RESEARCH ARTICLE

Asian Research Association

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION

DOI: 10.54392/irjmt2528

Feature Engineering Trends in Text-Based Affective Computing: Rules to Advance Deep Learning Models

Geeta Pattun [a, *], Pradeep Kumar [a]

[a] Department of Computer Science and Information Technology, Maulana Azad National Urdu University, Hyderabad, 500032, India

* Corresponding Author Email: geeta781@manuu.edu.in

DOI: https://doi.org/10.54392/irjmt2528

Received: 20-07-2024; Revised: 03-03-2025; Accepted: 12-03-2025; Published: 21-03-2025

Abstract: Understanding emotions in textual data, particularly within dynamic social media platforms such as YouTube, Facebook, and Twitter, presents significant challenges. This paper aims to provide a comprehensive review of emotion detection techniques in affective computing, highlighting key advancements, challenges, and ethical concerns. The key contributions of this review include an examination of foundational theories of NLP-based emotion recognition, an analysis of the role of affect lexicons in emotional classification, and a review of commonly used datasets for training emotion detection models. Additionally, it explores various feature extraction techniques, including lexicon-based approaches such as SentiWordNet and NRC Emotion Lexicon, statistical and syntactic features like n-grams and POS tags, and semantic embeddings from deep learning models such as Word2Vec, GloVe, BERT, RoBERTa, and GPT. Findings show that while deep learning and transformer models improve contextual understanding, they also introduce challenges such as high computational costs, data imbalance, and domain adaptability issues. Bias in training data poses ethical risks, potentially reinforcing stereotypes and enabling manipulative applications like targeted advertising and misinformation. Key research gaps include the need for improved feature representations, bias mitigation, enhanced model accuracy and fairness. Traditional models struggle with real-world complexities, while transformer-based models face challenges related to scalability, dataset limitations, and interpretability. Addressing these challenges will enhance affective computing accuracy, fairness, and applicability across industries such as healthcare, education, and human-computer interaction.

Keywords: Affective Computing, Emotion Detection, Feature Engineering, Text Based, Embedding

1. Introduction

Affective Computing is an emerging field of research that aims to develop smart systems with cognitive, interpretive, and quick response capabilities [1]. It encompasses various research areas, including sentiment analysis and emotion analysis. Although sentiment analysis and emotion analysis are distinct, with the former focusing on broad categories like positive or negative and the latter exploring nuanced emotional states [2], both fall under the umbrella of affective computing. The term sentic refers to sentiments and sensations, highlighting the connection between these concepts in affective computing. Social media platforms have proven to be vital public spaces where stakeholders across different domains express their views and opinions through textual and multimedia content [3, 4]. This wealth of data can significantly enhance sentiment analysis by providing diverse perspectives. Emotion theories play a crucial role in strengthening sentiment analysis by moving beyond binary classifications. They enable NLP systems to detect nuanced emotional states more effectively. Moreover, their application in dataset annotation and labeling improves training data quality, leading to more robust emotion detection models that are adaptable across diverse domains such as affective computing, psychology [5], educational technology [6], and neuroscience.

In the context of text-based emotion detection within affective computing, extracting emotions from text is quite complicated compared to extracting emotions from audio or visual cues like facial expressions or physiological data such as skin temperature [7, 8]. The author emphasized that recognizing emotional information requires extracting meaningful clues from textual content [9]. Lexicon dictionaries have been indispensable for conventional models by providing meaningful emotional clues through techniques such as expanding lexicons using seed words or incorporating corpus-based lexicon expansion. The choice of database significantly impacts feature extraction techniques for inferring emotions from text. Labeled datasets simplify this process, whereas unlabeled datasets pose challenges. By leveraging advancements in natural language processing (NLP) such as deep learning models (e.g., BERT), researchers can significantly enhance feature engineering for text-based embeddings. These models excel at capturing contextual information essential for nuanced emotion detection [10]. Overall, this integration highlights how advancements in feature engineering contribute to developing more robust affective computing systems capable of handling real-world complexities through improved text-based emotion recognition.

The literature on affective computing reveals persistent gaps that hinder progress. Conventional rule-based systems, despite their structured nature, struggle with real-world complexities such as informal language found in social media [11]. Traditional machine learning models face challenges like reliance on static features, data imbalance, and limited adaptability across contexts. These issues impede accurate detection of nuanced emotional expressions, underscoring the need for advanced methods. Recent NLP advancements have introduced transformer-based models (e.g., BERT and GPT), which offer improvements by capturing contextual cues and enabling transfer learning. However, a significant research gap exists in applying these models within affective computing due to unexplored potential and challenges such as high computational demands, need for high quality datasets, need for diverse datasets, transparency and interpretabilities issues.

This review is novel as it uniquely traces the evolution of feature emphasis in affective computing over the past decade (2013–2023), from early rule-based systems to modern transformer-based deep learning approaches. It provides a comprehensive comparison of methodologies, highlighting their strengths, limitations, and the impact of feature utilization advancements on effectiveness. The review underscores the limitations of conventional systems that rely on affect lexicons and static feature engineering, which struggle with dynamic textual data. In contrast, it examines transformer models' strengths in capturing global context while addressing challenges related to computational demands, data quality, and ethical transparency [12, 13]. Importantly, this review integrates ethical considerations by emphasizing transparency, fairness, and interpretability in emotion recognition systems. By bridging technical and ethical perspectives through a hybrid approach that leverages both traditional and transformer models, it offers actionable insights for developing robust, adaptive, and trustworthy affective computing systems capable of navigating real-world complexities while ensuring privacy and ethical integrity.

The PubMed trend visualization, shown in Figure 1, was prepared based on data retrieved in February 2025 using four search queries: (affective computing AND textual), (affective computing AND text), (emotion analysis AND text), and ((emotion classification AND textual) OR (emotion detection AND text)) for the 2020–2025 period. Its numeric data is presented in Table 1 referred from emotion detection and text - Search Results – PubMed (https://pubmed.ncbi.nlm.nih.gov/?term=emotion+detection+and+text&filter=years.2020-2025). The visualization and data highlight the need for further research in affective computing, particularly in leveraging textual data to develop intelligent systems with cognitive, interpretive, and responsive capabilities crucial for emotion-aware applications across various platforms.

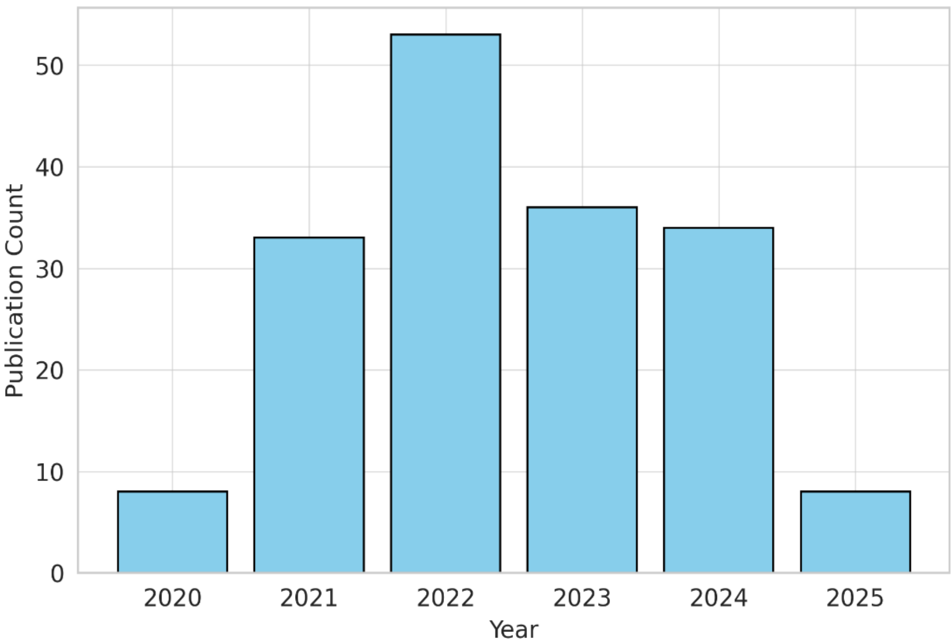

Figure 1. PubMed-based trend analysis of affective computing (2020 - 2025)

Table 1. PubMed query results for affective computing (2020–2025)

S.No	Year	Search Query-1	Search Query-2	Search Query-3	Search Query-4	count
1	2020	4	2	1	1	8
2	2021	7	8	6	12	33
3	2022	9	10	10	24	53
4	2023	4	7	9	16	36
5	2024	6	3	11	14	34
6	2025	2	2	3	1	8

1.1 Affect lexicons and Emotion Models

Emotion refers to strong feelings, as given in WordNet, and it is a large lexical database of English. Affective term selection and understanding are crucial for realizing appropriate expression in conversation, such as relating direct reference words to emotion states (e.g., fear, cheerful) and indirect references to emotion states (e.g., monster, ghost) [14]. Affective terms help focus on the classification of emotion and valence (positive and negative polarity) in context, and they also help instigate a viable connection between emotions and semantics [15]. The emotion in the text can be expressed explicitly using affective words (e.g., happy, or guilty), or implicitly using emotion-bearing words.

Emotion theories play a fundamental role in NLP-based emotion recognition systems by providing a standardized framework for categorizing and interpreting emotions. They ensure consistency in emotion detection by offering structured representations, such as discrete (joy, anger, sadness) and dimensional (valence-arousal) models. These theories guide feature extraction by helping identify linguistic, syntactic, and semantic clues essential for recognizing emotions in text. Additionally, they enhance contextual understanding, reducing misclassification by interpreting emotions within different textual and situational contexts. Psychologists have conceptualized human emotions in two or three dimensions. The emotion models presented in Figures 2, 3, and 4 have been widely adopted in the literature across various platforms as a foundation for a deeper understanding of affective computing studies.

The Ekman Model of Emotions, developed by Paul Ekman in the 1970s and illustrated in Figure 2, is a discrete emotion model that identifies universally recognized emotions based on cross-cultural studies of facial expressions. Unlike dimensional models, which represent emotions on a spectrum, Ekman's model classifies emotions into six primary categories: happiness, sadness, fear, anger, surprise, and disgust. These emotions were validated through studies conducted in diverse cultural settings, including remote tribes, demonstrating that they are biologically hardwired rather than culturally learned.

Ekman's model is particularly valuable for affective computing and NLP-based emotion recognition, as it provides a structured framework for emotion classification in text, speech, and facial analysis. By associating emotions with specific linguistic and facial cues, it enables rule-based and machine learning models to detect emotional expressions in human communication. However, a limitation of this model is that it does not account for emotion intensity or mixed emotions, which can be important for capturing subtle variations in sentiment. Despite this, it remains widely used in sentiment analysis, chatbot development, and human-computer interaction, offering a reliable foundation for emotion detection systems.

The Circumplex Model of Emotions, developed by James A. Russell in 1980 and shown in Figure 3, takes a dimensional approach to understanding emotions by mapping them onto a continuous two-dimensional space. Unlike discrete models that categorize emotions into fixed types, the Circumplex Model represents emotions based on Valence (pleasant to unpleasant) and Arousal (high to low energy). This framework captures the relationships between different emotions, showing how they transition smoothly rather than existing as isolated states. For example, joy is characterized by high arousal and positive valence, while sadness is associated with low arousal and negative valence.

A key advantage of the Circumplex Model is its ability to represent emotion intensity and gradation, making it particularly valuable for affective computing and NLP-based emotion recognition. By incorporating valence and arousal dimensions, computational models can more accurately analyze emotional expressions in text, distinguishing between emotions with similar sentiment but different energy levels (e.g., anger vs. frustration). Additionally, the model helps in recognizing context-dependent emotional variations, improving emotion classification in dialogue systems, sentiment analysis, and human-computer interaction.

Figure 2. Ekman's Model of Emotions (1970s)

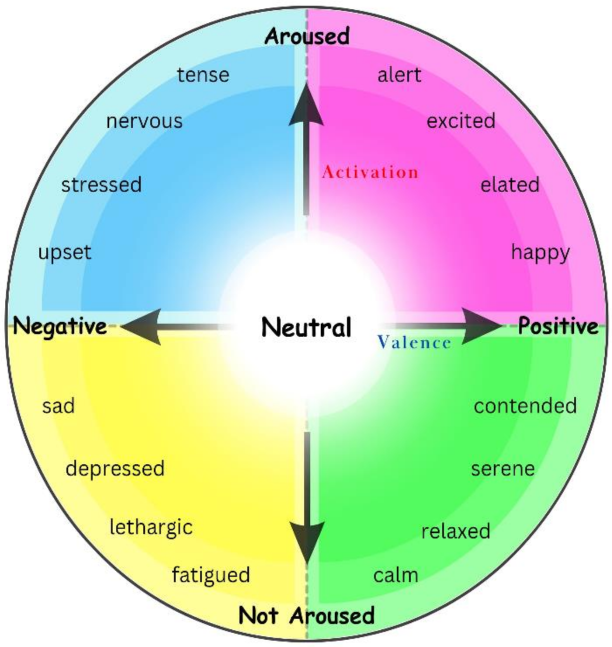

Figure 3. Circumplex Model of Emotions (Russell, 1980)

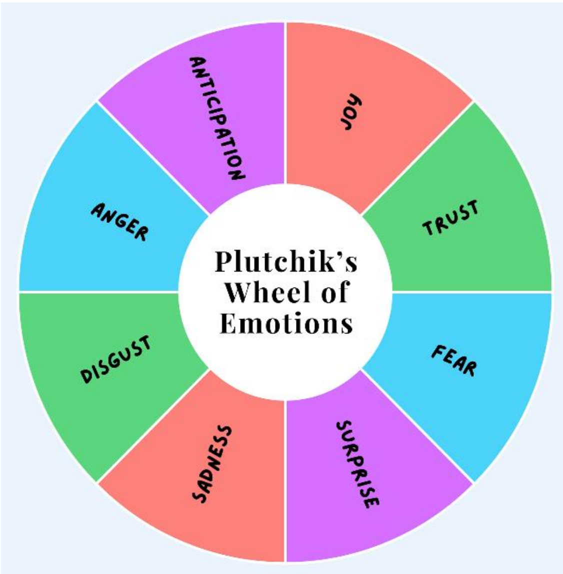

Figure 4. Plutchik's Wheel of Emotions (1980)

`Its continuous nature makes it well-suited for deep learning-based approaches, enabling more nuanced and flexible emotion detection in real-world applications.

Plutchik's Model of Emotions, shown in Figure 4 and developed in 1980, presents a structured approach to understanding emotions by categorizing them into primary, secondary, and complex forms. Unlike discrete models that classify emotions into fixed categories, Plutchik's model emphasizes the intensity, blending, and adaptive function of emotions. It identifies eight primary emotions—joy, trust, fear, surprise, sadness, disgust, anger, and anticipation—considered fundamental across cultures. These emotions vary in intensity, forming a spectrum where, for example, joy ranges from serenity to ecstasy, while anger ranges from annoyance to rage. Additionally, emotions can mix to form secondary emotions, such as love (joy + trust) or remorse (sadness + disgust), illustrating the dynamic and interconnected nature of human feelings.

Beyond classification, Plutchik's model highlights the adaptive function of emotions, suggesting that emotions evolved to enhance survival. Fear triggers a fight-or-flight response, trust fosters social bonding, and disgust helps avoid harmful substances. The model is often visualized as a wheel, where opposing emotions (e.g., joy vs. sadness, fear vs. anger) are placed opposite each other to illustrate contrasts and relationships. In affective computing, Plutchik's framework is valuable for NLP-based emotion recognition, enabling models to capture emotional nuances by incorporating intensity variations and emotion blending, improving the accuracy of text-based emotion classification.

1.2 Emotion Databases and its relevance

The temporal evolution of commonly used text-based datasets, shown in Figure 5 for the period 1990 to 2022, highlights the progression of data collection practices, annotation methodologies, and the increasing complexity of emotion modeling. Early literature datasets on affectively study were focused primarily on discrete emotion categories, while more recent datasets, such as those from SemEval-2017 and CMU-MOSEI, have expanded to include dimensional models of emotion, emotion intensity, and multimodal data. This temporal evolution demonstrates the increasing sophistication of emotion detection research, which has shifted from simple sentiment analysis to more nuanced and context-aware emotion recognition systems. By examining this evolution, we gain insights into the changing needs and approaches in the field, making it clear that newer datasets address more complex challenges in emotion detection, such as detecting subtle emotional variations, understanding emotional intensity, and integrating multimodal data sources.

The datasets described in Table 2 are crucial for affective computing, as they cover a range of emotion categories, from discrete emotions (e.g., anger, sadness, joy) to dimensional emotions (e.g., arousal, valence). They also include key attributes such as dataset name, description, year, source/creator, emotion labels, and size, providing a foundational resource for developing and improving emotion recognition models. Among these, datasets like ISEAR, SemEval-2017 Task 4, EmoBank, and WASSA-2017 provide a comprehensive range of emotion-labelled data essential for emotion detection from text. These datasets cover various domains, from social media to daily conversations, and include diverse emotional categories, allowing for the development of robust models capable of identifying a wide range of emotional expressions. Datasets like EmoBank and SemEval-2018 offer structured annotations for subtle emotion distinctions, while WASSA-2017 and the Toxic Comments dataset provide a focus on emotion intensity and negative emotional states like anger, making them invaluable for more granular emotion recognition.

Additionally, datasets such as Daily Dialog, Grounded Emotions, and CMU-MOSEI enrich emotion detection research by providing data that reflects natural, conversational text and multimodal inputs. These datasets are critical for training models that detect emotions in real-world applications, from chatbots to social media content moderation. Resources like the OANC dataset and Cecilia Ovesdotter Alm's Affect Data contribute to building emotion detection systems that generalize well across different languages and cultural contexts. Collectively, these datasets provide the diverse, high-quality data necessary to advance emotion detection research and create systems that understand and respond to emotional cues in text.

Figure 6 visualizes the emotion classes reported in datasets used in affective computing studies over the last five years, presented through a bar chart. Highlighting the significance of these emotion classes is essential for fostering positive emotional experiences and mitigating negative ones across various domains such as education, workplaces, and social interactions. A deeper understanding of these emotions can contribute to improved personal development, productivity, and overall quality of life.

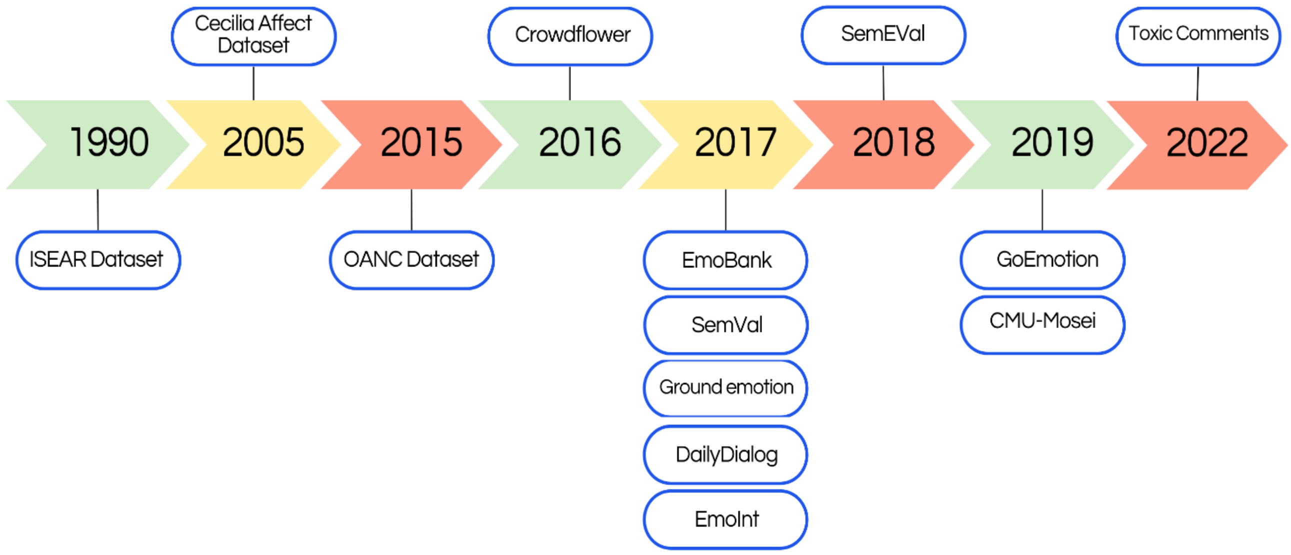

Figure 5. Evolution of text-based datasets (1990–2022) in affective computing

Table 2. Text-based Emotion Dataset reported in the literature

Dataset Name	Description	Year	Source/Creator	No of Emotions	Size
ISEAR	International Survey on Emotion Antecedents and Reactions	1990	University of Geneva	7 emotions	~7,666 instances
SemEval -2017	Affective Text Dataset	2017	Sem Eval	6 emotions	1250 instances
EmoBank	Affective Text Dataset for sentiment and emotion recognition	2017	University of Melbourne	6 emotions	~10K sentences
WASSA 2017 EmoInt	Constructed from tweets	2017	WASSA	6 emotions	8,000 instances
Cecilia Ovesdotter Alm's Affect Data	Constructed from tales	2005	Cecilia Ovesdotter Alm	6 emotions	~2,800 sentences
Daily Dialog	Daily dialogues dataset with emotions	2017	HCTI	7 emotions	13,118 dialogues
Crowd Flower	Emotion analysis of text from various crowdsourcing tasks	2016	Crowd Flower	13 emotions	39,740 instances
Grounded Emotions 2017	Dataset for emotion recognition in textual content	2017	WASSA	2 emotions	2557 tweets
SemEval 2018	Affective text dataset for emotion classification	2018	Sem Eval	11 emotions	6838 instances
Toxic Comments Dataset	Dataset of toxic comments for classification	2022	Jigsaw	6 forms of toxicity	15,9,571comments
OANC Dataset	Open American National Corpus for emotion analysis	2015	Linguistic Data Consortium	6 emotions	14 million words
CMU-MOSEI	Multimodal sentiment analysis dataset	2016	CMU	7 emotions	23,000 videos

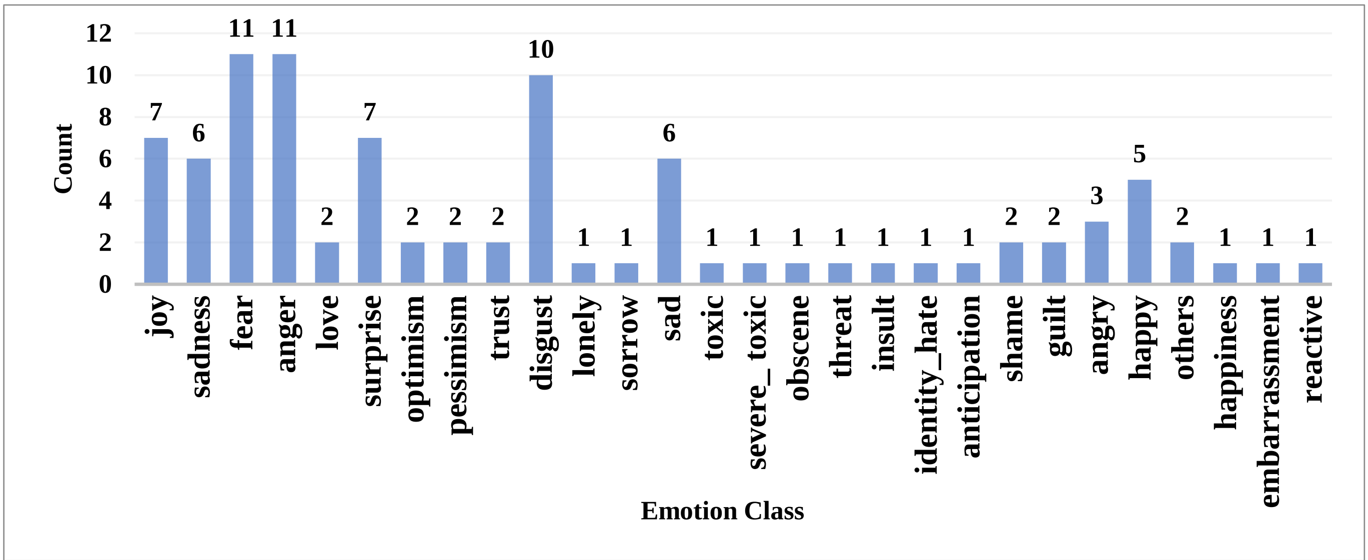

Figure 6. Emotion classes reported in affective computing datasets

2. Conventional Rule Based Models strengths

Lexicons provide a structured framework for mapping emotions to words, ensuring standardized and systematic emotion recognition in text. Resources such as WordNet Affect, LIWC, ANEW, and SentiWordNet enable models to identify both discrete and dimensional emotions, covering a spectrum from basic feelings like joy and sadness to nuanced states like valence and arousal. Without lexicons, emotion detection systems would struggle with inconsistencies and errors, particularly in informal language contexts. Lexicon-based approaches address challenges posed by slang, symbols, and informal language in real-world data.

Emoticon lexicons and slang dictionaries enhance models ability to interpret non-standard expressions, such as emojis and internet slang, which are prevalent in social media and digital communication. This adaptability ensures robust performance across diverse textual genres, from formal writing to casual conversations. Lexicons like NRC EmoLex and Ren-CECps further refine emotion detection by incorporating cultural and contextual insights, improving generalizability. Self-tailored lexicons enhance domain-specific accuracy, allowing systems to adapt rule-based methods to specialized contexts

Table 3 presents a comparative analysis of lexicon dictionaries based on language, granularity, purpose, and emotion types. It also highlights the relevance of specific lexicons as features in affective computing within conventional models.

Table 3. Comparison of lexicons as features based on language, granularity, purpose, and emotion

Lexicon Resource	Reference	Language	Granularity	Purpose	Source	Emotion Types
WordNet Affect	[15-17]	English	Word-level	Emotion detection based on WordNet	WordNet	Affective states (anger, fear, etc.)
LIWC	[18]	English (and others)	Word-level	Psychological analysis of text	Linguistic Inquiry	Emotions, cognitive states
ANEW	[19]	English	Word-level	Measures emotional intensity of words	University of Florida	Arousal, Valence, Dominance
SentiWordNet	[20, 21]	English	Word-level	Sentiment analysis using WordNet	WordNet	Positive, Negative
SentiWordNet 3.0	[22, 23]	English	Word-level	Updated sentiment analysis	WordNet	Positive, Negative, Neutral
Emoticon lexicon	[24-26]	English	Emoticon-level	Mapping emoticons to emotions	Online resources	Various (joy, sadness, etc.)
Slang Lexicon	[27-29]	English	Word-level	Maps slang words to meanings/emotions	Online sources	Varies (colloquial emotions)
NRC emoLex	[30, 31]	English (and others)	Word-level	Emotion and sentiment lexicon	NRC	Joy, Sadness, Anger, etc.
Ren-CECps	[32, 33]	Chinese	Word-level	Emotion detection for Chinese text	Renmin University	Joy, Sadness, Anger, etc.
Phrasal Verbs	[34- 37]	English	Multi-word level	Emotion mapping for phrasal verbs	Online resources	Various emotions

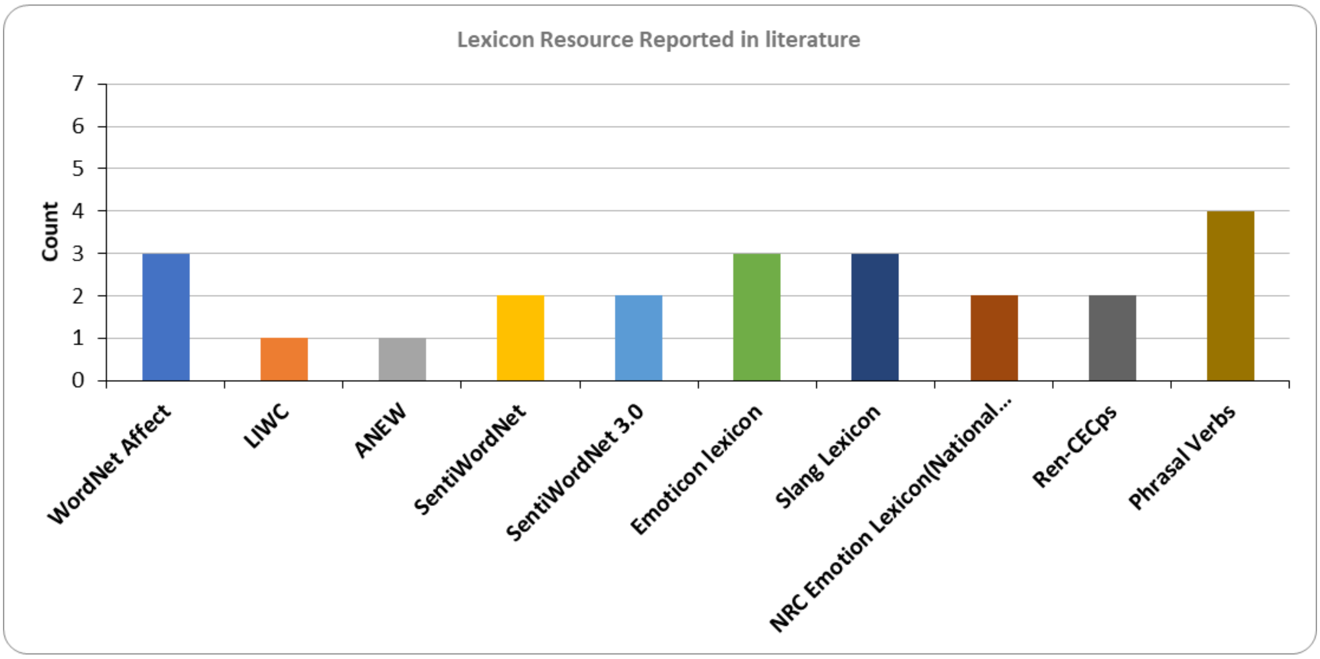

Figure 7. Usage trends of lexicons in affective computing studies

This comparison aids researchers in selecting the most appropriate lexicon for their studies, ensuring informed decisions for accurate and context-aware emotion detection. Lexicons used in emotion detection research are illustrated in Figure 7, showcasing the prevalence of resources such as WordNet-Affect, NRC EmoLex, and SentiWordNet. This visualization helps researchers identify influential lexicons and choose suitable resources for emotion recognition. It also emphasizes that despite the growing use of machine learning and hybrid models, lexicon-based approaches remain essential for emotion detection.

2.1 Limitations of Conventional Rule Based Models

As per authors of [28], [38] whether conducting sentence-level sentiment analysis or a document-level emotion detection of social media data, affective computing studies still face significant challenges. These challenges include deficient coverage of emotion words, limited awareness of polarity shifters and negations, restricted availability of lexicon resources for emoticons, and a lack of lexicon resources for slang terms in social media.

2.2 Research Gaps in conventional Rule Based models

To perform affective computing using emotional words, certain rules were formulated to classify the emotions behind user-generated text. However, these rules often lack scalability and adaptability to different domains [39-41]. This emotion detection process from text is a challenging and time-consuming [42] with issues such as language ambiguity, text bearing multiple emotions, or text not being associated with any emotion category [43] are major concerns.

3. Machine Learning models and its strength

Machine learning models have transformed affective computing by automating emotion recognition through advanced feature extraction and classification techniques. Traditional rule-based methods rely on lexicons such as WordNet Affect, SentiWordNet, and NRC Emotion Lexicon, but they often struggle with context ambiguity and informal language variations. To overcome these limitations, machine learning models incorporate diverse feature sets, enabling context-aware and data-driven emotion analysis. Feature engineering plays a critical role in improving the accuracy of machine learning-based emotion detection.

Lexicon-based features (e.g., Extended SentiWordNet, emoticon lexicons, and slang lexicons) provide semantic insights, while word embeddings such as Word2Vec, GloVe, Fast Text, and Google Embeddings capture contextual relationships. Additionally, techniques like CBOW, TF-IDF, and Bag of Words (BoW) extract statistical patterns from text, aiding in emotion classification. The features used in affective computing literature are shown in Figure 8, highlighting the key features employed by machine learning classifiers in affective computing studies.

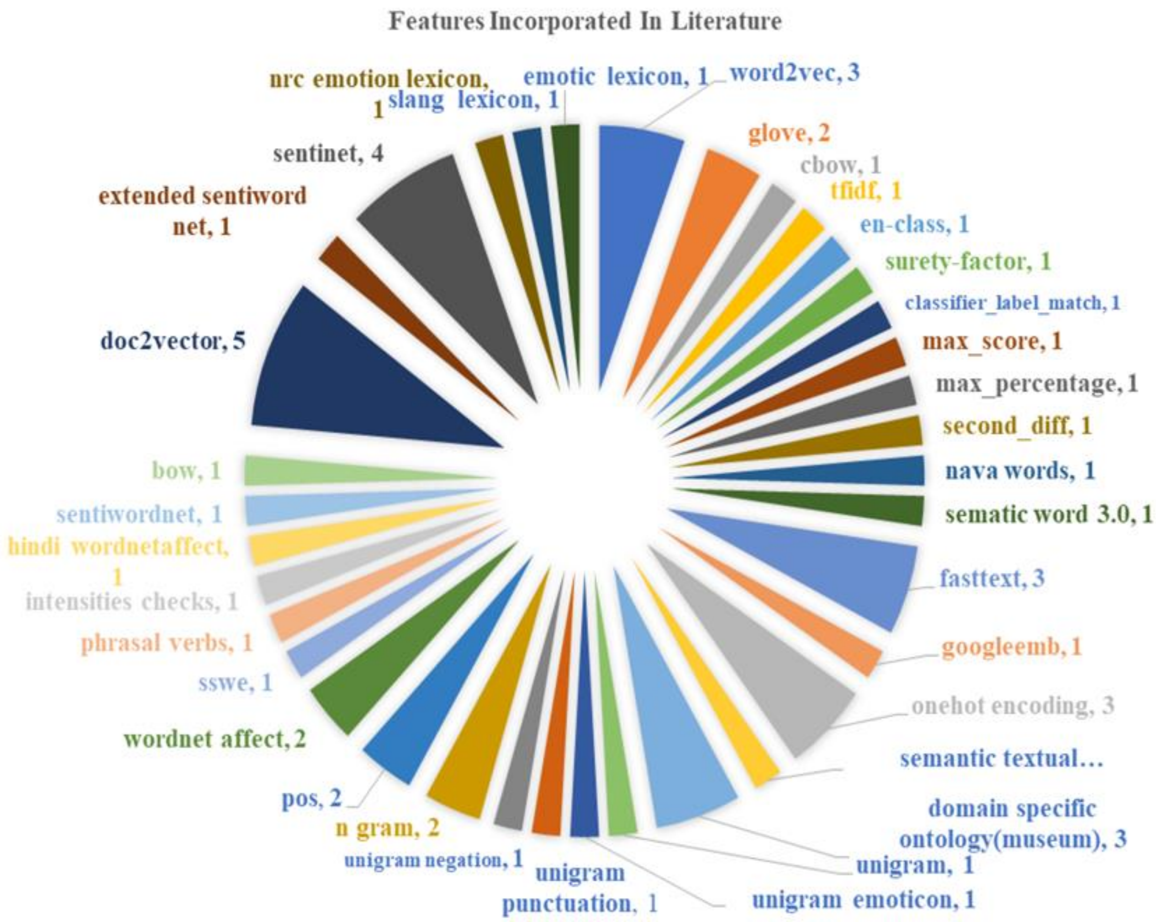

Figure 8. Feature incorporated in Machine Learning Model

3.1 Challenges of Machine Learning Model

Shah, *et al* [16] proposed a novel approach that includes emotion lexicon and other feature vector conversions for Emotion Detection using Machine learning. Herzig *et al* [44] combined Word2vec and GloVe embeddings with traditional features such as CBow, TFIDF and EN-Class to achieve high accuracy, however this resulted in longer training time. A study by [25] on tweet emotions explored features like surety factor, classifier label matches, and maximum score but focused solely on single emotions, neglecting the possibility of mixed emotions within a tweet. While utilizing Nava words and semantic words 3.0 was effective, the study suggests incorporating a domain-specific semantic corpus for even better performance [45]. Polignano, *et al* [46] compared various word embeddings like GloVe, FastText, Doc2vec and Google word embeddings, highlighting the potential for more fine-grained approaches to capture subtler emotional nuances. As per [47] the study using word-level features for utterance-level tasks (e.g., GloVe) emphasize the need for interpretable features to enhance model performance and improve generalizability.

The study in [48] employed word embedding, one-hot encoding, and post-padding techniques revealed the negative impact of imbalanced datasets on model effectiveness. It suggested data balancing techniques for improvement; similarly, another model using semantic textual similarity and a domain-specific ontology also faced challenges due to imbalanced datasets [49].The study in [50] explored features such as unigram, unigram emoticon, punctuation, and negation, but found inconsistencies in labeling due to non-expert text classification. A model incorporating N-grams, emoticons, POS tagging, negation, WordNet Affect, and SSWE achieved some success, however its training process lacked context awareness, as essential aspect for accurate emotion detection [51]. The author [35] identified sarcasm as challenge due to its reliance on context, when utilizing POS tagging, phrasal verbs, intensity checks, and negation. While resources like WordNet Affect, Hindi WordNet Affect, SentiWordNet, and unigrams provided rich semantic information, their high computational load necessitates exploring lighter alternatives or optimizing model architecture [52].

3.2 Research Gaps in Machine Learning-Based Affective Computing

A study [28] investigated sentiment lexicons such as Extended SentiWordNet, sentiNet, NRC Emotion Lexicon, slang lexicon, and emoticon lexicon, but these were not evaluated on a domain-specific corpus.

Table 4. Feature incorporated and Research Gaps in Machine Learning model for Affective Computing

Features	Research Gap	Reference
Word2vec, GloVe embedding in combination with CBow, TFIDF, EN- Class	Training time requires attention	[44]
Surety factor, Classifier_label_match, max_score, max_percentage, second_diff	Tweet with multiple emotions were not considered for training	[25]
Nava words, Semantic word 3.0	A Domain-specific Semantic Corpus is missing	[45]
GloVe, FastText, Google Embedding	Tools are needed to prepare significant for a level of granularity	[46]
Word-level features extracted for utterance-level features (eg. GloVe)	Lack of generalizability	[47]
Word embedding, one-hot encoding, post-padding	The Dataset was imbalanced	[48]
Semantic textual similarity, domain-specific ontology(museum)	Imbalanced Dataset	[49]
Unigram, Unigram emoticon, Unigram punctuation, Unigram negation	Classification texts by Non-experts are not adequately reliable. Experts labeling is a costly.	[57]
N gram, emoticon, POS tagging, negation, wordNet affect, SSWE	The training process did not adequately account for the context of the dialogue	[51]
POS tagging, phrasal verbs, intensities checks, negation	Sarcasm was not handled, the context-wise model suffered	[35]
Wordnet Affect, Hindi Wordnet Affect SentiWordNet, unigrams	Heavy computational load	[52]
BOW, Word2Vector,doc2Vector	High training time complexity	[53]
Extended SentiWord Net, sentinet, NRC Emotion Lexicon, slang Lexicon, emotic lexicon	Domain-specific techniques not investigated.	[28]

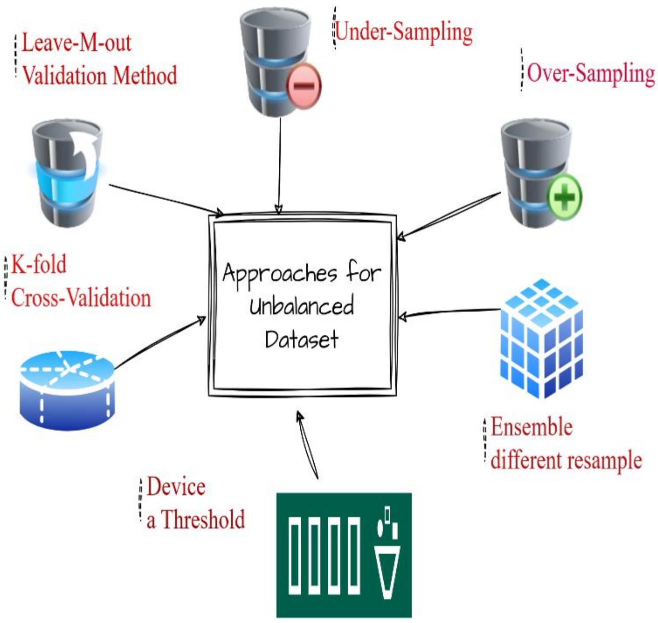

Figure 9. Techniques to handle unbalanced dataset

The research gaps in feature selection for training machine learning algorithms are summarized in Table 4 for better visibility and to highlight areas for improving performance in specialized applications. Imbalanced datasets remain a major challenge in emotion detection from text. For example, if a dataset consists mostly of happy tweets, it can bias the model's predictions.

A trained model might excel at identifying happiness while missing more subtle expression of sadness. This bias arises because the model prioritizes learning patterns from the dominant emotion. Common metrics like accuracy can be misleading, since high scores may affect the performance on underrepresented emotions. This limitation hinders the model's ability to adapt to real-world data with varying emotional distributions. The study referenced in [53] achieved superior results with features like BOW, Word2Vector, and doc2Vector, however it noted high training time complexity. Techniques for dimensionality reduction and model architecture optimization are recommended for future work. Figure 9 illustrates how imbalanced datasets can be effectively addressed through resampling methods (Under-sampling [54], Over-sampling [55]), K-fold Cross-Validation, leave-M-out validation methods can help in dealing with serious issue of imbalance dataset but further research is required in this direction. Additionally employing Ensemble techniques with different resample datasets at various ratios tailored to specific requirement can further enhance model performance [56].

4. Deep Learning Models and their strength

We are in the era where the machines can anticipate human needs; they are increasingly equipped to understand aspects of human emotions. Affective Computing is paved the way for machine to provide tailored responses in real-world applications such as Chatbots and dialog-based system [51]. These applications encourage further exploration of deep learning techniques in Affective Computing, moving away traditional feature-based machine learning methods SS-BED, is system a new deep learning framework, that can address emotion understanding in text conversations by combining semantic and sentiment analysis. Banothu, *et al* [48] developed a deep learning framework using LSTM network and incorporated web embedding, one-hot encoding, post padding, and lexicon-based emotion resource for extraction and classification of Twitter data, this framework achieved an accuracy of 92% and detected 6 basic emotions.

Table 5. Embeddings incorporated as features in advanced deep learning for affective computing

Reference	Word Embedding	Model Type	Training Data	Language	Contextual	Usage
[46]	FastTextEmb	Shallow, CBOW, Skipgram	Wikipedia, Common Crawl	Multilingual	No	Text classification, similarity tasks
[27, 56-61]	GloVeEmb	Matrix factorization	Wikipedia, Gigaword	Multilingual	No	Semantic analysis, similarity tasks
[62, 63]	Word2Vec	Skipgram, CBOW	Google News	Multilingual	No	Text similarity, sentiment analysis
[61]	BERT	Transformer (deep)	Books Corpus, Wikipedia	English (and others)	Yes	NLP tasks, sentence classification
[64]	BanglaFast-Text word	CBOW, Skipgram	Bangla text corpora	Bangla	No	Bangla text classification, similarity
[65-68]	GPT-2	Transformer (large)	WebText	English	Yes	Text generation, conversational AI
[69-71]	GPT 3	Transformer (large)	Common Crawl, WebText, etc.	English	Yes	Text generation, summarization, QA
[68,69,72]	GPT- 3.5	Transformer (larger)	Common Crawl, WebText, etc.	English	Yes	Text generation, AI chatbots
[75,77]	GPT- 4	Transformer (largest)	Web Data, more diverse sources	Multilingual	Yes	Advanced text generation, multi-modal

Additionally Studies have proposed a Multi-task Learning model for multi-modal emotion recognition and sentiment analysis. This model detects six emotion classes (anger, disgust, fear, happy, sad, surprise) in videos, utilizing GloVe and Facet 2 as semantic resources [47]. The modalities considered included (16216+ 4625) statements, acoustic (1838), and substantial number of visuals were considered for training the RNN framework, the author emphasized the importance of contextual utterance for classification and validated the results with a F1 score of 75%. The framework was tested on CMU MOSEI Dataset. A largest data set contains 23l utterances [47, 58].

The author proposed a classification approach based on BILSTM, CNN [59], and self-attention Mechanism [46, 60]. Experiments demonstrate that the embedding layer significant impacts emotion classification. The model, which utilized three different embedding layers across three datsets with 3 different embedding layer on 3 different dataset ISEAR, SEMEval 2018 and SemVal 2019 have achieved F1 Scores of 84%, F1 scores of 70.3 % compared to just 44% of random forest classifier.

Feature extracted using Word2Vec, GloVe, and FastText model are configured with LSTM structure comprising of 200 hidden units and dropout rate of 3. Various embeddings used as features in deep learning-based affective computing literature are summarized in Table 5. The Toxic Comment dataset, comprising 150,000 comments labeled into six classes of abuse and harassment, and the SemEval 2018 Task 1 dataset, containing 7,000 tweets labeled into 11 emotion classes, were used in a proposed ensemble deep learning model for multi-label and binary classification of user-generated content.

4.1 Challenges of Deep learning Models

Emotion detection has not been thoroughly explored across all domains, while considerable amount of research has been conducted in the retail market and e-commerce market. Other areas such as crime detection, patient analysis for assessing the depression levels, real-time emotion analysis of a learner in the e-learning environment, have received limited attention. Self-learning approaches have encouraged researcher to adopt machine learning algorithms and deep learning algorithm. Dealing with unbalanced dataset often remains a challenging task. These classifiers often faces dual burden of handling very heavy computation and providing real time responses. It has been observed that an individual's cultural affiliation significantly impacts their emotional expression in specific situations. Limited corpus resources in various languages pose a significant challenge to cross-language research. Additionally, the casual style of written text remains a concern, along with reliability issues in corpus annotation and the lack of labeled corpora for multi-class problems. Over fitting continues to be a major challenge, and the research approaches proposed so far have proven to be only partially effective in addressing these issues.

5. Transformers: A Paradigm Shift

The literature offers various techniques for extracting contextual information; however there is growing demand for the robust method. Transformer based embedding have significantly enhanced the quality of contextual information extraction, which comes with some limitations. Ameer, *et al* [73] proposed use of multiple attention mechanisms and Transformer Networks for multi-label emotion classification. Hasan, *et al* [57] developed DeepEmotex models that leverage sequential transfer learning from pre-trained USE and BERT models. The DeepEmotex-BERT models also outperform a baseline Bi-LSTM [57, 74] by 23% on the EmoInt and Stimulus benchmark datasets, showcasing the power of transfer learning from large pre-trained language models. The researchers [13] trained 11 RoBERT a-large models on seven annotator-labeled models datasets using annotator-rated and self-reported emotion test sets. The self-reported models achieved higher average F1 scores in self-reported emotion test sets.

Imran, (2024) conducted a comparative analysis of state-of-the-art pre-trained transformers (PTMs) for emotion classification on GitHub and Stack Overflow datasets [75]. They evaluated BERT, RoBERTa, ALBERT, DeBERTa, CodeBERT and Graph CodeBERT against SEntiMoji, revealing improvements in F1 scores ranging from 1.17-16.79%.Incorporating polarity features in attention demonstrated gain of 1.0 - 10.23% over baseline PTMs. The experiments involved [76] BiLSTM and KNN for non-LLM-based techniques. They introduced two novel methods: the Multi-Iteration Agentic Workflow and the Multi-Binary-Classifier Agentic Workflow, while were employed in the SemEval 2024 Task [77] 10 on ERC and EFR, employing ensemble methods and transformer approaches. Graterol, *et al* [49] suggested a system based on the EMONTO that uses a social robot to identify emotion from text while concurrently storing the information in a semantic repository (an ontology to represent emotion).

This framework was examined for its performance in a tour guide robot at a museum, which detected emotions expressed by visitors regarding specific artwork. The author adapted a transformer architecture that supports simple transformer library, achieving an accuracy of 52% [49, 60].The framework trained on SEMEval 2018 Task [78], which contains tweets of English statements labeled with vector representing emotion across a total of 11 emotion classes such as anger, anticipation, disgust, fear, joy, optimist, pessimist, sadness, lone, surprise, and trust. Classification methods can further improved by using a transformer as a feature for textual semantic similarity for instance, MLSMOTE has proved effective in addressing imbalance dataset issue. Tesfagergish *et al* [79] proposed a two-stage sentiment analysis approach combining zero-shot learning and ensemble classification. Nedilko & Y. Chu [80-82] experimented with multi-class emotion classification on short English essays about harmful events. They found that fine-tuning a GPT-3 model outperformed baseline character-based XGBoost Classifier.

5.1 Challenges of Transformer Models

Emotion Detection in Transformers including GPT models are well suited for understanding emotions in text owing to their self-attention mechanisms. Hence, can be extended to more complex tasks, such as, emotion prediction on diverse datasets, but involves certain challenges. Transformers face performance challenges owing to the lack of high-quality datasets [2]. Wan & Woźniak [83] suggested frameworks that fully leverage the capabilities of pretrained models to process diverse and complex emotional expressions in online comments. Strategies are needed to make transformer models more cost-effective and sustainable in terms of computational resources and energy consumption Transformer models are computationally expensive, and further research is needed to make them more efficient for large-scale applications.

As per [75] need extensive refinement of pre-trained models (PTMs) is required to better address contextual nuances in empathetic software intelligence. Few-shot learning with generative pre-trained models such as ChatGPT and GPT-4 has inherent limitations due to constrained context windows, cost inefficiencies, and unpredictable model behavior [80]. There is a need to address the limitations of GPT models in causal reasoning, emotion intensity prediction, and coping behavior modeling, especially in high goal relevance scenarios [84]. Table 6 presents a comparative analysis of affective computing literature over the past five years, highlighting key limitations in transformers and deep learning models. According to [68] ChatGPT models struggle with tasks such as measuring engagement, assessing personality, detecting sarcasm, and identifying subjectivity. Understanding their decision-making process is crucial for establishing trust.

6. Ethical and Privacy Concerns of Affective Computing System

The collection of emotional data without meaningful consent can lead to significant violation of privacy. [85-87] Continuous monitoring of emotions can resemble surveillance, undermining personal autonomy and creating discomfort for individuals unaware of being analyzed. The lack of transparency in emotional data collection and processing intensifies concerns regarding trust in the system [85]. Further research is required to enhance transparency and interpretability.

6.1 Bias in Data

Using biased data to train emotion detection systems raises significant ethical concerns as it directly affects the fairness and reliability of these systems. Extensive data use raises privacy concerns. When datasets lack diversity, systems often fail to recognize or accurately interpret emotional expressions across different cultural, gender, or demographic groups [78], [88-90]. This result in outcomes that disproportionately disadvantage underrepresented populations perpetuates inequality.

Furthermore, biases can reinforce harmful stereotypes, such as associating certain emotions with specific groups. Additionally owning to the generative nature, GPT models can produce harmful or biased content. Cheng *et al* [76] found that the LLMS failed to reproduce the exact inference result, and also the latency of the LLMs was quite high for longer inferences. Guidelines are needed to prevent the generation of harmful or biased content using generative models. Develop ethical guidelines and safeguards to prevent such issues. Ethically [86, 91] curated diverse and representative datasets, ensuring inclusion across cultural, gender, age, and linguistic groups to prevent under representation. Bias audits and fairness metrics should be employed to evaluate system performance across demographics and identify disparities.

6.2 Potential Misuse of Emotional Insights

Emotion detection systems [69, 88] pose significant risks of misuse, particularly when emotional insights are exploited for unethical purposes. In marketing, such systems can be used to manipulate consumer behavior and target vulnerabilities to drive purchases or decisions. In workplaces, continuous emotion monitoring could lead to intrusive surveillance, eroding employee trust and potentially fostering discrimination based on perceived emotional states. Governments and organizations could misuse emotional insights for coercion, propaganda, or social control, undermining individual freedoms and privacy.

Additionally, these systems may amplify biases, unfairly categorizing certain groups based on stereotypes. To mitigate such risks, robust ethical guidelines, transparency, and legal safeguards are essential to ensure that emotional insights are used responsibly and solely for legitimate, consensual purposes. Research [85, 92-95] is required on social bias, cost reduction, privacy preservation issues.

Table 6. Critical Analysis of Affective Computing literature in last 5 years

Ref	Year	Approach	Research Finding	Research Gap
[2]	2024	EmoLeverage model using CMU-MOSEI dataset ,SST-2 and IMDB datasets	SST Accuracy 95 and IMDB Accuracy 95	Scarcity of high quality dataset noted
[75]	2024	Incorporating polarity features in attention on GitHub and Stack Overflow datasets.	F1 score 1.0-10.23% gains over baseline PTMs	Need refinement of contextual gaps in advance empathy
[76]	2024	Multi-Iteration Agentic Workflow and the Multi-Binary-Classifier Agentic Workflow	F1-score 0.16 gains over baseline	Need to focus on reproducing the exact inference result, need to reduce model latency
[77]	2024	ERC and EFR are employed with ensemble methods and transformer approaches.	F1 score = 79%	Need for innovation in designing scalable, accurate, and domain-adaptive sentiment analysis systems.
[73]	2023	Multiple attention mechanisms and Transformer Networks on SemEval-2018 Task-1C	RoBERTa-MA Acuuracy = 62.4% and XLNet-MA outperformed with 45.6% accuracy	There is Need to explore more advanced techniques for addressing class imbalance beyond traditional one.
[80]	2023	Fine-tuning a GPT-3 mode and prompt engineering for zero-shot and few-shot learning using ChatGPT and GPT-4 models.	Micro F1 score = 70% F1 macro = 64%	Need for innovative strategies that balance cost, performance, and scalability, particularly for advanced GPT models.
[13]	2023	11 RoBERTa-large models with annotator-rated and self-reported emotion	F1 score =84%	Lacks generalizability and scalability of emotion detection models
[68]	2023	GPT 3.5,GPT4	GPT 3 Accuracy =75% GPT3.5 Accuracy = 74%	Lack of work on reinforced prompt design
[92]	2023	GPT 3.5 Davinci -003, GPT 4	Accuracy = 97% recall = 91%	Focus is required for social bias, cost reduction, privacy issues
[74]	2023	Bert+CNN,Bert+RNN, Bert+BILSTM	Accuracy = 93%	Explore more comprehensive emotion detection frameworks using advanced transformer models like RoBERTa and hybrid architectures.
[84]	2023	GPT 3.5,GPT 4	GPT 3.5 R2 =49%,GPT4 R2= 74%	It Should address the limitations of GPT models in causal reasoning, emotion intensity prediction, and coping behavior modeling, especially in high goal relevance scenarios.

[93]	2022	Naive Bayes,LSTM,RNN	Accuracy =70%	Faces issues like data imbalance, contextual understanding, and adaptability across diverse datasets.
[94]	2022	DVA-BERT model with sentiment intensity prediction	R2 SCORE =72%	Exploring diverse masking strategies involving multiple sentiment words and experimenting with other attention-based models can further enhance the accuracy sentiment analysis.
[79]	2022	Two-stage Sentiment analysis approach combining zero-shot learning and ensemble classification on SemEval 2017 dataset.	Accuracy= 87.3%	Under-researched languages could provide deeper insights into the robustness and generalizability
[49]	2021	Simple Transformer	F1 Score = 76%, Accuracy = 53%, Recall = 59%, Precision = 76%	Need to develop multimodal emotion detection systems that integrate facial expressions, posture, and contextual information to enhance the accuracy of emotion recognition.
[78]	2020	Ensemble classifier (stacked and weighted ensemble)	SE Accuracy = 97%, WE. Accuracy = 97% TOXIC dataset, SE. Accuracy =87%. WE Accuracy = 84% for Sem2018.	Need for effective text data augmentation methods

7. Conclusion

Affective computing has emerged as a significant field aimed at developing intelligent systems with cognitive, interpretive, and responsive capabilities. Despite substantial progress in emotion detection, several challenges persist, particularly in handling ambiguous language, multiple emotion expressions, and imbalanced datasets. Traditional rule-based models often struggle with scalability and adaptability, limiting their applicability across diverse domains. While various sentiment lexicons and feature extraction techniques have been explored, their effectiveness remains constrained due to the absence of domain-specific corpora. Furthermore, privacy concerns and ethical implications surrounding continuous emotion monitoring highlight the need for more transparent and interpretable models. Addressing these challenges is crucial for advancing affective computing and ensuring its practical and ethical deployment across various fields such as healthcare, education, and security.

8. Future Directions

Developing high-quality, domain-specific datasets is essential for improving model performance. Refining contextual gaps in sentiment analysis can enhance the accuracy of emotion recognition, particularly in empathetic AI applications. Improved annotation processes are needed to ensure reliable and consistent emotion classification. Exploring transformer-based models such as RoBERTa and hybrid architectures can enhance the accuracy of emotion detection. Addressing limitations in GPT models, particularly in causal reasoning, emotion intensity prediction, and coping behavior modeling, is crucial for advancing real-world applications. Transformers, while effective, face challenges due to the lack of high-quality datasets and high computational costs. Research is needed to make transformer models more efficient, cost-effective, and sustainable in terms of computational resources and energy consumption. Refining pre-trained models (PTMs) can improve their ability to capture contextual nuances in emotional intelligence applications. Addressing data imbalance is a critical challenge in affective computing. Developing innovative resampling techniques beyond traditional under-sampling and over-sampling methods can improve model fairness and accuracy. Ensemble learning approaches with optimized sampling ratios tailored to specific datasets can further enhance performance. Need for techniques to Reduce model latency while balancing computational cost for real-time applications. Few-shot learning using generative pre-trained models such as ChatGPT and GPT-4 has inherent limitations, including constrained context windows, cost inefficiencies, and unpredictable model behavior. Research is needed to overcome these constraints, especially in high goal relevance scenarios. Expanding research into under-researched languages can enhance the robustness and generalizability of emotion detection models. Developing multimodal emotion recognition systems that integrate textual, facial, and contextual cues can improve accuracy and adaptability. Additionally, employing effective text data augmentation techniques can strengthen training datasets, leading to more reliable deep-learning models. Ensuring fairness in emotion detection models requires addressing social biases that may arise due to imbalanced training data. Strengthening privacy safeguards and improving transparency in emotional data collection and processing are crucial for maintaining trust in affective computing systems. Developing strategies for reinforced prompt design can enhance the interpretability and reliability of AI-driven emotion recognition. A comparative analysis of affective computing literature over the past five years indicates that existing models, including ChatGPT, face challenges in tasks such as engagement measurement, personality assessment, sarcasm detection, and subjectivity identification.

By addressing these challenges and future directions, affective computing can achieve greater accuracy, adaptability, and ethical integrity, leading to broader applications across industries such as healthcare, education, security, and human-computer interaction.

References

[1] H. Nida, K. Mahira, M. Mudasir, M. Mudasir Ahmed, M. Mohsin, Automatic emotion classifier, in Progress in Advanced Computing and Intelligent Engineering, Advances in Intelligent Systems and Computing, Springer, 713, (2019) 565–572. https://doi.org/10.1007/978-981-13-1708-8_52

[2] E. Parker, W. Nasir, O. Benson, (2024) Enhancing Emotion Detection with Sentiment Analysis Insights. Preprints, 1-11. https://doi.org/10.20944/preprints202404.0059.v1

[3] J. Suttles, N. Ide, (2013) Distant supervision for emotion classification with discrete binary values. In International Conference on Intelligent Text Processing and Computational Linguistics, Berlin, 121-136. https://doi.org/10.1007/978-3-642-37256-8_11

[4] I. Blekanov, M. Kukarkin, A. Maksimov, S. Bodrunova, (2018) Sentiment Analysis for Ad Hoc Discussions Using Multilingual Knowledge-Based Approach. in Proceedings of the 3rd International Conference on Applications in Information Technology, 117-121. https://doi.org/10.1145/3274856.3274880

[5] S. Srivastava, N. Arora, V. Gupta, An Efficient Deep Learning Model using Harris-Hawk Optimizer for Prognostication of Mental Health Disorders. International Research Journal of Multidisciplinary Technovation, 6(4), (2024) 106-123. https://doi.org/10.54392/irjmt2449

[6] O. A. Nasr, A. A. Alajab, A. A. Ahmed, K. Mohiuddin, F. Bin Zafrah, Asynchronous learning techniques during the COVID 19 Pandemic. International Research Journal of Multidisciplinary Technovation 3(2), (2021) 39-48. https://doi.org/10.34256/irjmt2127

[7] J. Zhang, Z. Yin, P. Chen, S. Nichele, Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. Information fusion, 59, (2020) 103-126. https://doi.org/10.1016/j.inffus.2020.01.011

[8] H. Gunes, M. Pantic, Automatic, dimensional and continuous emotion recognition. International Journal of Synthetic Emotions (IJSE), 1(1), (2010) 68-99. https://doi.org/10.4018/jse.2010101605

[9] R.W. Picard, Affective computing. Choice Reviews, 35(9), (1998) 35-5124.

[10] F. Draxler, L. Hirsch, J. Li, C. Oechsner, S.T. Völkel, A. Butz, Flexibility and Social Disconnectedness: Assessing University Students' Well-Being Using an Experience Sampling Chatbot and Surveys Over Two Years of COVID-19. DIS '22: Proceedings of the 2022 ACM Designing Interactive Systems Conference, (2022) 217–231. https://doi.org/10.1145/3532106.3533537

[11] P. Pereira, H. Moniz, J. P. Carvalho, Deep emotion recognition in textual conversations: a survey. Artificial Intelligence Review, 58(1), (2024) 10. https://doi.org/10.1007/s10462-024-11010-y

[12] B.K. Nagaraj, A. Kalaivani, R.S. Begum, S. Akila, H.K. Sachdev, N.S. Kumar, The Emerging Role of Artificial Intelligence in STEM Higher Education: A Critical Review. International Research Journal of Multidisciplinary Technovation, 5(5), (2023) 1–19. https://doi.org/10.54392/irjmt2351

[13] S.J. Lee, J. Lim, L. Paas, H.S. Ahn, Transformer transfer learning emotion detection model: synchronizing socially agreed and self-reported emotions in big data. Neural Computing and Applications, 35(15), (2023) 10945–10956. https://doi.org/10.1007/s00521-023-08276-8

[14] H. Binali, C. Wu, V. Potdar, Computational approaches for emotion detection in text. In 4th IEEE International Conference on Digital Ecosystems and Technologies, IEEE, United Arab Emirates. https://doi.org/10.1109/DEST.2010.5610650

[15] C. Strapparava, R. Mihalcea, SemEval-2007 task

14. in Proceedings of the 4th International Workshop on Semantic Evaluations, (2007) 70-74. https://doi.org/10.3115/1621474.1621487

[16] F.M. Shah, A.S. Reyadh, A.I. Shaafi, S. Ahmed, F.T. Sithil, (2019) Emotion detection from tweets using AIT-2018 dataset. in 2019 5th International Conference on Advances in Electrical Engineering ICAEE, IEEE, Bangladesh. https://doi.org/10.1109/ICAEE48663.2019.8975433

[17] S. Fathy, N. El-haggar, M.H. Haggag, A Hybrid Model for Emotion Detection from Text. International Journal of Information Retrieval Research, 7(1), (2017) 32–48. https://doi.org/10.4018/IJIRR.2017010103

[18] C.K. Chung, J.W. Pennebaker, Linguistic Inquiry and Word Count (LIWC). Applied Natural Language Processing, (2012) 206–229.

[19] F. Å. Nielsen, A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint, 718, (2011) 93–98.

[20] A. Esuli, F. Sebastiani, SentiWordNet: a high-coverage lexical resource for opinion mining. Evaluation, 17(1), (2007) 26.

[21] I.B. Araújo, C.R.N. Brito, I.A. Urbano, V.A. Dominici, M.A. Silva Filho, W.L. Silveira, B.P.G.L. Damasceno, A.C. Medeiros, E. Egito, Similarity between the in vitro activity and toxicity of two different fungizone™/lipofundin™ admixtures. Acta cirurgica brasileira, 20, (2005) 129-133. https://doi.org/10.1590/S0102-86502005000700022

[22] S. Baccianella, A. Esuli, F. Sebastiani, Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In Lrec, 10, (2010) 2200-2204.

[23] A. Esuli, F. Sebastiani, Sentiwordnet: A publicly available lexical resource for opinion mining. In LREC, 6, (2006) 417-422.

[24] L. Vinet, A. Zhedanov, A ‘missing’ family of classical orthogonal polynomials. Journal of Physics A: Mathematical and Theoretical, 44(8), (2011) 085201. https://doi.org/10.1088/1751-8113/44/8/085201

[25] B. Gaind, V. Syal, S. Padgalwar, Emotion Detection and Analysis on Social Media. Global Journal of Engineering Science and Researches, (2019) 78–89. https://doi.org/10.48550/arXiv.1901.08458

[26] J. Posner, J.A. Russell, B.S. Peterson, The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and psychopathology, 17(3), (2005) 715-734. https://doi.org/10.1017/S0954579405050340

[27] K. K. Agustiningsih, E. Utami, M.A. Alsyaibani, Sentiment analysis of COVID-19 vaccines in Indonesia on Twitter using pre-trained and self-training word embeddings. Jurnal Ilmu Komputer dan Informasi, 15(1), (2022) 39-46. https://doi.org/10.21609/jiki.v15i1.1044

[28] M. Z. Asghar, A. Khan, A. Bibi, F. M. Kundi, H. Ahmad, Sentence-level emotion detection framework using rule-based classification. Cognitive Computation, 9, (2017) 868-894. https://doi.org/10.1007/s12559-017-9503-3

[29] G. Sidorov, S. Miranda-Jiménez, F. Viveros-Jiménez, A. Gelbukh, N. Castro-Sánchez, F. Velásquez, I.D. Rangel, S. Suárez-Guerra, A. Treviño, J. Gordon, (2013) Empirical study of machine learning based approach for opinion mining in tweets. In Advances in Artificial Intelligence, Mexico. https://doi.org/10.1007/978-3-642-37807-2_1

[30] S. Mohammad, P. Turney, Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text, (2010) 26-34.

[31] S.M. Mohammad, P.D. Turney, Crowdsourcing a word–emotion association lexicon. Computational intelligence, 29(3), (2013) 436-465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

[32] L. Wang, C. Quan, Y. Bao, F. Ren, (2014) Construction of a Chinese Emotion Lexicon from Ren-CECps. In Intelligent Computing Methodologies: 10th International Conference, ICIC 2014, Springer International Publishing, China. https://doi.org/10.1007/978-3-319-09339-0_63

[33] H. Krishnan, M.S. Elayidom, T. Santhanakrishnan, Emotion detection of tweets using naïve bayes classifier. Emotion, 4(11), (2017) 457-462.

[34] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, M. Stede, Lexicon-Based Methods for Sentiment Analysis. Computational Linguistic, 37(2), (2011) 267–307. https://doi.org/10.1162/COLI_a_00049

[35] D. Seal, U.K. Roy, R. Basak, (2020) Sentence-Level Emotion Detection from Text Based on Semantic Rules. Information and Communication Technology for Sustainable Development, Springer Singapore. https://doi.org/10.1007/978-981-13-7166-0_42

[36] F.A. Acheampong, C. Wenyu, H. Nunoo-Mensah, Text-based emotion detection: Advances, challenges, and opportunities. Engineering Reports, 2(7), (2020) e12189. https://doi.org/10.1002/eng2.12189

[37] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, Lexicon-Based Methods for Sentiment Analysis, Computational Linguistics, 37(2) (2011)

267–307. https://doi.org/10.1162/COLI_a_00049

[38] L. Oneto, F. Bisio, E. Cambria, D. Anguita, Statistical Learning Theory and ELM for Big Social Data Analysis. IEEE Computational Intelligence Magazine, 11(3), (2016) 45–55. https://doi.org/10.1109/MCI.2016.2572540

[39] D. Seal, U. K. Roy, and R. Basak, "Sentence-Level Emotion Detection from Text Based on Semantic Rules," in Information and Communication Technology for Sustainable Development, Advances in Intelligent Systems and Computing, vol. 933, Springer Singapore, (2020) 423–430. https://doi.org/10.1007/978-981-13-7166-0_42

[40] S. Y. M. Lee, Y. Chen, C. Huang, S. Li, Detecting Emotion Causes With A Linguistic Rule-Based Approach 1. Computational Intelligence, 29(3), (2013) 390-416. https://doi.org/10.1111/j.1467-8640.2012.00459.x

[41] S. Shaheen, W. El-Hajj, H. Hajj, S. Elbassuoni, Emotion recognition from text based on automatically generated rules. IEEE International Conference on Data Mining Workshop, IEEE, China. https://doi.org/10.1109/ICDMW.2014.80

[42] U. Gupta, A. Chatterjee, R. Srikanth, P. Agrawal, A Sentiment-and-Semantics-Based Approach for Emotion Detection in Textual Conversations. arXiv. https://doi.org/10.48550/arXiv.1707.06996

[43] U. Orizu, (2018) Implicit Emotion Detection in Text. Aston University.

[44] J. Herzig, M. Shmueli-Scheuer, D. Konopnicki, (2017) Emotion Detection from Text via Ensemble Classification Using Word Embeddings. in Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval, 269–272. https://doi.org/10.1145/3121050.3121093

[45] D. Naresh Kumar, G. Deepak, A. Santhanavijayan, A novel semantic approach for intelligent response generation using emotion detection incorporating NPMI measure. Procedia Computer Science, 167, (2020) 571-579. https://doi.org/10.1016/j.procs.2020.03.320

[46] M. Polignano, P. Basile, M. de Gemmis, G. Semeraro, A comparison of word-embeddings in emotion detection from text using bilstm, cnn and self-attention. In Adjunct publication of the 27th conference on user modeling, adaptation and personalization, (2019) 63-68. https://doi.org/10.1145/3314183.3324983

[47] M.S. Akhtar, D. Chauhan, D. Ghosal, S. Poria, A. Ekbal, P. Bhattacharyya, "Multi-task Learning for Multi-modal Emotion Recognition and Sentiment Analysis. Proceedings of the 2019 Conference of the North, Stroudsburg, PA, USA: Association for Computational Linguistics, arXiv, (2019) 370–379. https://doi.org/10.48550/arXiv.1905.05812

[48] S. Banothu, S. Akula, V. Akarapu, T.R.L. Rao, Emotion Extraction and Classification from Twitter Text. Proceedings of the International Conference on IoT Based Control Networks & Intelligent Systems - ICICNIS 2021, SSRN. https://dx.doi.org/10.2139/ssrn.3884771

[49] W. Graterol, J. Diaz-Amado, Y. Cardinale, I. Dongo, E. Lopes-Silva, C. Santos-Libarino, Emotion Detection for Social Robots Based on NLP Transformers and an Emotion Ontology. Sensors, 21(4), (2021) 1322. https://doi.org/10.3390/s21041322

[50] M. Hasan, E. Rundensteiner, E. Agu, Automatic emotion detection in text streams by analyzing Twitter data. International Journal of Data Science and Analytics, 7, (2019) 35–51. https://doi.org/10.1007/s41060-018-0096-z

[51] A. Chatterjee, U. Gupta, M.K. Chinnakotla, R. Srikanth, M. Galley, P. Agrawal, Understanding Emotions in Text Using Deep Learning and Big Data. Computers in Human Behavior, 93, (2019) 309–317. https://doi.org/10.1016/j.chb.2018.12.029

[52] V. K. Jain, S. Kumar, S. L. Fernandes, Extraction of emotions from multilingual text using intelligent text processing and computational linguistics. Journal of Computational Science, 21, (2017) 316–326. https://doi.org/10.1016/j.jocs.2017.01.010

[53] F. Ghanbari-Adivi, M. Mosleh, Text emotion detection in social networks using a novel ensemble classifier based on Parzen Tree Estimator (TPE). Neural Computing and Applications, 31(12), (2019) 8971-8983. https://doi.org/10.1007/s00521-019-04230-9

[54] W. Wang, L. Chen, K. Thirunarayan, A. P. Sheth, (2012) Harnessing Twitter 'Big Data' for Automatic Emotion Identificatio. International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing, IEEE, Netherlands. https://doi.org/10.1109/SocialCom-PASSAT.2012.119

[55] A. Chatterjee, K.N. Narahari, M. Joshi, P. Agrawal, (2019) SemEval-2019 Task 3: EmoContext Contextual Emotion Detection in Text. in Proceedings of the 13th International Workshop on Semantic Evaluation, Stroudsburg, Association for Computational Linguistics, USA, https://doi.org/10.18653/v1/S19-2005

[56] K.N. Singh, S.D. Devi, H.M. Devi, A. K. Mahanta, A novel approach for dimension reduction using word embedding: An enhanced text classification approach. International Journal of Information Management Data Insights, 2(1), (2022) 100061. https://doi.org/10.1016/j.jjimei.2022.100061

[57] M. Hasan, E. Rundensteiner, E. Agu, (2021) DeepEmotex: Classifying Emotion in Text Messages using Deep Transfer Learning. 2021

IEEE International Conference on Big Data (Big Data), IEEE, USA. https://doi.org/10.1109/BigData52589.2021.9671803

[58] A.B. Zadeh, P.P. Liang, S. Poria, E. Cambria, L.P. Morency, Multimodal language analysis in the wild: Cmu-mosei dataset and interpretable dynamic fusion graph. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, 1, (2018) 2236-2246. https://doi.org/10.18653/v1/P18-1208

[59] X. Wang, W. Jiang, Z. Luo, Combination of convolutional and recurrent neural network for sentiment analysis of short texts. In Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers, (2016) 2428-2437.

[60] P. Zhong, D. Wang, C. Miao, Knowledge-enriched transformer for emotion detection in textual conversations. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, 165–176. https://doi.org/10.18653/v1/D19-1016

[61] J. Piskorski, G. Jacquet, TF-IDF Character N-grams versus Word Embedding-based Models for Fine-grained Event Classification: A Preliminary Study. In Proceedings of the Workshop on Automated Extraction of Socio-political Events from News, (2020) 26-34.

[62] T. Mikolov, E. Grave, P. Bojanowski, C. Puhrsch, A. Joulin, (2017) Advances in pre-training distributed word representations. arXiv preprint arXiv:1712.09405.

[63] M. Mohd, R. Jan, N. Hakak, Enhanced bootstrapping algorithm for automatic annotation of tweets. International Journal of Cognitive Informatics and Natural Intelligence (IJCINI), 14(2), (2020) 35–60. https://doi.org/10.4018/IJCINI.2020040103

[64] M. Kowsher, M.S.I. Sobuj, M.F. Shahriar, N.J. Prottasha, M.S. Arefin, P.K. Dhar, T. Koshiba, An enhanced neural word-embedding model for transfer learning. Applied Sciences, 12(6), (2022). 2848. https://doi.org/10.3390/app12062848

[65] C. Li, W. Xing, Natural Language Generation Using Deep Learning to Support MOOC Learners. International Journal of Artificial Intelligence in Education, 31(2), (2021) 186–214. https://doi.org/10.1007/s40593-020-00235-x

[66] S. Goel, Emotion Classification Using Nature Based Optimization with Transformers And Transfer Learning. Journal of Pharmaceutical Negative Results, (2022) 3052-3071.

[67] V. Kieuvongngam, B. Tan, Y. Niu, (2020) Automatic Text Summarization of COVID-19 Medical Research Articles using BERT and GPT-2. arXiv preprint arXiv:2006.01997.

[68] M.M. Amin, R. Mao, E. Cambria, B.W. Schuller, A wide evaluation of ChatGPT on affective computing tasks. IEEE Transactions on Affective Computing, IEEE, 15(4), (2024) 2204 – 2212. https://doi.org/10.1109/TAFFC.2024.3419593

[69] K. Kheiri, H. Karimi, (2023) Sentimentgpt: Exploiting gpt for advanced sentiment analysis and its departure from current machine learning. arXiv preprint.

[70] J. Kocoń, I. Cichecki, O. Kaszyca, M. Kochanek, D. Szydło, J. Baran, M. Gruza , A. Janz , K. Kanclerz , A. Kocoń , B. Koptyra , W. Mieleszczenko-Kowszewicz , P. Miłkowski , M. Oleksy , M. Piasecki , Ł. Radliński , K. Wojtasik , S. Woźniak, P. Kazienko, ChatGPT: Jack of all trades, master of none. Information Fusion, 99, (2023)101861. https://doi.org/10.1016/j.inffus.2023.101861

[71] M.U. Aytekin, O. Ayhan Erdem, Generative Pre-trained Transformer (GPT) Models for Irony Detection and Classification. 4th International Informatics and Software Engineering Conference (IISEC), IEEE, Turkiye. https://doi.org/10.1109/IISEC59749.2023.10391005

[72] F.M. Plaza-Del-Arco, J. Collado-Montañez, L.A. Ureña-López, M.T. Martín-Valdivia, Empathy and Distress Prediction using Transformer Multi-output Regression and Emotion Analysis with an Ensemble of Supervised and Zero-Shot Learning Models. Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis, Association for Computational Linguistics, (2022) 239–244. https://doi.org/10.18653/v1/2022.wassa-1.23

[73] I. Ameer, N. Bölücü, M.H.F. Siddiqui, B. Can, G. Sidorov, A. Gelbukh, Multi-label emotion classification in texts using transfer learning. Expert Systems with Applications, 213, (2023) 118534. https://doi.org/10.1016/j.eswa.2022.118534

[74] A. Bello, S.C. Ng, M.F. Leung, A BERT Framework to Sentiment Analysis of Tweets. Sensors, 23(1), (2023) 506. https://doi.org/10.3390/s23010506

[75] M.M. Imran, (2024) Emotion Classification in Software Engineering Texts: A Comparative Analysis of Pre-trained Transformers Language Models. NLBSE '24: Proceedings of the Third ACM/IEEE International Workshop on NL-based Software Engineering, 73 – 80. https://doi.org/10.1145/3643787.3648034

[76] L. Cheng, Q. Shao, C. Zhao, S. Bi, G.A. Levow, TEII: Think, Explain, Interact and Iterate with Large Language Models to Solve Cross-lingual Emotion Detection. Proceedings of the 13th

Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, Association for Computational Linguistics, (2024) 495–504. https://doi.org/10.18653/v1/2024.wassa-1.49

[77] N. Wan, S. Au, E. Ubale, D. Krogh, UCSC NLP at SemEval-2024 Task 10: Emotion Discovery and Reasoning its Flip in Conversation (EDiReF). Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024), Association for Computational Linguistics, (2024) 1492–1497. https://doi.org/10.18653/v1/2024.semeval-1.214

[78] G. Haralabopoulos, I. Anagnostopoulos, D. McAuley, Ensemble Deep Learning for Multilabel Binary Classification of User-Generated Content. Algorithms, 13(4), (2020) 83. https://doi.org/10.3390/a13040083

[79] S.G. Tesfagergish, J. Kapočiūtė-Dzikienė, R. Damaševičius, Zero-Shot Emotion Detection for Semi-Supervised Sentiment Analysis Using Sentence Transformers and Ensemble Learning. Applied Science, 12(17), (2022) 8662. https://doi.org/10.3390/app12178662

[80] A. Nedilko, Y. Chu, Team Bias Busters at WASSA 2023 Empathy, Emotion and Personality Shared Task: Emotion Detection with Generative Pretrained Transformers. Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, Stroudsburg, Association for Computational Linguistics, (2023) 569–573. https://doi.org/10.18653/v1/2023.wassa-1.53

[81] M. Alshehri, A. Alamri, A.I. Cristea, C.D. Stewart, Towards Designing Profitable Courses: Predicting Student Purchasing Behaviour in MOOCs. International Journal of Artificial Intelligence in Education, 31(2), (2021) 215–233. https://doi.org/10.1007/s40593-021-00246-2

[82] S.M. Mohammad, F. Bravo-Marquez, E. van der Goot, WASSA-2017 shared task on emotion intensity. Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Association for Computational Linguistics, (2017) 34–49. https://doi.org/10.18653/v1/W17-5205

[83] Wan, J., & Woźniak, M. (2024). A Sentiment Analysis Method for Big Social Online Multimodal Comments Based on Pre-trained Models. Mobile Networks and Applications, 1-14. https://doi.org/10.1007/s11036-024-02303-1

[84] A.N. Tak, J. Gratch, (2023) Is GPT a Computational Model of Emotion?. International Conference on Affective Computing and Intelligent Interaction (ACII), IEEE, USA. https://doi.org/10.1109/ACII59096.2023.103881 19

[85] J.K. Leung, I. Griva, W.G. Kennedy, J.M. Kinser, S. Park, S.Y. Lee, (2023) The Application of Affective Measures in Text-based Emotion Aware Recommender Systems. arXiv Prepr. https://doi.org/10.5220/0012143900003541

[86] S. Stajner, R. Klinger, Emotion Analysis from Texts. Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: Tutorial Abstracts, Stroudsburg, Association for Computational Linguistics, (2023) 7–12. https://doi.org/10.18653/v1/2023.eacl-tutorials.2

[87] M.H. Tuama, W.M. Mashloosh, H.M. Albehadili, M. Alazzawi, M.A. Al-shareeda, Beyond Polarity: The Potential Applications and Impacts of Sentiment Analysis and Emotion Detection. AlKadhum J. Sci., 1(2), (2023) 44–51. https://doi.org/10.61710/akjs.v1i2.51

[88] K. Arjun Raj, M. Afthab Aslam, S. Sreenivasan Panatte, A.K Das, M. Mohyiddeen, Advancements in Emotion Detection: A Comprehensive Review of Text and Audio-Based Approaches. International Journal for Research in Applied Science & Engineering Technology (IJRASET), 12(3), 2875–2879. https://doi.org/10.22214/ijraset.2024.59451

[89] S. M. Mohammad, Ethics Sheet for Automatic Emotion Recognition and Sentiment Analysis. Computational Linguistics, 48(2), (2022) 239–278. https://doi.org/10.1162/coli_a_00433

[90] Z. Chen, Y. Cao, H. Yao, X. Lu, X. Peng, H. Mei, X. Liu, Emoji-powered sentiment and emotion detection from software developers' communication data. ACM Transactions on Software Engineering and Methodology (TOSEM), 30(2), (2021) 1-48. https://doi.org/10.1145/3424308

[91] H. Yu, C. Miao, C. Leung, T.J. White, Towards AI-powered personalization in MOOC learning. npj Science of Learning, 2(1), (2017) 1–5. https://doi.org/10.1038/s41539-017-0016-3

[92] K. Kheiri, H. Karimi, (2023) Sentimentgpt: Exploiting gpt for advanced sentiment analysis and its departure from current machine learning. arXiv preprint.

[93] T. Sharma, M. Diwakar, P. Singh, S. Lamba, P. Kumar, K. Joshi, Emotion Analysis for predicting the emotion labels using Machine Learning approaches. IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), IEEE, India. https://doi.org/10.1109/UPCON52273.2021.9667562

[94] J.L. Wu, W.Y. Chung, Sentiment-based masked language modeling for improving sentence-level valence–arousal prediction. Applied Intelligence, 52(14), (2022) 16353–16369. https://doi.org/10.1007/s10489-022-03384-9

[95] A. Yadav, A. Gupta, An emotion-driven,

transformer-based network for multimodal fake news detection. International Journal of Multimedia Information Retrieval, 13(1), (2024) 7. https://doi.org/10.1007/s13735-023-00315-3

Authors Contribution Statement

Pradeep Kumar: Research problem Supervision, Editing, Validation; Geeta Pattun: Relatable research analysis, research questions analysis and research finding, Writing-Original draft, paper Methodology, Investigation, Conceptualization. All authors have read and approved the final manuscript.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes

About the License