A Systematic Literature Review of Emotion Recognition in Conversation Using Natural Language Processing

Indresh Yadav[1], Kamlesh Raghuwanshi[2], Dr. Surabhi Karsoliya[3]

Abstract-*Text-based emotion detection, especially in multi-speaker conversations, is now a very important area of research in Natural Language Processing (NLP). It has many useful uses in fields like customer service, healthcare, and social media analysis. This systematic literature review compiles the development of methodologies in this domain, mapping the transition from fundamental techniques to the contemporary state of the art. We start by looking at early lexicon-based and rule-based methods. These were fast to compute, but they had trouble with the subtleties of conversational language. The review then talks about how machine learning models like Support Vector Machines and Naive Bayes classifiers have made things better. These models used data-driven pattern recognition, but they were often limited by features that were made by hand. Deep learning is at the heart of modern emotion detection. Sequential models like LSTMs and transformer-based architectures like BERT are setting new performance standards by capturing temporal dependencies and bidirectional context. We also talk about hybrid and ensemble models that use more than one method to make predictions more accurate. This review also shows how important multimodal and contextual approaches are becoming. These approaches combine text with audio-visual cues and model how conversations work. Even though a lot of progress has been made, problems like dataset dependency, multilingual support, and ethical biases still exist. This paper concludes by delineating these challenges and charting future research trajectories, underscoring the necessity for more resilient, multimodal, and ethically-informed systems to enhance human-computer interaction.*

Keywords: *Emotion Detection, Natural Language Processing (NLP), Emotion Recognition in Conversation (ERC), Deep Learning, Machine Learning, Sentiment Analysis, Multimodal Emotion Recognition.*

I. Introduction

Natural Language Processing (NLP) for detecting emotions in conversations is becoming an increasingly important area of research with far-reaching effects in many fields, including healthcare, customer service, education, and social media analysis. Since conversations often happen with more than one person talking at once, it's important to understand the emotional subtleties and dynamics between people in order to make good emotion detection systems. This literature review combines recent progress in this field, pointing out different methods, their effects, and gaps in knowledge that need more research. People often think of emotion as the mental state that goes along with thoughts, feelings, and actions. Stoics, including Cicero, classified emotions into four categories: metus (fear), aegritudo (pain), libido (lust), and laetitia (pleasure). Subsequently, evolutionary theories of emotion were established in the late 19th century by Darwin and Prodger [1]. He posited that emotions evolved via natural selection, thereby possessing cross-cultural universality. Plutchik [2] recently classified emotions into eight fundamental categories, represented by the wheel of emotions. Ekman [3] also made the case for a link between emotion and facial expression. Natural language frequently reflects an individual's emotional state.

Consequently, emotion recognition has gained prominence in the domain of NLP [4], [5], owing to its extensive applications in opinion mining, recommender systems, healthcare, and various other fields. Strapparava and Mihalcea [6] tackled the challenge of identifying emotions in news headlines. To solve the problem of recognising emotions in text, several emotion lexicons [7], [8] have been made. The NLP community has only recently started to pay attention to emotion recognition in conversation (ERC) [9]–[12] because there is more public conversational data available. You can use ERC to look at conversations that happen on social media. It can also help with real-time conversation analysis, which can be very useful in legal trials, interviews, e-health services, and more. In contrast to vanilla emotion recognition of sentences or utterances, ERC necessitates the contextual modelling of each utterance. This context can be ascribed to the preceding utterances and is contingent upon the temporal sequence of those utterances. In contrast to the recently published studies on ERC [10]–[12], both lexicon-based [8], [13], [14] and contemporary deep learning-based [4], [5] emotion recognition methodologies are ineffective on ERC datasets, as these studies overlook conversation-specific factors such as presence.

Because technology has changed so much and the Internet has grown so much, digital social platforms have huge amounts of digitised data in the form of text, images, videos, and more. You can find digitised text on digital social media in blogs, news articles, customer reviews of different products and services, discussion forums, review/recommendation systems, conversational agents, and more. These online social networks also let people share their thoughts and feelings. All of these social media sites help businesses figure out what people are feeling and thinking, look at social trends, follow up on customer feedback to improve their business strategy, help customers make decisions, and more. Consequently, text-based emotion detection has evolved into a prominent field of study.
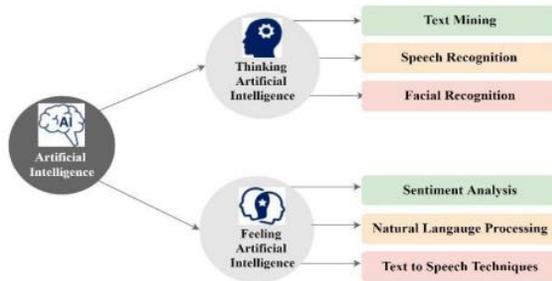

Figure 1 Domains of Artificial Intelligence

Text-based emotion detection has been utilized in numerous practical applications; [15] the author delineates an innovative system for emotion analytics that leverages social media content to facilitate information gathering, situational awareness, and administrative decision-making. [16] Looked into using automatic sentiment analysis to figure out whether a review of a product or service is positive or negative. [17] Talked about how to classify emotions in news articles. [15] Used a system that takes user reviews into account when making e-commerce recommendations. [18] Suggest a way for social robots to figure out how people feel. [19] Looked into conversational agents in health care. [20] Created a system for users to review films to help them make decisions. Finding emotions in a lot of digital text on social media sites is the key to getting more research done in the field. This will help people and machines work together better and make better decisions in many areas. TBED encompasses multidisciplinary aspects, thereby being applicable to various domains, including psychology, sociology, human-computer interaction (HCI), and data mining. This Systematic Literature Review (SLR) not only gives an overview of the current state of research on text-based emotion detection, but it also highlights the different digital social platforms that can be used for this purpose, the techniques that have been used, the datasets that have been used in the research, and the future directions of the field.

II. Related Work

In the domain of emotion detection (ED) systems, the portrayal of emotions via diverse models constitutes a fundamental component, influencing the conceptualisation and analysis of emotions. These models seek to encapsulate the complex nature of emotions, which manifest in various states, thereby requiring distinct differentiation among them. Establishing a suitable emotion model is an essential preliminary step in any endeavour related to ED, as it guides the design and functionality of the system [21]. The literature identifies two principal frameworks for representing emotions: discrete emotion models (DEMs) and dimensional emotion models (DiEMs), each providing unique methodologies for categorising and comprehending emotional states.

Discrete emotion models concentrate on categorising emotions into distinct groups, highlighting their uniqueness. Paul Ekman's model is one of the most well-known frameworks in this area. It says that there are six basic emotions—happiness, sadness, anger, disgust, surprise, and fear—that come from different neural systems [22][23][24][25]. Ekman's theory asserts that fundamental emotions emerge from an individual's interpretation of a situation and can amalgamate to create more intricate emotions, including guilt, shame, or pride. Robert Plutchik's model builds on this idea by adding acceptance/trust and anticipation to Ekman's six main emotions, making a total of eight [26]. Plutchik categorises these emotions into opposing pairs: joy versus sadness, trust versus disgust, anger versus fear, and surprise versus anticipation. This implies that combinations of these primary emotions produce more intricate emotional states. He also stresses that the intensity of emotions changes depending on how the person experiences them. The Ortony, Clore, and Collins (OCC) model, on the other hand, questions the idea of "basic emotions" while agreeing that emotions come from how we see events and change in strength [27]. The OCC model significantly broadens the spectrum by delineating 22 distinct emotions, such as relief, envy, reproach, and admiration, thereby offering a more comprehensive representation than the frameworks proposed by Ekman or Plutchik. The selection of DEM is contingent upon the researcher's aims, with the OCC model providing a more exhaustive classification owing to its broad spectrum of emotional categories.

needs. Each has its own strengths when it comes to showing the complicated world of human emotions.
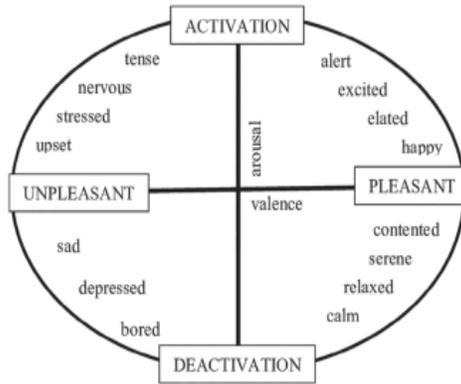

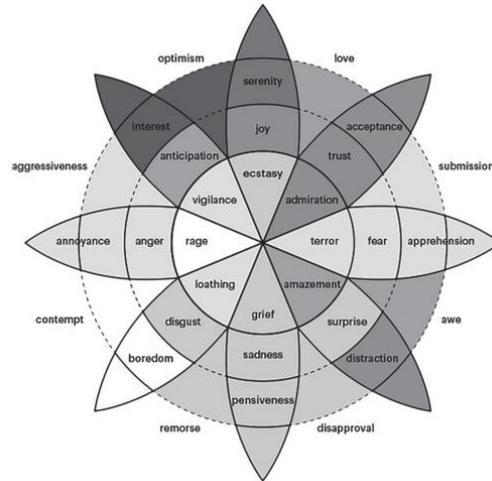Figure 2 Russell's circumplex model of affects

Dimensional emotion models (DiEMs), on the other hand, suggest that emotions are connected to each other rather than being separate, and they place them in a spatial context [21][28]. These models, encompassing unidimensional and multidimensional frameworks, generally represent the essential behavioral states associated with positive and negative experiences [29]. Two-dimensional and three-dimensional frameworks are examples of multidimensional models that are used more often. James Russell's circumplex of affect is a well-known two-dimensional model that shows how emotions are connected by placing them along two axes: Arousal (activation versus deactivation) and Valence (pleasantness versus unpleasantness) [30]. In the same way, Plutchik's two-dimensional wheel of emotions uses Valence and Arousal to show emotions in concentric circles, with the main emotions in the center and their combinations spreading out from there. This shows how they relate to each other. Russell and Mehrabian enhance this framework by introducing a three-dimensional model that includes Valence, Arousal, and Dominance, with Dominance indicating the level of control an individual has over their emotions [31]. This extra dimension makes the model better by taking into account the experiential side of emotional regulation.

The literature indicates that discrete models are extensively utilized for emotion classification tasks due to their simplicity [32]. But they can't capture the full range of emotional subtleties and intensities as well as dimensional models can. DiEMs are especially useful for projects that need a deep understanding of how emotions are related to each other because they can show how emotions are similar and different. In the end, the choice between DEMs and DiEMs depends on what the ED system

Figure 3 Plutchik's wheel of emotions

Table 1 Emotion Detection in Conversation or Using NLP

Authors	Year	Key Methods/Contributions	Datasets Used
Canales L, Martínez-Barco P [32]	2014	A thorough review of lexicon-based and machine learning methods for classifying emotions in text, emphasizing difficulties such as negation management and sarcasm detection..	ISEAR, various textual corpora
Poria, Soujanya, et al. [33]	2019	Introduces ERC as a subfield of NLP and talks about context, speaker dependencies, and multimodal fusion. It also looks at early graph-based and RNN models for conversational emotions..	IEMOCAP, MELD, EmoryNLP
Acheampong, Felix A., et al. [34]	2020	Surveys rule-based, lexicon, and hybrid DL methods; focuses on adding emojis and using them on social	Twitter, Facebook posts

		media; finds gaps in multilingual support.	
Hazarika, Deepanway, et al. [35]	2020	Suggests cross-modal attention for combining text and audio in conversations; gets the best results on multimodal ERC by keeping features that are unique to each modality..	CMU-MOSEI, IEMOCAP
Kim, Evgeny, and Roman Klinger [36]	2020	A systematic review of textual and nonverbal cues in dialogues, concentrating on sequence models such as LSTM for turn-taking emotions, and addressing ethical biases..	DailyDialog, MELD
Peng, Sancheng, et al. [37]	2022	It sorts DL methods for monolingual and cross-lingual text, looks at CNN-RNN hybrids for conversation flow, and focuses on emoji-based models..	Twitter, Sina Weibo
Machová, Kristína, et al. [38]	2023	Compares ML models (SVM, NB) with DL (CNN-LSTM) for text emotion; hybrid neural networks work best; this applies to chatbot conversations.	Custom social media texts
Flor Miriam, et al. [39]	2024	It looks at more than 50 datasets and models and focuses on moral issues like bias in ERC. It also makes plans for integrating personality types in a multimodal and dynamic way.	IEMOCAP, MELD, M3ED

III. Emotion Detection in Conversations Using NLP

Emotion detection in conversations is a complicated and changing area of Natural Language Processing (NLP) that tries to find and sort emotional states that are shown in spoken or written conversations. This job is very important for things like human-computer interaction, sentiment analysis, keeping an eye on mental health, and automating customer service. Conversations are always changing and depend on the situation, and emotional cues are often very subtle, which makes things very difficult. Researchers have created a variety of methods over time, moving from basic lexicon-based methods to more advanced machine learning, deep learning, and multimodal methods. This part looks at how these methods have changed over time, pointing out their pros, cons, and contributions to the field.

a. Early Approaches: Lexicon-Based and Rule-Based Methods

Early attempts to detect emotions relied a lot on lexicon-based and rule-based methods. Lexicon-based techniques, including those utilizing WordNet-Affect [40], assigned words to established emotional classifications (e.g., joy, anger, sadness). These methods were easy to use and didn't take up too much computer time, so they were good for early uses. But they had trouble with emotions that depended on the situation, sarcasm, and subtle expressions because they relied on fixed word-emotion links. Rule-based systems, frequently based on psychological frameworks such as Russell's circumplex model of affect [41], classified emotions according to dimensions of valence (positive-negative) and arousal (high-low). These methods were the first steps toward finding emotions, but their inability to deal with how conversations change led to the creation of more advanced methods.

b. Machine Learning Techniques

The introduction of machine learning marked a significant advancement in emotion detection, enabling models to learn patterns from data rather than relying on predefined rules.

- **Support Vector Machines (SVMs)**: Because they can work with high-dimensional data and do well with little training data, SVMs have been used a lot for text classification tasks, such as emotion detection [42]. SVMs effectively differentiate between emotional categories by mapping text features (e.g., n-grams, bag-of-words) into a high-dimensional space.

Because they are strong in small datasets, they work well with early conversational datasets that don't have many annotations, like the DailyDialog dataset [43].

- **Naive Bayes Classifiers**: Individuals like these probabilistic models because they are simple and work well for detecting emotions in text [44]. Naive Bayes classifiers assume that features are independent. This is a simplification that makes calculations faster and works well for tasks like classifying emotions and sentiments. But because they depend on hand-crafted features (like part-of-speech tags and lexical features), they can't pick up on the subtle emotional changes that happen in conversations.

Machine learning methods built on lexicon-based ones by adding data-driven feature extraction. However, their performance was limited by the quality and variety of features and the number of labeled datasets that were available.

c. Deep Learning Techniques

Deep learning has changed the way we find emotions by letting models find complicated, non-linear patterns in conversational data, especially in situations where there is a lot of context.

- **Neural Networks**: Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are examples of sequential neural networks that are good at processing data that comes in a sequence, like conversations [45]. These models can capture how dialogue changes over time, which makes them good for understanding subtle emotional expressions that change over several turns. For instance, LSTMs have been employed to model emotional transitions in datasets such as IEMOCAP [46], which comprises dyadic conversational data.

- **Transformers**: Since their attention mechanisms capture contextual relationships in text [47], transformer-based models like BERT and RoBERTa have become the best way to find emotions. These models are very good at understanding the two-way context of conversations, which lets them pick up on small emotional cues. Fine-tuned models like EmoBERT have gotten the best results on conversational datasets like MELD, which has multimodal dialogue data [48].

Deep learning models have made a big difference in how well they can detect emotions, but they are hard to use in places with limited resources because they are very complicated and need a lot of labeled data.

d. Hybrid and Ensemble Models

Researchers have created hybrid and ensemble models that combine the best parts of different machine learning and deep learning methods to get around the problems with each one.

- **Hybrid Models**: These models combine machine learning and deep learning parts to use the speed of traditional methods and the power of neural networks to express ideas [45]. For instance, a hybrid model could use word embeddings (like GloVe or BERT embeddings) as input to a dense neural network layer. This would combine strong classification with contextual understanding. These kinds of models have been very good at finding emotions in complicated conversation datasets, striking a good balance between accuracy and speed.

- **Ensemble Techniques**: Ensemble methods merge predictions from several models to boost performance, frequently attaining superior precision and recall compared to individual models [44]. For example, an ensemble of SVMs, Naive Bayes classifiers, and LSTMs can do better than any one model by using their strengths together. Ensemble methods work especially well when dealing with the fact that emotional expressions change depending on the conversation..

e. Multimodal and Contextual Approaches

Recent advancements have concentrated on integrating supplementary data sources and contextual information to improve emotion detection.

- **Multimodal Integration**: Multimodal approaches combine text with other forms of communication, like speech and visual cues, to give a deeper emotional context [49]. The MELD dataset [48] has text, audio, and video data, which lets models look at tone, facial expressions, and text content all at once. Tensor fusion networks and other multimodal fusion techniques have outperformed unimodal models by capturing interactions between different types of data.

- **Contextual Analysis**: Conversations are naturally sequential, and emotions

frequently rely on preceding statements and conversational history. As discussed in [50], contextual analysis models the effect of past statements and two-way information flow to make it easier to find emotions. DialogueRNN and other techniques explicitly model speaker states and turn-taking dynamics to deal with problems like changes in emotion and speaker dependency.

IV. Challenges in Text-Based Emotion Detection

Even though there have been big improvements, the field of text-based emotion detection still has a lot of problems that make it hard for current models to be accurate and useful. These problems come from the fact that human language is naturally complicated and the fact that current methods have technical problems.

One of the main problems is that emotions depend on the situation. Early lexicon-based and rule-based methods, although computationally efficient, frequently failed due to their dependence on static word-emotion associations. They had a hard time understanding subtle expressions, sarcasm, and emotions that change a lot depending on the conversation. Even though machine learning and deep learning models have made it easier to model context, it is still hard to get the full dynamics of a conversation right. Emotions in a conversation aren't fixed; they change based on what was said before, who is speaking, and how the conversation is going overall. Many models still have trouble fully understanding this. Another big problem is that we need a lot of high-quality labeled datasets. Deep learning models, especially transformer-based architectures like BERT, have changed the field by finding complicated patterns in text. But their performance really depends on having a lot of annotated data. Making these datasets takes a lot of resources, and the fact that there isn't much well-annotated conversational data can make it hard to train models and make them work well. This reliance makes it very hard to use advanced models in places where resources are limited.

Additionally, numerous current models encounter challenges in managing multilingual and multimodal complexities. Most studies have concentrated on monolingual text, resulting in a deficiency in multilingual support for emotion detection systems. Also, text by itself doesn't always show how someone is feeling completely. Tone of voice and facial expressions are two examples of nonverbal cues that people use in conversations. Multimodal approaches that combine text, audio, and visual data have shown promise, but combining these different types of data to capture interactions between them is a difficult problem that needs advanced methods.

Lastly, ethical issues, like bias, are becoming a bigger problem for the field. Models trained on biased datasets can reinforce and magnify societal biases, resulting in unjust or erroneous emotional analyses. Ensuring fairness and addressing ethical concerns in Emotion Recognition in Conversation (ERC) is an essential domain that necessitates additional research and the creation of resilient, impartial models. The performance of machine learning models is also limited by the quality and variety of the features they use. Hand-crafted features may not be able to capture the full range of emotional expression.

V. Conclusion

This review has looked at how text-based emotion detection has changed over time. It started with simple lexicon-based methods and has quickly moved on to the advanced deep learning models that are now the best in the field. Initial methodologies, such as lexicon-based and rule-based systems, established a fundamental framework yet were constrained by their inadequacy in addressing context, sarcasm, and the fluid characteristics of human language. The later addition of machine learning methods, like Support Vector Machines and Naive Bayes classifiers, was a big step forward because they made it possible to recognize patterns based on data, even though hand-crafted features often limited performance. Deep learning, especially transformer-based architectures like BERT, has changed the field in a big way. They can capture complex, non-linear patterns and contextual nuances in conversational data like no other technology can. These models, along with sequential networks like LSTMs, have done very well at figuring out the temporal dependencies and bidirectional context that are necessary for accurate Emotion Recognition in Conversation (ERC). The creation of hybrid and ensemble models has also shown that mixing the best parts of different methods can make systems that are more reliable and accurate, balancing speed with expressive power. Even with these improvements, there are still big problems to solve. The need for big, high-quality annotated datasets, the difficulty of supporting multiple languages, and the need for good multimodal integration of text, audio, and visual cues are all ongoing problems. As we said before, a key area that needs more innovation is accurately modeling the context of a conversation, including

speaker dependencies and changes in emotion. As these technologies are used more and more, ethical issues like reducing bias in training data and making sure things are fair are becoming more and more important.

Future research is poised to look into a number of promising areas. The combination of multimodal data is a key area of research. Techniques that can effectively combine different modalities to capture cross-modal interactions have the potential to greatly improve accuracy. Another important area of research is creating models that can change based on the personalities of the people they are talking to and how the conversation is going. As emotion detection systems become more common in real-world situations, like healthcare, customer service, social robotics, and mental health analysis, it will be very important to deal with the problems that already exist and follow these new paths. The ongoing improvement of these technologies holds the potential to not only improve the way people and computers interact, but also to give us a better understanding of the complicated world of human emotions.

Reference

[1] C. Darwin and P. Prodger, The Expression of the Emotions in Man and Animals. Oxford, U.K.: Oxford Univ. Press, 1998.

[2] R. Plutchik, ''A psychoevolutionary theory of emotions,'' Social Sci. Inf., vol. 21, nos. 4–5, pp. 529–553, 1982.

[3] P. Ekman, ''Facial expression and emotion,'' Amer. Psychol., vol. 48, no. 4, p. 384, 1993.

[4] B. Kratzwald, S. Ilic, M. Kraus, S. Feuerriegel, and H. Prendinger, ''Deep learning for affective computing: Text-based emotion recognition in decision support,'' 2018, arXiv:1803.06397. [Online]. Available: https://arxiv.org/abs/1803.06397

[5] N. Colneric and J. Demsar, ''Emotion recognition on Twitter: Comparative ˆ study and training a unison model,'' IEEE Trans. Affect. Comput., to be published.

[6] C. Strapparava and R. Mihalcea, ''Semeval-2007 task 14: Affective text,'' in Proc. 4th Int. Workshop Semantic Eval., Stroudsburg, PA, USA: Association for Computational Linguistics, 2007, pp. 70–74. [Online]. Available: http://dl.acm.org/citation.cfm?id=1621474.1621487

[7] C. Strapparava and A. Valitutti, ''Wordnet affect: An affective extension of WordNet,'' in Proc. LREC, vol. 4, 2004, p. 40.

[8] S. M. Mohammad and P. D. Turney, ''Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon,'' in Proc. NAACL HLT Workshop Comput. Approaches Anal. Gener. Emotion Text. Los Angeles, CA, USA: Association for Computational Linguistics, 2010, pp. 26–34.

[9] S.-L. Yeh, Y.-S. Lin, and C.-C. Lee, ''An interaction-aware attention network for speech emotion recognition in spoken dialogs,'' in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), May 2019, pp. 6685–6689.

[10] S.-Y. Chen, C.-C. Hsu, C.-C. Kuo, T.-H. Huang, and L.-W. Ku, ''Emotionlines: An emotion corpus of multi-party conversations,'' 2018, arXiv:1802.08379. [Online]. Available: https://arxiv.org/abs/1802.08379

[11] N. Majumder, S. Poria, D. Hazarika, R. Mihalcea, A. Gelbukh, and E. Cambria, ''DialogueRNN: An attentive RNN for emotion detection in conversations,'' 2018, arXiv:1811.00405. [Online]. Available: https://arxiv.org/abs/1811.00405

[12] H. Zhou, M. Huang, T. Zhang, X. Zhu, and B. Liu, ''Emotional chatting machine: Emotional conversation generation with internal and external memory,'' in Proc. 32nd AAAI Conf. Artif. Intell., 2018, pp. 730–738.

[13] C.-H. Wu, Z.-J. Chuang, and Y.-C. Lin, ''Emotion recognition from text using semantic labels and separable mixture models,'' ACM Trans. Asian Lang. Inf. Process., vol. 5, no. 2, pp. 165–183, 2006.

[14] S. Shaheen, W. El-Hajj, H. Hajj, and S. Elbassuoni, ''Emotion recognition from text based on automatically generated rules,'' in Proc. IEEE Int. Conf. Data Mining Workshop, Dec. 2014, pp. 383–392.

[15] Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv 2014, arXiv:1412.3555.

[16] Sameh Al-Natour, Ozgur Turetken (2020), A comparative assessment of sentiment analysis and star ratings for consumer reviews. Elsevier International Journal of Information Management 54, https://doi.org/10.1016/j.ijinfomgt.2020.102132

[17] S. Godfrey Winster, M. Naveen Kumar (2020), Automatic classification of emotions in news articles through Ensemble decision tree classification techniques. Journal of Ambient Intelligence and Humanized Computing, https://doi.org/10.1007/s12652-020-02373-5

[18] Graterol. W, Díaz-Amado J, Cardinale Y, Dongo I, Lopes-Silva E, Santos-LibarinoC(2021) Emotion Detection for Social Robots Based on NLP Transformers and an Emotion Ontology. Sensors, 21,1322. https://doi.org/10.3390/s21041322

[19] Joao Luis Zeni Montenegro, Cristiano Andréda Costa ∗, Rodrigo da Rosa Righi (2019) Survey of conversational agents in health, Expert Systems withApplications 129, 56–67, 0957-4174, Elsevier https://doi.org/10.1016/j.eswa.2019.03.054

[20] Kamil Topal, Gultekin Ozsoyoglu, (2016) Movie Review Analysis: Emotion Analysis of IMDb Movie Reviews, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 978-1-5090-2846-7/16

[21] Borod JC. The Neuropsychology of Emotion. Oxford, UK: Oxford University Press; 2000.

[22] Ekman P. Basic emotions. Handbook Cognit Emot. 1999;98(45-60):16.

[23] Hendrick C, Lilly RS. The structure of mood: a comparison between sleep deprivation and normal wakefulness conditions. J Pers.1970;38(4):453-465.

[24] Clyde Dean J. Manual for the Clyde mood scale. Coral Gables, Fla.: Biometric Laboratory, University of Miami; 1963.

[25] Borgatta EF. Mood, personality, and interaction. J Gen Psychol. 1961;64(1):105-137.

[26] Plutchik R. A general psychoevolutionary theory of emotion. Amsterdam, Netherlands: Elsevier; 1980 (pp. 3–33).

[27] Ortony A, Clore GL, Collins A. The Cognitive Structure of Emotions. Cambridge, MA: Cambridge University Press; 1990.

[28] Lang PJ. The emotion probe: studies of motivation and attention. Am Psychol. 1995;50(5):372.

[29] Scherer KR, Wallbott HG. Evidence for universality and cultural variation of differential emotion response patterning. J Pers Soc Psychol. 1994;66(2):310.

[30] Russell JA. A circumplex model of affect. J Pers Soc Psychol. 1980;39(6):1161.

[31] Russell JA, Mehrabian A. Evidence for a three-factor theory of emotions. J Res Pers. 1977;11(3):273-294.

[32] Canales L, Martínez-Barco P. Emotion detection from text: a survey. Paper presented at: Proceedings of the Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days; 2014:37--43; ACM.

[33] Poria, Soujanya, et al. "Emotion recognition in conversation: Research challenges, datasets, and recent advances." IEEE access 7 (2019): 100943-100953.

[34] Acheampong, Francisca Adoma, Chen Wenyu, and Henry Nunoo-Mensah. "Text-based emotion detection: Advances, challenges, and opportunities." Engineering Reports 2.7 (2020): e12189. Canales, Lea, and Patricio Martínez-Barco. "Emotion Detection from Text: A Survey." Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC), 2014, pp. 37-43.

[35] Hazarika, Devamanyu, Roger Zimmermann, and Soujanya Poria. "Misa: Modality-invariant and-specific representations for multimodal sentiment analysis." Proceedings of the 28th ACM international conference on multimedia. 2020.

[36] Kim, Evgeny, and Roman Klinger. "Frowning and Saying 'No!': A Systematic Review of Affective Analysis of Nonverbal Behavior in Conversations." ACL Anthology, 2020.

[37] Peng, Sancheng, et al. "A survey on deep learning for textual emotion analysis in social networks." Digital Communications and Networks 8.5 (2022): 745-762.

[38] Machová, Kristína, et al. "Detection of emotion by text analysis using machine learning." Frontiers in Psychology 14 (2023): 1190326.

[39] Plaza-del-Arco, Flor Miriam, et al. "Emotion analysis in NLP: Trends, gaps and roadmap for future directions." arXiv preprint arXiv:2403.01222 (2024).

[40] Strapparava, Carlo, and Alessandro Valitutti. "Wordnet affect: an affective extension of wordnet." Lrec. Vol. 4. No. 1083-1086. 2004.

[41] Russell, James A. "A circumplex model of affect." Journal of personality and social psychology 39.6 (1980): 1161.

[42] Goudjil, Mohamed, et al. "A novel active learning method using SVM for text classification." International Journal of Automation and Computing 15.3 (2018): 290-298.

[43] Li, Yanran, et al. "Dailydialog: A manually labelled multi-turn dialogue dataset." arXiv preprint arXiv:1710.03957 (2017).

[44] Shukla, Diksha, and Sanjay K. Dwivedi. "A comparative study of text-based emotion detection techniques for emotion recognition on social media data." 2023 IEEE 7th Conference on Information and Communication Technology (CICT). IEEE, 2023.

[45] Singh, Gurpreet, et al. "Potential approach for text-based emotion detection using NLP coupled with deep learning of sentiment analysis." 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2024.

[46] Tripathi, Samarth, Sarthak Tripathi, and Homayoon Beigi. "Multi-modal emotion recognition on iemocap dataset using deep learning." arXiv preprint arXiv:1804.05788 (2018).

[47] Acheampong, Francisca Adoma, Henry Nunoo-Mensah, and Wenyu Chen. "Transformer models for text-based emotion detection: a review of BERT-based approaches." Artificial Intelligence Review 54.8 (2021): 5789-5829.

[48] Kamath, Patrick S., and W. Ray Kim. "The model for end-stage liver disease (MELD)." Hepatology 45.3 (2007): 797-805.

[49] Zhang, Dong, et al. "Modeling both intra-and inter-modal influence for real-time emotion detection in conversations." Proceedings of the 28th ACM International Conference on Multimedia. 2020.

[50]Rahman, Tauhidur, and Carlos Busso. "A personalized emotion recognition system using an unsupervised feature adaptation scheme." 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2012.