNN-LSTM hybrid model for sequential data categorization using image sequences as input. Each time step (image frame) is processed individually by the TimeDistributed CNN, which uses convolutional and pooling layers to extract spatial data. These features are flattened and sent to an LSTM layer, which detects temporal dependencies in the sequence. Class probabilities are output from the final dense layer.The model is trained with categorical cross-entropy and the Adam optimizer, and its accuracy, classification report, and confusion matrix are all evaluated. This design efficiently manages both spatial and temporal data, resulting in increased classification accuracy.

3.5 Artificial Neural Network (ANN)

In this project, we use TensorFlow and Keras to build a simple Artificial Neural Network (ANN) for multi-class classification. The input data consists of 20 characteristics and 5 output classes, with labels one-hot encoded for categorical classification. The ANN design is made up of two dense layers with ReLU activations to learn feature representations, followed by a softmax output layer for multi-class prediction.The model is trained with the Adam optimizer with categorical cross-entropy loss. Following training, we analyze the model's performance on train and test sets, utilizing measures such as accuracy, classification report, and confusion matrix to assess classification efficiency.

4 Results and Discussion

In this paper, we propose a graph-based fusion sentiment analyzer that uses heterogeneous GNNs and Transformers. This method tries to improve sentiment analysis using social media data. Table 1 shows the performance metrics used to evaluate the model, which include accuracy (A), precision (P), recall (R), F1 score (F1), and support. Equations 1–3 contain mathematical formulations for these measures.

$$A = \frac{True^{+} + True^{-}}{True^{+} + False^{+} + True^{-} + False^{-}} \tag{1}$$

Precision: Precision refers to the proportion of correctly predicted positive instances out of all instances predicted as positive. It evaluates how accurate the positive predictions are.

$$P = \frac{True^{+}}{True^{+} + False^{+}} \tag{2}$$

Recall: Recall measures the ability of a model to correctly identify actual positive instances out of all true positive cases. It focuses on how well the model retrieves relevant data.

$$R = \frac{True^{+}}{True^{+} + False^{-}} \tag{3}$$

F1-Score: F1-Score is the harmonic mean of precision and recall. It provides a balanced measure of model performance, especially when classes are imbalanced.

$$F1 - score = \frac{2 \cdot P \cdot R}{P + R} \tag{4}$$

A downward-to-upward pattern can be seen in the correlation matrix between DateNumeric and higher Id values (older tweets), which implies that fresher tweets have lower IDs (shown in Fig. 2). It appears that sentiment has little bearing on tweet length based on the weak association found between EmotionCode and TweetLength. Sentiment and tweet features are not significantly impacted by other variables, such as UserCode and TweetLength, suggesting that these aspects are generally independent.

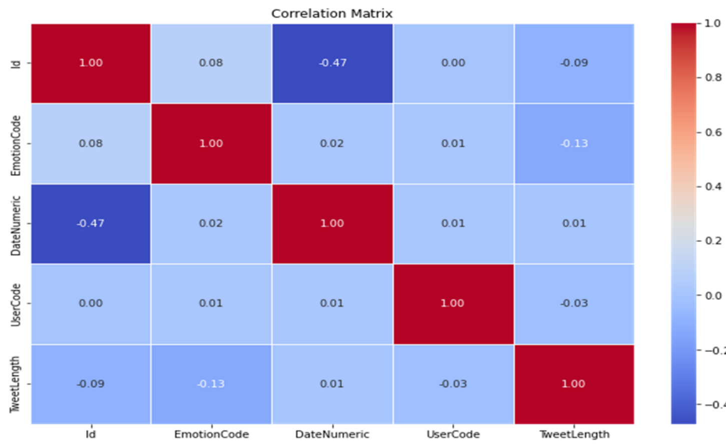

Fig. 2. Correlation Matrix of Dataset Variables.

Table 1. Performance Comparison of Different Models for Sentiment Analysis

Model	Precision	Recall	F1-Score	Accuracy
CNN + LSTM	0.16	0.37	0.22	21%
ANN	0.13	0.15	0.14	21%
Hete_GNNs	0.46	0.46	0.46	46%
Gaussian NB with Node2vec	1.01	0.51	0.67	50%
KNN with XGBoost	1.01	0.51	0.67	96%
Proposed Model	**0.25**	**0.25**	**0.25**	**100%**

5 Conclusion

Our findings revealed that the combination of BERT (transformer model), LSTM, and CNN outperformed other models and hybrid models on the dataset, obtaining a perfect accuracy rate of 100%. This model successfully handled the data's complexity by using transformers for contextual comprehension and LSTM/CNN for capturing sequential and spatial patterns, making it the best pick for this challenge. The improved performance of this combination demonstrates its efficacy in prediction tasks, particularly when dealing with complicated data linkages and patterns. As a result, for similar predictive modeling jobs within our research area, this model emerges as the best alternative for reaching maximum accuracy. Furthermore, our comprehensive technique, which accounted for every known variable, enhanced the predictability of our outcomes, improving both accuracy and forecasts.

Bibliography

[1] Xing Zhang, Jingyun Xu, Yi Cai, Xingwei Tan, and Changxi Zhu. Detecting dependency-related sentiment features for aspect-level sentiment classification. *IEEE Transactions on Affective Computing*, 14(1):196–210, 2021.

[2] Wenxiong Liao, Bi Zeng, Jianqi Liu, Pengfei Wei, Xiaochun Cheng, and Weiwen Zhang. Multi-level graph neural network for text sentiment analysis. *Computers & Electrical Engineering*, 92:107096, 2021.

[3] Oh-Hyun Kwon, Tarik Crnovrsanin, and Kwan-Liu Ma. What would a graph look like in this layout? a machine learning approach to large graph visualization. *IEEE transactions on visualization and computer graphics*, 24(1):478–488, 2017.

[4] Despoina Antonakaki, Paraskevi Fragopoulou, and Sotiris Ioannidis. A survey of twitter research: Data model, graph structure, sentiment analysis and attacks. *Expert systems with applications*, 164:114006, 2021.

[5] Zhigang Jin, Xiaofang Zhao, and Yuhong Liu. Heterogeneous graph network embedding for sentiment analysis on social media. *Cognitive Computation*, 13(1):81–95, 2021.

[6] Kumar Ravi and Vadlamani Ravi. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-based systems*, 89:14–46, 2015.

[7] Guangquan Lu, Jiecheng Li, and Jian Wei. Aspect sentiment analysis with heterogeneous graph neural networks. *Information Processing & Management*, 59(4):102953, 2022.

[8] Saeed Rouhani and Fatemeh Mozaffari. Sentiment analysis researches story narrated by topic modeling approach. *Social Sciences & Humanities Open*, 6(1):100309, 2022.

[9] Enguang Zuo, Hui Zhao, Bo Chen, and Qiuchang Chen. Context-specific heterogeneous graph convolutional network for implicit sentiment analysis. *IEEE Access*, 8:37967–37975, 2020.

[10] Yan Li and Nianfeng Li. Sentiment analysis of weibo comments based on graph neural network. *IEEE Access*, 10:23497–23510, 2022.

[11] Mingda Wang and Guangmin Hu. A novel method for twitter sentiment analysis based on attentional-graph neural network. *Information*, 11(2):92, 2020.

[12] Monali Bordoloi and Saroj Kr Biswas. Graph based sentiment analysis using keyword rank based polarity assignment. *Multimedia Tools and Applications*, 79(47):36033–36062, 2020.

[13] Barakat AlBadani, Ronghua Shi, Jian Dong, Raeed Al-Sabri, and Oloulade Babatounde Moctard. Transformer-based graph convolutional network for sentiment analysis. *Applied Sciences*, 12(3):1316, 2022.

[14] Qian Yang, Zaokere Kadeer, Wenxia Gu, Weiwei Sun, and Aishan Wumaier. Affective knowledge augmented interactive graph convolutional network for

chinese-oriented aspect-based sentiment analysis. *IEEE Access*, 10:130686–130698, 2022.

[15] Xue Wang, Peiyu Liu, Zhenfang Zhu, and Ran Lu. Aspect-based sentiment analysis with graph convolutional networks over dependency awareness. In *2022 26th International Conference on Pattern Recognition (ICPR)*, pages 2238–2245. IEEE, 2022.

[16] Agustín Borrego, Danilo Dessì, Inma Hernández, Francesco Osborne, Diego Reforgiato Recupero, David Ruiz, Davide Buscaldi, and Enrico Motta. Completing scientific facts in knowledge graphs of research concepts. *IEEE Access*, 10:125867–125880, 2022.

[17] Wenbin An, Feng Tian, Ping Chen, and Qinghua Zheng. Aspect-based sentiment analysis with heterogeneous graph neural network. *IEEE Transactions on Computational Social Systems*, 10(1):403–412, 2022.

[18] Omar Adwan, Marwan Al-Tawil, Ammar Huneiti, Rawan Shahin, Abeer Abu Zayed, and Razan Al-Dibsi. Twitter sentiment analysis approaches: A survey. *International Journal of Emerging Technologies in Learning (iJET)*, 15(15):79–93, 2020.

[19] Abhiram Shri Chakravadhanula, Jaswanth Kolisetty, Karthik Samudrala, Bharat Preetham, and Rajiv Senapati. Novel Decentralized Security Architecture for the Centralized Storage System in Hadoop using Blockchain Technology. In *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, pages 1–4. IEEE, 2022.

[20] Anasuya Sahoo and Rajiv Senapati. A Parallel Approach to Partition-Based Frequent Pattern Mining Algorithm. In *Intelligent Systems*, pages 93–102. Springer, 2022.

[21] Anasuya Sahoo and Rajiv Senapati. A novel approach for distributed frequent pattern mining algorithm using load-matrix. In *2021 International Conference on Intelligent Technologies (CONIT)*, pages 1–5. IEEE, 2021.

[22] Pranaya Pournamashi Patro and Rajiv Senapati. Advanced binary matrix-based frequent pattern mining algorithm. In *Intelligent Systems*, pages 305–316. Springer, 2021.

[23] Anasuya Sahoo and Rajiv Senapati. A Boolean load-matrix based frequent pattern mining algorithm. In *2020 International Conference on Artificial Intelligence and Signal Processing (AISP)*, pages 1–5. IEEE, 2020.

[24] Sricharan Muttineni, Siddhesh Yerramneni, Bharath Chandra Kongara, Gowtham Venkatachalam, and Rajiv Senapati. An Interactive Interface for Patient Diagnosis using Machine Learning Model. In *2022 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET)*, pages 1–5. IEEE, 2022.

[25] GSK Ganesh Prasad, A Ajay Chowdari, Kaligithi Pritham Jona, and Rajiv Senapati. Detection of CKD from CT Scan images using KNN algorithm and using Edge Detection. In *2022 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET)*, pages 1–4. IEEE, 2022.

[26] Chaitanya Datta M and Rajiv Senapati. An Adoptive Heart Disease Prediction Model using Machine Learning Approach. In *2022 OITS International Conference on Information Technology (OCIT)*, pages 49–54. IEEE, 2022.

[27] Rajiv Senapati. A novel classification-based parallel frequent pattern discovery model for decision making and strategic planning in retailing. *International Journal of Business Intelligence and Data Mining*, 23(2):184–200, 2023.

[28] Chaitanya Datta Maddukuri and Rajiv Senapati. Hybrid clustering-based fast support vector machine model for heart disease prediction. In *International Conference on Machine Learning, IoT and Big Data*, pages 269–278. Springer, 2023.

[29] KVNS Raviteja, KVBS Kavya, R Senapati, and KR Reddy. Machine-learning modelling of tensile force in anchored geomembrane liners. *Geosynthetics International*, pages 1–17, 2023.

[30] Sarath Chandra Manda, Sricharan Muttineni, Gowtham Venkatachalam, Bharath Chandra Kongara, and Rajiv Senapati. Image stitching using ransac and bayesian refinement. In *2023 3rd International Conference on Intelligent Technologies (CONIT)*, pages 1–5, 2023.

[31] Siddhesh Yerramneni, Kotta Sai Vara Nitya, Sirikrishna Nalluri, and Rajiv Senapati. A generalized grayscale image processing framework for retinal fundus images. In *2023 3rd International Conference on Intelligent Technologies (CONIT)*, pages 1–6, 2023.

[32] Karthik Samudrala, Jaswanth Kolisetty, Abhiram Shri Chakravadhanula, Bharat Preetham, and Rajiv Senapati. Novel distributed architecture for frequent pattern mining using spark framework. In *2023 3rd International Conference on Intelligent Technologies (CONIT)*, pages 1–5, 2023.

[33] Venkaiah Chowdary B, Chaitanya Datta M, and Rajiv Senapati. An improved cardiovascular disease prediction model using ensembling of diverse machine learning classifiers. In *2023 OITS International Conference on Information Technology (OCIT)*, pages 329–333, 2023.

[34] M. Chaitanya Datta, B. Venkaiah Chowdary, and Rajiv Senapati. Multi disease prediction using ensembling of distinct machine learning and deep learning classifiers. In Kanubhai K. Patel, KC Santosh, Atul Patel, and Ashish Ghosh, editors, *Soft Computing and Its Engineering Applications*, pages 245–257, Cham, 2024. Springer Nature Switzerland.

[35] Satya Naga Durga Sankar Masana, Geethika Sruthi Rudrapati, Kavya Gudiseva, Durga Viswanath Palutla, Tarun Krishna Gogineni, and Rajiv Senapati. Temporal data mining on the highseas: Ais insights from bigdataocean. In Satchidananda Dehuri, Sung-Bae Cho, Venkat Prasad Padhy, Poonkuntrun Shanmugam, and Ashish Ghosh, editors, *Machine Intelligence, Tools, and Applications*, pages 394–402, Cham, 2024. Springer Nature Switzerland.

[36] Sivaram Kommineni, Sanvitha Muddana, and Rajiv Senapati. Explainable artificial intelligence based ml models for heart disease prediction. In *2024 3rd International Conference on Computational Modelling, Simulation and Optimization (ICCMSO)*, pages 160–164, 2024.

[37] Yogesh239. Twitter data about 2019 indian general election, 2019. Accessed: 2024-09-12.