NAME: *Smruti Rajesh Sonekar* **COLLEGE** :*Sardar Patel Institute of Technology*
Super40 Batch

Topic: Twitter Sentimental Analysis

Table: Literature Survey

Sr. no	***Title***	***Author***	***Year***	***Findings/Remarks***
1.	Aspect-level Sentiment Analysis for Social Media Data in the Political Domain using Hierarchical Attention and Position Embeddings	Renny Pradina Kusumawardani ,Muhammad Wildan Maulidani	2020	• In this paper,over 3000 tweets of Indonesia Political and it consist of higher number of linguistic data . • It has used LSTM model , which is Long Short Term Memory which is better verison of RNN it remembers patterns of the sequence of data and it has performed better than LRU and RNN. • They have also use HAPN architecture which uses hierarchical attention and its widely used in sentiment analysis , if sentence is positive , neutral , negative. • This paper focused on understanding the word embeddings (special , linguistic words) to understand informal language , and for better results one should explore BERT model as well.
2	Robust Sentiment Detection on Twitter from Biased and Noisy Data Luciano Barbosa AT&T Labs			• In this paper , a method to automatically detect sentiments on Twitter messages by exploring the characteristics of tweets and the meta-information of the words that compose them. • The method leverages sources of noisy labels as training data provided by a few sentiment detection websites over Twitter data. • They have tried to map the word withs part of speech and used a lot of classification algorithm to generalize it. • They have used 2 step approach for sentiment detection in which first they check if the tweet which was made was opinion , belief, categorizing them into subjective and objective and then used polarity detection which means sentiment analysis of the tweet if its neutral, positive, negative.

3	Study on Machine learning based Social Media and Sentiment analysis for medical data applications	R. Meena1 , Dr. V. Thulasi Bai2 1Assistant Professor, Prathyusha Engineering College, Chennai, India meenarajeswaran@gmail.com 2 Professor, KCG College of Technology , Chennai, India	2019	• This study explores the use of social media data for cancer queries on Google Trends, Twitter, and online forums, using sentiment analysis. • ADRMine uses sentiment analysis features to test tweets in various domains using sentiment polarity and sentiment load. Cancer treatment data can be analyzed using Twitter data, while online health forums can reveal adverse health effects of banned drug usage. Content analysis can reveal symptoms of mental health using depression-related chats. • It also highlighted that social media can be integrating with healthcare management to create awareness about drugs, smoking, cancer , mental health. • Sentiment analysis was performed on the extracted tweets, categorizing them into Positive, Neutral, and Negative categories. The results showed that positive and neutral tweets were more influential than negative ones. • The findings suggest that sentiment analysis and text-based mining can be used as a tool for understanding user perception and public health interventions, with specific methodologies applied to the data to achieve desired results.
4	Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis	Lei Zhang1,2, Riddhiman Ghosh1, Mohamed Dekhil1, Meichun Hsu1, Bing Liu2	January 2011	• The highlight of this research paper is that they have used Lexicon based approach in which experts and researches create a list of words and phrases along with their sentiment labels (positive, negative, or neutral). This lexicon becomes a reference point for the analysis. • Now every word in tweets is compared against the lexicon and If a match is found, the sentiment label of the matched word is assigned to that part of the text. • The proposed method is more desirable for practical applications due to its nature of no manual involvement and its ability to automatically adapt to new fashions in language, neologisms, and trends. Experimental results show that the proposed method dramatically improves the recall and F-score, outperforming state-of-the-art baselines.

5	TagNet: Toward Tag-based Sentiment Analysis of Large Social Media Data	Yang Chen	28 May 2018	• In this paper,The integration of hashtags and replies in social media messages allows for the measurement of continuous sentiments towards fast-evolving subjects, such as social events, topics, and conversations. • TagNet is a novel visualization approach designed for tag-based sentiment analysis of large social media data. • Tag selection is based on a TFIDF weighting scheme, with the goal being to display the most representative tags over a specified time range. The topology of tags inherits the generality of traditional tag clouds to represent their relationships, more specifically temporal relations and co-occurrences. • Central Tags: Imagine a central area where common topics (hashtags, keywords, etc.) are displayed as tags. These tags act like a reference point for comparing sentiment across different datasets. • Sentiment Glyphs: On opposite sides of each central tag, TagNet displays visual elements (glyphs) that represent the overall sentiment (positive or negative) associated with that topic across the datasets. • Trend Glyphs: These are dynamic visuals that depict how sentiment towards a particular topic (represented by the central tag) changes over time within each dataset. They essentially show the "sentiment journey" for each topic.
6	Sentiment Analysis of Twitter Data: A Survey of Techniques	Vishal A. Kharde Department of Computer Engg, Pune Institute of Computer Technology ,Pune University of Pune (India)	April 2016	• In this paper, twitter tweets are analyzed using different algorithms and approach . • First the dataset is preprocessed in which it includes removal of urls, removing stop words, punctuations, and then feature extractions is used and Naives Bayes, Support Vector Machine,have highest accuracy. • Many approaches were tested , lexicon based , corpus based approach,cross domain,Crosslingual and lexion based approach was more effective. • The results show that SVM outperforms Naive Bayes and Maximum Entropy for sentiment analysis of Twitter data.

				• When using unigrams and bigrams (considering one or two words together) and removing stop words, SVM achieved an accuracy of 77.73%.
7	TRABSA: Interpretable Sentiment Analysis of Tweets using Attention-based BiLSTM and Twitter-RoBERTa	Md Abrar Jahin, Md Sakib Hossain Shovon, and M. F. Mridha	30 Mar 2024	• In this paper, researchers have used new method TRABSA which is known as Transformer-based Robusta Attention BLISTm for sentimental analysis • I t first preprocesses the data and then it uses Pre trained transformer model which analyzes complex relationship . • TRABSA also employs BiLSTMs (Bidirectional Long Short-Term Memory networks). These networks can analyze text in both directions (forward and backward) to capture the context and sentiment flow within a sentence. • To train the model, TRABSA uses sentiment lexicons - dictionaries that associate words with positive, negative, or neutral sentiment. The paper compares different approaches for labeling sentiment in the training data and selects the most optimal one.
8	Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach	Yuxing Qi1 and Zahratu Shabrina 2,3	Published online 09 February 2023	• Focuses on analyzing Tweets during the third lockdown in England. • Three main methods used: lexicon-based, machine-learning-based approaches, and hybrid techniques. • Lexicon-based approach uses VADER (Valence Aware Dictionary and Sentiment Reasoner) and Python TextBLob library. • SentiWordNet-based approach uses new evaluation metrics. • Machine learning approach uses multinomial naive bayes theorem, Support Vector Classification (SVC), Random forest, and Hyperparameters optimization. • Lexcoin-based approach yields better results, while SVC has the highest accuracy in machine learning approach.
9	Twitter Vigilance: a Multi-User platform for Cross-Domain Twitter Data Analytics, NLP and Sentiment Analysis	Daniele Cenni, Paolo Nesi, Gianni Pantaleo, Imad Zaza	28 June 2018	• Multipurpose tool for Twitter data collection, analysis, and monitoring. • Uses Twitter Search API for flexibility and filtering query results. • Collects Twitter data based on keywords, user IDs, or locations.

				• Provides real-time and historical data analysis. • Offers various analytics including tweet volume metrics, sentiment analysis, Natural Language Processing (NLP), user network analysis, and geolocation-based data analysis. • Supports creation of personalized dashboards for data visualization. • Enables efficient search with filtering options for trend exploration. • Allows exporting processed data for independent analysis. • Features include high recall, multi-user platform, real-time analysis, advanced features, and data visualization.
10	AffinityFinder: A System for Deriving Hidden Affinity Relationships on Twitter Utilizing Sentiment Analysis	Abdelmounaam Rezgui Daniel Fahey Ian Smith Department of Computer Science & Engineering New Mexico Tech	18 October 2016	• System collects and analyzes tweets to derive relationship scores reflecting affinity degrees. • Aims to make smart inferences about relationships using nontraditional information. • Data includes tweets containing mentions of other users. • Language analytics classifies messages to reveal potential relationship types. • Affinity scores are derived between user pairs based on mentions in tweets. • Workflow includes data collection, Bayesian analysis, and affinity graph building. • Utilizes Twitter REST API and Mongo DB, Bayes theorem for training, and TextBlob python library for sentiment anal