Sentiment Analysis on Text based Data of Social Media using Deep Learning

Shruti Sudhakar Chavan; Dr S.S.Sherekar
Department Of Computer Science
Sant Gadge Baba Amravati University
Amravati Maharashtra India

Abstract – Sentiment Analysis Using Deep Learning Is The Process Of Automatically Identifying And Categorizing Emotions Expressed In Textual Data By Leveraging Neutral Network Models That Learn Complex Pattern In Language. This Paper Presents A Robust Framework For Sentiment Classification Of Social Media Text Using A Bidirectional Long Short – Term Memory (BiLSTM) Model . A BiLSTM Based Model Is Use To Analyse Social Media Data, With Preprocessing, Class Balancing, And Feature Extraction Stages. The Model Achieves 95% Accuracy , Effectively Predicting Positive, Negative , And Neutral Sentiments , Demonstrating Its Strength For Real-Time Sentiment Monitoring .

Keywords – Sentiment Analysis , Social Media, Deep Learning , Text Based Data

I.Introduction

Sentimental Analysis Refers To The Task Of Determining The Sentimental Polarity – Positive , Negative , Or Neutral – Expressed Within A Given Block Of Block Of Text. **Deep Learning** , A Subfield Of Machine Learning , Utilizes Artificial Neutral Network Architectures To Model Complex Pattern In Data . In This Study , A Deep Learning – Based Approach Employing A **Bidirectional Long Short – Term Memory (Bilstm)** Network Is Utilized To Train On Labeled Text Data And Perform Sentimental Classification, Effectively Predicting Whether The Sentiment Is Positive , Negative , Or Neutral

A. Sentiment Analysis

Sentiment Analysis Is Nothing But Classifying Opinion Of People Into Negative , Positive, Neutral .Tweet Data Holds Valuable Insights That Can Guide Company Policy Decisions Through Sentiment Analysis. This Involes Using Text Mining And Natural Language Processing To Extract And Classify Sentiments From Tweets. [1] Public Comments And Reviews From Platforms Like Twitter And Facebook Are Key For Analysing Overall Sentiment On Products , Services, Or Global Events . [2]

Twitter Captures Public Emotions Through Short Posts, And Sentiment Analysis Interpret These Opinions On Event Or Policies [3]. Sentiment Polarity In Tweets About The Digital India Mission , Showing Promising And Outlining In Tweets About The Digital India Mission , Showing Promising Results And Outlining Plans To Enhance Accuracy By Addressing Challenges Like Sarcasm , Negation, And Emotions[4] Sentiment-Specific Word Embedding Using Modifies Delta Tf-Idf To Enhance Twitter Sentiment Analysis , With Future Work Focused On Capturing Word Order.[5]

B. Deep Learning

Deep Learning Is A Subset Of Ai Which Use Artificial Neural Networks To Learn From Data. Cnn- Based Deep Learning Model For Classifying Tweet Sentiments Into Positive, Negative, And Neutral , Achieving 63.59% Test Accuracy Using Preprocessed Tweets And Fasttext Word Embedding. [6]

DOI: https://doi.org/10.5281/zenodo.15637232

C. Sentiment Analysis On Text Based Data Of Social Media Using Deep Learning

The Text Is Preprocessed By Cleaning And Tokenizing , Then Analysed Using The Labelled Semeval Tweet Dataset With Sentiments Categorized As Negative, Positive , Neutral.[7]

Different Models Are Use After That [8]. We Can Also Do Political Analysis By Using Sentiment Analysis [9].After Training Model We Can Do Final Sentiment Prediction Weather It Is Positive, Negative, Neutral.[10]

II.Related Work

Hong Soon Goo Et.Al(2017)[22] The Paper Emphasizes The Growing Popularity Of Social Media And Its Potential For Tasks Like Prediction And Sentiment Analysis , Particularly Using Platforms Like Twitter. Due To The Massive Volume Of Unstructured Data , Machine Learning – Specially Deep Learning With Deep Feedforward Neutal Networks – Is Essential For Effective Analysis , Achieving Around 75% Accuracy In Experimental Results.

Qurat Tul Ain Et.Ai(2017) [23] The Paper Highlights The Vast Amount Of User Generated Data On Web – Like Opinions And Emotions On Social Network, Forums , And Blogs – And The Growing Importance Of Sentiments Analysis In Organizing This Unstructured Data . It Notes The Challenge Of Limited Labelled Data In Nlp And Emphasizes The Integration Of Deep Learning Techniques, Such As Deep Neural Network And Cnns, For Effective Sentiment Classification And Addressing Issues Like Cross – Lingual Analysis

Yogesh Chandra Et.Al [24] The Paper Emphasizes The Need For Sentiment Analysis Of User Generated Data On Social Media To Aid Decision Making For Government And Defense Organization , Especially During Critical National Event . By Analyxing Opinions Shared On Platforms Like Twitter And Facebook , Insight Can Be Gained Without Violating Oublic Sentiment. The Study

Employs Machine Learning Classifiers, Polarity-Based Analysis, And Deep Learning Model To Classify Tweets As Positive Or Negative , Aiming To Handle The Diverse Range Of Opinions And Enabling Real Time Sentiment Classification.

Prasanna Kumar Et.Al (2024) [20] This Paper Aim To Improve Sentiment Analysis On Social Media By Developing A Custom Framework(Smsa) That Include A Specially Created Emotion Thesaurus. It Explores How Different Deep Learning Model Perform Under Varied Settings And Highlights A Novel Approach To Building Sentiment Dictionaries . The Study Shows Improved Accuracy And Insights Compared To Earlier Methods, Offering Promising Directions For Future Applications

Dilesh Tanna Et.Al(2020)[26] This Paper Introduces A Custom Social Media Platform With Built-In Sentiments Analysis To Access User Activity. It Rates Sentiments, Generates Reports For Admins , And Helps Identity Issues Like Emotional Distress. The Platform Also Personalizes Content And Supports Sharing Across Other Social Media.

IV. Litreture Survey

Wadzani Et.Al [11] This Give A Detail Overview Of The Use Of Machine Learning And Deep Learning Model In Detecting Depression On Social Media . The Deep Learning Model Turn Out To Be More Realistic And Gives Better Accuracy, In This Paper Cnn,Lstm And Rnn Model Are Used .[11]

Aditya Patil Et.Al(2023)[12] In This Research Paper An Approach Is Used To Enhance Sentiment Analysis On Social Media Data Through Text Analysis An Predictive Modelling [12]

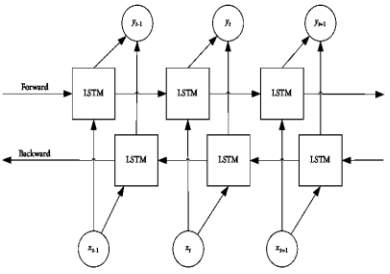
Kottala Sri Yogi Et.Al[13] This Research Explores The Use Of Machine Learning , Particularly Nb,Svm And Deep Learning Model Like Cnns And Rnns, For Sentiment Analysis On Social Media. It Highlights Deep Learning's Superior Accuracy And Ability To

DOI: https://doi.org/10.5281/zenodo.15637232

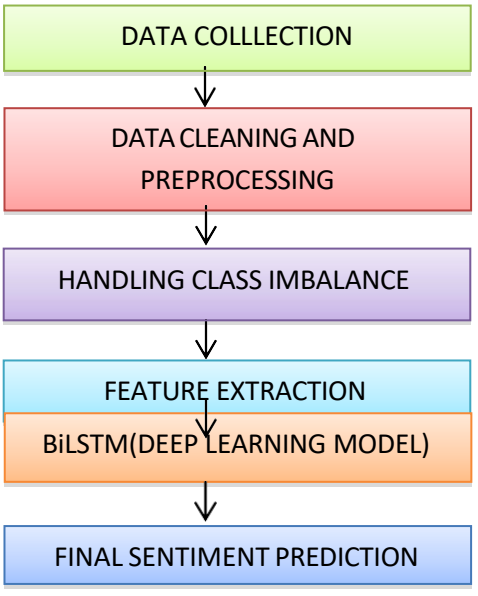
Capture Subtle Emotions In Text , Despite Challenges In Real Time Processing. [13]
Israt Jahan Et.Al (2024) [14] This Research Shows Random Forest Outperforms Other Models In Predicting Heart Attacks Using Patient And Ecg Data . It Emphasizes The Importance Of Advanced Machine Learning In Improving Early Diagnosis And Patient Outcomes. [14]
Youngchaul Et.Al [15] This Study Demonstrated That Incorporating Sentiment Polarities From Twitter Data Into Predictive Model Like Lstm And Gru Improves The Accuracy Of Forecasting Confirmed Covid-19 Cases . It Highlights The Value Of Qualitative Opinion Mining For Epidemic Prediction And Suggests Future Research Should Expand Data Sources And Explore Additional Modelling Approaches. [15]
Sashank Saya [16] This Flask – React App Performs Real – Time Sentiment Analysis On Text And Images Using Machine Learning , Achieving 82% Accuracy With Gradient Boosting And Vgg 16 , Supporting Use Cases Like Social Media Monitoring And Content Moderation . [16]
Mr Swapnil Et.Al [17] This Study Analyses Political Youtube Comments On Bjp And Inc Using The Afinn Lexicon And Various Machine Learning Models, Finding That Svm And Random Forest Offer The Best Performance For Sentiment Analysis , While The Afinn Lexicon And K-Nn Showed Lower Accuracy .
Sunasi Et.Al [18] This Paper Present An Ensemble Method Using Random Forest , K – Nearest Neighbor , Naïve Bayes And Svm To Detect Sarcasm And Its Types In Social Media , Improving Sentiment Analysis Accuracy With Proven Reliability [18].
Nabaa Et.Al [19] This Study Explores Statistical Challenges In Social Media Data Analysis , Focusing On Sentiment Tracking . By Using Advanced Statistical Techniques Like Nlp And Machine Learning , It Addresses Issues Such As Data Sparsity ,High Dimentionality, And Noise , Improving Sentiment Classification Accuracy By 15% And Reducing Noise By 20%. [19].
Mohammad Et.Al [21] This Paper Introduces A Hybrid System Combining Text Mining And Neural Networks For Sentiment Classification , Comparing It With Machine Learning And Deep Learning Algorithm Using A Dataset Of 1 Million Tweets, The System Achieved An Accuracy Rate Of 83.7% Outperforming Standard Supervised Approaches [21]

V. Proposed Method

The Design Of This Proposed Method Is Illustrated In Figure 1. The Collected Database Of Social Media Text Is Labelled With Sentiments Like Positive ,Negative , Neutral . The Data Will Go Through Preprocessing Stages . Then The Next Stage Is Handling Class Imbalance . The Next Stage Is Feature Extraction . The Bilstm Deep Learning Model Is Used In Training Phase . The Final Stage Is Sentiment Prediction



[26] - Bilstm Structure



DOI: https://doi.org/10.5281/zenodo.15637232

Fig 1 : Overview Of Model

A. Data Collection

The Data Set Utilized In This Research Contains Text Used For Differ Social Media Platform Such As Twitter , Instagram ,Telegram , Whatsapp , Telegram ,Quora . We Will Give Sentiment Label To This Text ,Then By Using Deep Learning Model We Will Do Sentiment Classification To Get Our Desired Output .

B. Data Cleaning and Preprocessing

Data Cleaning And Preprocessing Is Essential To Transform Raw Text Into A Format Suitable For Deeo Learning Model. It Starts With **Text Normalization** Which Convert All Text To Lowercase For Consistency And Remove Extra Whitespace , Table And Newline Charecters, Then **Noise Removal** For Eliminate Url, Hashtags, Mentions, Special Character, And Emoji And Remove Numbers Unless They Carry Meaningful Sentiments. **Tokenization** Split Text Into Individual Words(Tokens) For Easier Processing And **Stopword Removal** Filter Out Comman Words (E.G., “The”,”Is”,”In”) That Do Not Contribute Significantly To Sentiment Analysis, Retain Negative Words (E.G , “Not”,”Never”,”Isn’t”) As They Influence Sentiment Meaning. Finally **Stemming** Reduce Words To Their Root Form (E.G “Running”→”Run”) To Standardize Different Forms Of Same Word

C. Handling Class Imbalance

In Sentiment Datasets, Certain Classes (Like Positive) Often Dominate . This Imbalance Can Cause The Model To Favour The Majority Class

Techniques To Address Imbalance :

- Resampling :
  - Oversampling : Increase Samples Of Minority Classes
  - Undersampling : Reduce Samples Of Majority Class
- Data Augmentation
  - Generate Synthetic Data Using Paraphrasing Or Translation Techniques
- Class Weight Adjustment :
  - Assign Higher Weights In Minority Classes During Odel Training To Ensure Balanced Learning

C. Feature Extraction

It Starts With Tokenization Which Convert Text Into Numerical Tokens That Model Can Process , Then There Is Need For Padding In Which Neural Network Required Input Sequences Of Equal Length.

D.Deep Learning Model

The System Uses **Bidirectional Long Short Term Memory (Bilstm)** Network , Ideal For Sequential Data Like Text Because It Captures Both Past And Future Context In Sentences .

Model Architecture

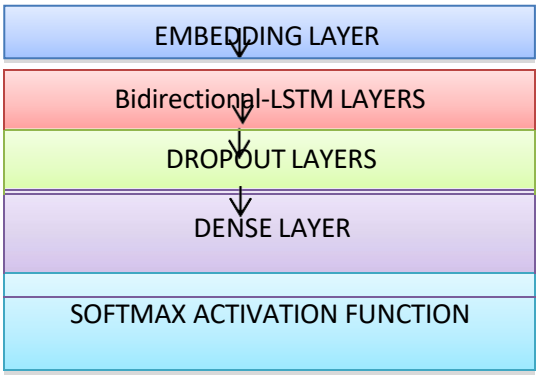


Fig 2. : Deep Learning Model Architecture

- Embedding Layer

Convert Tokens Into Dense Vectors, Representing Semantic Relationships

- Bi-Directional LSTM Layers:

Processes Text In Both Forward And Backward Directions To Capture Context Especially Effective For Understanding Sentences-Level Meanings

- Dropout Layers:

Randomly Deactivates Neurons Training To Prevent Overfitting

- Dense Layers:

Learn Complex Patterns And Relationships In The Data .

Softmax Activation Function :

DOI: https://doi.org/10.5281/zenodo.15637232

Outputs Probabilities For Each Sentiment Class

The Class With The Highest Probability Is Chosen As The Final Prediction

In Model Training , The Loss Function Is Used And Adam Optimizer Combines Benefits Of Adagrad And Rmsprop , Adjusts Learning Rates Dynamically For Faster Convergence. The Regularization Techniques Suh As Early Stopping And Dropout Regularization Is Used In Model Training.

D. Final Sentiment Prediction

Final Sentiments Predictions Reveals Weather The Text Is Positive , Negative Or Neutral.

VI. Result and Discussion

In This Section We Will Discuss The Result Of Our Model That Is Performance Metric

Table 1 : Result Obtained Of Proposed Model

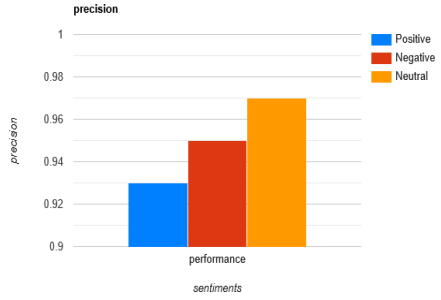


Fig 3 : Precision

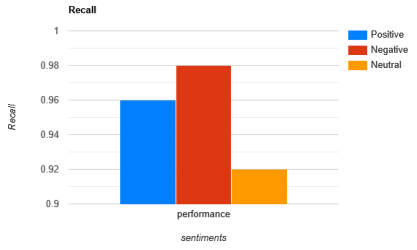


Fig 4 : Recall

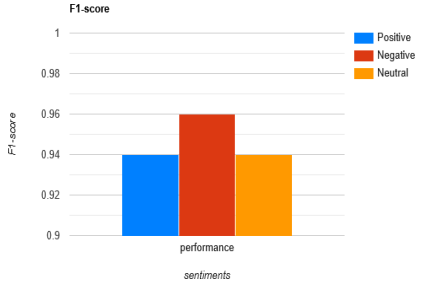
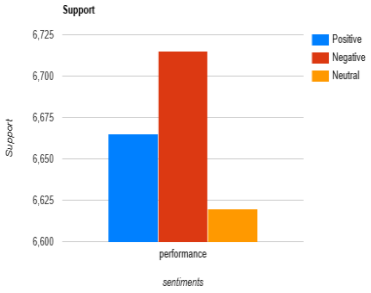


Fig 5 : F1-Score



For The Above Table We Understand That Fig (3) Shows **Precision** Of Positive Sentiments Is 93 % , Negative Is 95 % And

For Neutral Is 97 % , The Overall Precision Is 95 % . The Fig (4) Shows **Recall** Is 96 % For Positive Sentiments , 98 % For Negative Sentiments And 92 % For Neutral Sentiments , Therefore The Overall Recall Is 95.33 %.

Class	Positive	Negative	Neutral	Accuracy
Precision	**0.93**	**0.95**	**0.97**	**0.95**
Recall	**0.96**	**0.98**	**0.92**	**0,95**
F1 – Score	**0.94**	**0.96**	**0.94**	**0.95**
Support	**6665.0**	**6715.0**	**6620.0**	**1.0**

Fig (4) Shows **F1- Score** Tells The Overall Performance Of Model , So For Positive Sentiments It Is 94 % , Negative It Is 96% And Neutral It Is 94% And Therefore Overall Performance Of This Model Is 95 % .

The Overall **Accuracy** Is 95 % .

VII. Conclusion

To Ensure Consistency And Reduce Noise , All Text Was Lowercased , Punctuation Removed And Stopwords Filtered Out – Except For Crucial Negations Like “Not ” And “Never” , Which Significantly Affect Sentiment. To Address Sentiment Imbalance In Social Media Datat, Resampling Techniques Balanced The Dataset Across Positive, Negative, Neutral Classes. The Model Used A Bidirectional LSTM To Capture Context In Both Directions, With Dropout Layers To Prevent Overfitting And A Softmax Layer For Sentiment Classification, Performance Was Evaluated

DOI: https://doi.org/10.5281/zenodo.15637232

Using Precision , Recall , F1 – Score Ad Support , With Early Stopping And Checkpoints Optimizing Training . An Interactive Loop Enabled Real- Time Fig 6 : Support

References

[1] Kusrini,Machamad Mashuri,"Sentiment Analysis In Twitter Using Lexicon Based And Polarity Multiplication",IEEE,2019

[2] Huma Parveen,Prof. Shikha Panday,"Sentiment Analysis On Twitter Data-Set UsingNaive Bayes Algorithm",IEEE,2016

[3] Rupali Patil ,Nishat Gada,Krisha Gala ,"Twitter Data Visualization And Sentiment Analysis Of Article 370",IEEE

[4] Mishra, P., Rajnish, R., & Kumar, P. (2016). *Sentiment Analysis Of Twitter Data: Case Study On Digital India*. Amity University, Lucknow, India.

[5] Rania Othman,1 Youcef Abdelsadek,2 Kamel Chelghoum,2 Imed Kacem,2 Rim Faiz3,"Improving Sentiment Analysis In Twitterusing Sentiment Specific Word Embeddings", 10th IEEE International Conference On Intelligent Data Acquisition And Advanced Computing Systems: Technology And Applications18-21 September, 2019

[6] Pota, M., Esposito, M., Palomino, M. A., & Masala, G. L. (2018). A Subword-Based Deep Learning Approach For Sentiment Analysis Of Political Tweets. *2018 32nd International Conference On Advanced Information Networking And Applications Workshops (WAINA)*, 604–609. Https://Doi.Org/10.1109/WAINA.2018.00147

[7] Rania Othman,1 Youcef Abdelsadek,2 Kamel Chelghoum,2 Imed Kacem,2 Rim Faiz3,"Improving Sentiment Analysis In Twitterusing Sentiment Specific Word Embeddings", 10th IEEE International Conference On Intelligent Data Acquisition And Advanced Computing Systems: Technology And Applications18-21 September, 2019

[8] Parul Sharma ,Teng-Sheng Moh "**Prediction Of Indian Election Using Sentiment Analysis On Hindi Twitter**" Sentiment Predictions, And The Model Achieved Near – Perfect Accuracy With Balanced , Reliable Results.

IEEE International Conference On Big Data (Big Data) 2016

[9] Gaurav Dubey,Silpi Chawla,Kirandeep Kaur,"Social Media Analysis Of Indian Political Diplomat ",7th International Conference On Cloud Computing, Data Science & Engineering – Confluence 2017

[10] Zhaoxia WANG, Chee Seng CHONG, Landy LAN, Yinping YANG, Seng Beng HO And Joo Chuan TONG,"Fine-Grained Sentiment Analysis Of Social Media With Emotion Sensing"Future Technologies Conference 2016

[11] Adyan Marendra Ramadhani,Hong Soon Goo,"Twitter Sentiment Analysis Using Deep Learning Methods",7th International Annual Engineering Seminar (Inaes), Yogyakarta, Indonesia ,2017.

[12] Q. Tul Ain, M. Ali , A Riaz , A Noureen , M Kamran , B Hayat And A. Rehman" Sentiment Analysis Using Deep Learning Techniques: A Review" (IJACSA) International Journal Of Advanced Computer Science And Applications, Vol. 8, No. 6, 2017.

[13] Yogesh Chandra,Antoreep Jana,"Sentiment Analysis Using Machine Learning And Deep Learning",7th International Conference On Computing For Sustainable Global Development (Indiacom),2020

[14] Prasanna Kumar Rangarjan,Bharathi MohanGurusamy,Gayathri Muthurasu,Rithani Mohan,GundalaPallavi,Sulochana Vijayakumar,Ali Altalbe,"The Social Media Sentiment Analysis Framework: Deep Learning For Sentiment Analysis On Social Media",International Journal Of Electrical And Computer Engineering (IJECE) Vol. 14, No. 3, June 2024

[15] Tanna, D., Dudhane, M., Sardar, A., Deshpande, K., & Deshmukh, N. (2020). Sentiment Analysis On Social Media For Emotion Classification. In *Proceedings Of The International Conference On Intelligent*

DOI: https://doi.org/10.5281/zenodo.15637232

Computing And Control Systems (ICICCS 2020) (Pp. 233–237). IEEE. Https://Doi.Org/10.1109/ICICCS48265.2020.9161060

[16] Wadzani Aduwamai Gadzama , Danlami Gabi , Musa Sule Argungu , Hassan Umar Suru,"The Use Of Machine Learning And Deep Learning Models In Detecting Depression On Social Media: A Systematic Literature Review", Www.Sciencedirect.Com/Journal/Personalized-Medicine-In-Psychiatry,2024

[17] Aditya Patil, Swati Patil, Prajakta Patil,Dr.Vaishali Patila, Sanket Lodhac,Dr.Sudhir Chitnisb,Dr. Arun Patilb,"An Approach To Enhance Sentiment Analysis On Social Media Data Through Text Analytics And Predictive Modeling",Research Gate , Empirical Economics Letters, 22 (2): (August 2023)

[18] Kottala Sri Yogi,Dankan Gowda V,Galiveeti Poornima,KDV Prasad,Srinivas.D,P Vishnu Prasanth,"Comparative Analysis Of Machine Learning Techniques For Detecting Sentiments In Social Media",Research Gate ,May 2024.

[19] Israt Jahan, Md Nakibul Islam , Md Mahadi Hasan And Md Rafiuddin Siddiky, "Comparative Analysis Of Machine Learning Algorithms For Sentiment Classification In Social Media Text", World Journal Of Advanced Research And Reviews, 2024.

[20] Song Y, Yoon B. Prediction Of Infectious Diseases Using Sentiment Analysis On Social Media Data. Plos ONE. 2024;19(9):E0309842. Https://Doi.Org/10.1371/Journal.Pone.0309842

[21] Sashank Saya,Shreelaxmi K Malawade ,Smaranya Vijaya Krishna ,Vaishnavi K S ,Shylaja B ,"Sentiment Analysis For Social Media Presence",Research Square ,2024.

[22] **Goje, S. P., & Patil, R. H. (2024).** Sentiment Analysis Of Political Parties On Social Media: A Machine Learning And Lexicon-Based Approach. *International Journal Of Scientific Research In Computer Science, Engineering And Information Technology*, 10(5), 233–246. Https://Doi.Org/10.32628/CSEIT24105103

[23] Sunusi Kabir Alaramma1, Adamu Adamu Habu2, Badamasi Imamu Ya'u1, And Abdullahi Gamsha Madaki,"Sentiment Analysis Of Sarcasm Detection In Social Media",Gadau Journal Pure Alli Sci, 2(1): 76-82 (2023)

[24] Nabaa Muhammad Diaa1, Saba Sabah Ahmed2, Hayder Mahmood Salman3, Wafaa Adnan Sajid4,"Statistical Challenges In Social Media Data Analysis Sentiment Tracking And Beyond",Journal Of Ecohumanism 2024

[25] Mohammed H. Abd El-Jawad,Rania Hodhod,Yasser M. K. Omar,"Sentiment Analysis Of Social Media Networksusing Machine Learning"

[26] Wang, H., Wang, J., Cao, L., Li, Y., Sun, Q., & Wang, J. (2021). A Stock Closing Price Prediction Model Based On CNN-Bislstm. *Complexity*, 2021, 5360828. Https://Doi.Org/10.1155/2021/5360828

DOI: https://doi.org/10.5281/zenodo.15637232