ORIGINAL PAPER

A review on emotion detection by using deep learning techniques

Tulika Chutia[1] · Nomi Baruah[2]

Abstract

Along with the growth of Internet with its numerous potential applications and diverse fields, artificial intelligence (AI) and sentiment analysis (SA) have become significant and popular research areas. Additionally, it was a key technology that contributed to the Fourth Industrial Revolution (IR 4.0). The subset of AI known as emotion recognition systems facilitates communication between IR 4.0 and IR 5.0. Nowadays users of social media, digital marketing, and e-commerce sites are increasing day by day resulting in massive amounts of unstructured data. Medical, marketing, public safety, education, human resources, business, and other industries also use the emotion recognition system widely. Hence it provides a large amount of textual data to extract the emotions from them. The paper presents a systematic literature review of the existing literature published between 2013 to 2023 in text-based emotion detection. This review scrupulously summarized 330 research papers from different conferences, journals, workshops, and dissertations. This paper explores different approaches, methods, different deep learning models, key aspects, description of datasets, evaluation techniques, Future prospects of deep learning, challenges in existing studies and presents limitations and practical implications.

Abbreviations
NLP	Natural Language Processing
PICOC	Population, Intervention, Comparison, Outcomes, Context
CNN	Convolutional Neural Network
LSTM	Long Short Term Memory
BiLSTM	Bidirectional Long Short Term Memory

Tulika Chutia and Nomi Baruah have contributed equally to this work.

✉ Tulika Chutia
  tulikachutiadu24@gmail.com

  Nomi Baruah
  baruahnomi@gmail.com

[1] Computer Science and Applications, Dibrugarh University, Dibrugarh, Assam 786004, India

[2] Computer Science and Engineering, DUIET, Dibrugarh University, Dibrugarh, Assam 786004, India

SEDAT   Sentiment and Emotion Detection in Arabic Text
TF-IDF   Term Frequency-Inverse Document Frequency
ADASYN   Adaptive Synthetic Algorithm
TBED   Text Based Emotion Detection
BERT   Bidirectional Encoder Representations from Transformers
PRISMA   Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SLR   Systematic Literature Review
MSA   Modern Standard Arabic
VAD   Valence,arousal and dominance model
LSVM   Linear Support Vector Machine
SGD   Stochastic Gradient descent
SATD   Sentiment Analysis of Twitter Social Media Data
GRNN   General Regression Neural Network
COA   Chaotic coyote optimization algorithm
BOW   Bag of Words
PMI   Point-wise Mutual Information
SFS   Sequential Forward Selection
MTL   Multi Task Learning
MLEC   Multi-label Emotion Recognition
CC   Classifier Chain
GAN   Generative Adversial Network
ACD   Aspect Category Detection
TL-ERC   Transfer Learning- Emotion Recognition
WS-LSTM   Word-level Sentiment Analysis
SSentiA   Self-supervised Sentiment Analysis: ReviewRC

1 Introduction

Human language understanding and human language generation are the two aspects of natural language processing (NLP). In NLP applications different types of classifiers such as Support Vector Machines (SVM), Naive Bayes (NB), decision tree (DT), k-nearest neighbors (KNN), etc. are used to achieve the result of different domains (Itani et al. 2017). Speech recognition, document summarization, question answering, speech synthesis, machine translation, and other applications all employ NLP (Itani et al. 2017). Itani et al. (2017) also mentioned that these classifiers are also designed and used for different purposes such as question answering, movie sales or reviews, and many other applications. which means that the entity or corpora which they depend on need to be domain specific. The author also mentioned that in the field of sentiment analysis (SA), the classifier aims to identify whether the specific sentence or social media posts are positive, negative, neutral, happy, sad etc. The two critical areas of NLP are SA and emotion recognition (Nandwani and Verma 2021). Both names are used interchangeably, but different from a few angles. The author claims that due to ambiguities in natural language, the human language understanding is more difficult. Emotion is a conscious mental state subjectively experienced as strong feelings usually directed towards a specific object and typically accompanied by physiological and behavioral changes in body language. Textual emotion detection is devoted to automatically identifying emotion states in textual expressions, such as Happy, Sad, Angry etc. For instance, a person experiences joy when they receive good news and a

person experiences fear when they are threatened. Emotions have a strong influence on our lives. We make decisions based on our mind whether we are happy, angry, bored, sad, or frustrated etc.

Since an adequate amount of data is generated in social media which is known as big data, SA, also known as opinion mining was introduced to examine the data steadily and constructively (Nandwani and Verma 2021). There are three types of Sentiment and emotion analysis techniques that are Machine Learning Based Approach, Deep-learning Based Approach, and Lexicon-based Approach which are described in Sect. 3. There are three levels of analysis in SA Word level analysis, Sentence level, and document level analysis. Nandwani and Verma (2021) claimed that SA is a method for detecting the authors' viewpoints on a subject positive, or negative SA implies assessing whether the data is negative, positive, or neutral which assists the business person in understanding their customer's views better so that they may create the necessary changes for their reviews and products. So the authors (Nandwani and Verma 2021) define sentiment analysis as the process of finding the writer's perception/attitude, which may be negative, positive, or neutral. The motivation of SA is to regulate the categorized opinionated text and polarity of the text whether it is positive or negative. It does not only include positive and negative, it may include agreed or disagreed, good or bad (Nandwani and Verma 2021). Basically, the range can be quantified on a 5-point scale: agree, strongly agree, disagree, strongly disagree, or neutral. Nandwani and Verma (2021) also took an example of the Stock market, the investor's decisions are based on the investor's sentiment which can rapidly spread and amplify over the network. Nandwani and Verma (2021) mentioned the applications of NLP in the healthcare sector, where Twitter become a crucial source of information provided by healthcare citizens and professionals. For example: during the COVID pandemic, people use social media to share their thoughts, opinions, and feelings on COVID-19 by posting. They also refer to the applications of NLP in the Education sector. They said that nowadays educational institutions sometimes depend on Social media for marketing and advertisement purposes. The guardians and students act online research and learn more about the institution, courses offered, and professors. Thus the author claimed that by applying SA students can select the best institutions or professors in his admission or registration process.

In 2019, 2.4 billion people used Facebook, the biggest social media site on earth. Over a billion people use WhatsApp, Instagram, Twitter and YouTube, are other social media sites. In 2019, there were at least 3.5 billion people online and 7.7 billion in the world (Ortiz-Ospina and Roser 2023). This indicates that more than two-thirds of all internet users and one in three people worldwide use social media platforms. The world has been altered by social media. Our ability to find partners, access news information, and organize to demand political change is changing as a result of these technologies' quick and widespread adoption (Ortiz-Ospina and Roser 2023). People are using social media to communicate their feelings since Internet services have improved now-a-days. On social media like Facebook, Twitter, Instagram etc., people freely express their feelings, arguments, and opinions on various topics. In addition, many users give feedback and review of various products and services on e-commerce sites. User ratings and reviews on multiple platforms encourage vendors and service providers to enhance their current systems, goods, or services. Now-a-days almost every industry is undergoing some digital transaction, which gives resulting a vast amount of structured and unstructured data. The enormous task for companies is to transform unstructured data into informative one that can help them in decision-making (Nandwani and Verma 2021; Ahmad et al. 2020).

In human life emotions are an inseparable component (Nandwani and Verma 2021). The word "emotion" is derived from the French word emotion means a physical disturbance that came into existence in the seventeenth century. In the nineteenth century, the word emotion was considered as a psychological term. These emotions influence human decision-making and help us communicate to the world in a better way. Emotion detection, also known as emotion recognition, is the process of identifying a person's various feelings or emotions (for example, joy, sadness, or anger). Researchers have been working hard to automate emotion recognition for the past few years. However, some physical activities such as heart rate, shivering of hands, sweating, and voice pitch also convey a person's emotional state (Itani et al. 2017; Kratzwald et al. 2018), but emotion detection from text is quite hard. Emotion analysis plays a critical role in the education sector both for teachers and students. The efficiency of a teacher is decided not only by his academic credentials but also by his enthusiasm, talent, and dedication to his work. Taking timely feedback from students is the most effective technique for a teacher to improve teaching approaches (Sangeetha and Prabha 2021). Darwin and Prodger (1998) mentioned that emotion detection from text is comparatively more complex than SA. "Emotion" is the strong feeling extracted from a person's mood, or circumstances, according to the Oxford Dictionary. According to the Cambridge Dictionary "Emotion" is a strong feeling such as love and anger, or strong feelings in general. Darwin and Prodger (1998) also mentioned the different applications of Emotion Recognition such as systematic e-learning systems based on students' emotions, human-computer interaction, mental health analysis, improving business blueprint, political events, criminal psychology or terrorist mindset after a crime or terrorist attack, performance of chatbot and so on.

This paper is organized into 11 sections. In Sect. 1, we discussed the evolution of emotion research questions and their motivation, terms and terminologies, the contribution of this paper, the SLR diagram (Fig. 5), and inclusion and exclusion criteria with data extraction. In Sect. 2, we mentioned the different emotional models and their classifications. In Sect. 3, we discussed the related work that is done in text-based emotion analysis, key aspects of Machine Learning, and the Deep Learning approach. In Sect. 4, we discussed the existing surveys on contents and topics, surveys on methods of emotion detection, and overall methodologies of text-based emotion detection. In Sect. 5, we discussed the different document types whether it is collected from journals, conference papers, workshop papers, or dissertations. In Sect. 6, we discussed the dataset, data preprocessing, feature extraction, and prospective methodology for emotion detection using Deep learning methods and algorithms. The word level, sentence level, and document level analysis are discussed in Sect. 7. The different deep learning techniques are reviewed in Sect. 8 with their different methodologies. In Sect. 9, we discussed about the different evaluation techniques/metrics. Challenges that were identified in text-based emotion detection (TBED) are discussed in Sect. 10. The results and conclusions, practical implications, future direction of related studies, and limitations of this paper are discussed in Sect. 11.

1.1 Evolution of emotion

After completing some psychological experiments on facial expressions captured in different circumstances on both humans and animals, In 1872, Charles Darwin claimed that

Fig. 1 Evolution of emotion (Sailunaz et al. 2018)
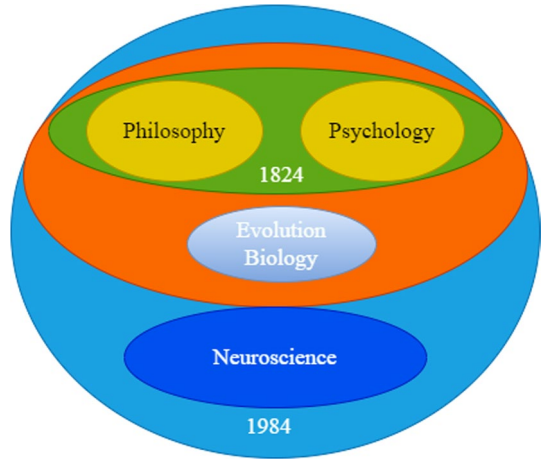
humans and other animals express emotions with similar expressions and behavior under similar circumstances (Darwin and Prodger 1998). His theory of emotion was also related to era and associated circumstances. According to his claims, emotions in humans and other animals were realized over time. He discussed about general principles of emotions in detail; means of expressions of emotions in both humans and animals; causes and effects of all possible emotions like anxiety, grief, dejection, despair, joy, love, devotion, etc.; explanation of emotions with images to show the expressions for particular emotions. Some philosophical and spiritual categorizations of emotions were present before that as described in Bell (1824) and Sailunaz et al. (2018). Figure 1 shows the evolution of emotion in different fields of study.

1.2 PICOC criteria (Widyassari et al. 2022)

See Table 1.

Table 1 PICOC criteria

Population	Emotion detection
Intervention	The method of emotion detection
Comparison	–
Outcomes	Survey of emotion detection
Context	Study in different fields with small and large datasets

1.3 Research question and motivation

See Table 2.

Table 2 Research question and motivation

ID	Research question	Motivation
RQ1	What are journal/Conference papers about emotion detection?	Identifying the most significant journal/conference papers in emotion detection
RQ2	What is the evolution of Emotion Detection?	Identifying the most popular experiments in the evolution of emotion detection
RQ3	What are the datasets that are used in Emotion Detection?	Identifying the most commonly used datasets in emotion detection
RQ4	What is the process of Emotion Detection?	Identifying the most often process used in emotion detection
RQ5	What are the different approaches in Emotion Detection?	Identifying most frequently used methods in emotion detection
RQ6	What are the different levels of Emotion Detection analysis?	Identifying how differently but most frequently discussed levels in emotion detection

1.4 Terms and terminologies

The terms that are used in emotion detection are Tyng et al. (2017):

- *Emotions* Emotions are described as a complex set of interactions between subjective and objective variables mediated by neural and hormonal systems. It is the psychoneural processes involved in controlling the strength and pattern of actions in the dynamic flow of intense behavioral exchange between animals as well as with specific objects important for survival. The term "emotion" represents an "overarching" concept that encompasses affective, cognitive, behavioral, expressive, and physiological changes; emotions are triggered by external stimuli and are associated with a combination of feeling and motivation
- *Moods* Moods last longer than feelings, which are also characterized by positive and negative moods. In contrast, feelings refer to mental experiences that necessarily have good or bad valence and are accompanied by internal physiological changes in the body, especially in the viscera, including the heart, lungs, and intestines, to maintain or restore homeostatic balance. Feelings are generally not caused by emotions
- *Drive* Drive is an inherent program of action necessary for the satisfaction of basic and instinctive (biologically predetermined) physiological needs, e.g., hunger, thirst, libido, exploration, play, and bonding with a partner (Tyng et al. 2017); this is sometimes referred to as "homeostatic drive" In short, a crucial common feature of emotions, moods, and feelings, affects, and drives is their inherent valence, which lies on the spectrum of positive and negative valence (pleasure-unpleasure/good-bad).

1.5 Contribution

The contributions of this paper are as highlighted:

- *Better accuracy in TBED* The deep learning and transformer-based models give better performance in Text-Based Emotion Recognition compared with Machine-Learning models. This paper presents surveys of comprehensive study and applications domains of TBED, whereas the SLR(Systematic Literature Review) diagram also presents a comprehensive and transparent overview of the article.
- *Multimodal Incorporation* The Deep Learning techniques allow the incorporation of textual data with video, audio, and different physiological signals for a better TBED system. With a particular focus on techniques, datasets, applications, key aspects of Machine learning and Deep learning approach, and potential research, we try to give an in-depth overview of TBED.
- *Existing pre-trained models and transfer learning* For fine-tuning the Emotion Recognition task with a large dataset different pre-trained models such as BERT, GPT, etc. are used. So, This paper discusses the different pre-training models in Sect. 11 and also evaluates the different evaluation matrices, various techniques, methods, approaches, and levels of analysis in TBED. We also summarize the key research findings, and trends in our field, which will make it easier for readers to follow the track. Additionally, we also provide a comprehensive description and summary of different publicly available datasets that can be used to support research in that domain.
- *Accountable and illustrating* The Deep Learning models provide a personalized area for a definite domain which gives more accurate results and gives discrete differences in different languages and expressions of emotions. We also concentrate on exploring the role of TBED in different application areas and examining the function of TBED.
- *Circumstantial understanding* Text-based emotion recognition depends on circumstantial information that is found in text or sentences. The Deep Learning techniques can capture more precisely for better results of the system. Also, the Transformer model is capable of reporting the context in which a phrase or word is used. We have also discussed various challenges that were identified in existing studies and future directions of the related studies for overcoming the challenges and limitations.

1.6 Inclusion and exclusion criteria

See Table 3.

Table 3 Inclusion and exclusion criteria

Inclusion criteria	Studies that emotion detection in the text includes topics, problems, datasets, techniques, methods, and challenges used. This research consisted of journals and papers from conferences that specifically discussed the emotion detection of text. The publications taken were studies from 2017 to 2023
Exclusion criteria	Studies that do not have experimental results and use datasets that are unclear Studies that discuss topics beyond emotion detection of text

Table 4 Data extraction

Property	Research question
Publication	RQ1
Evolution of emotion detection	RQ2
Emotion detection dataset	RQ3
Emotion detection method	RQ4
Emotion detection process	RQ5
Level of analysis	RQ6
Emotion detection challenges	RQ7

1.7 Data extraction

The data that are extracted from the mentioned research questions are represented in Table 4.

2 Emotion models

The existing emotion models can be divided into two classes that are Categorical and Dimensional (Sailunaz et al. 2018; Calvo and Mac Kim 2013). Table 5 summarizes some basic emotion models used frequently in the study and expresses almost all possible human emotions that are proposed by Ekman (1992), Shaver et al. (1987), Oatley and Johnson-Laird (1987), Plutchik (1980) and Lövheim (2012). Based on distinct emotion labels the categorical emotion model is focused. The categorical model assumes that there are discrete emotion categories. Plutchik (1980) define a set of eight basic bipolar emotions, consisting of a superset of Ekman's and with two additions: TRUST and ANTICIPATION. These eight emotions are organized into four bipolar sets: joy vs. sadness, anger vs. fear, trust vs. disgust, and surprise vs. anticipation and Ekman et al. (1999) concluded that the six basic emotions are anger, disgust, fear, happiness, sadness, and surprise. The

Table 5 Emotion models (Sailunaz et al. 2018)

References	Emotions	Approach	Structure
Calvo and Mac Kim (2013)	Anger,fear, joy, sadness, surprise, disgust	Categorical	NA
Shaver et al. (1987)	Anger, fear, joy, love, surprise, sadness	Categorical	Tree
Oatley and Johnson-Laird (1987)	Anger, anxiety, happiness, sadness, disgust	Categorical	NA
Plutchik (1980)	Acceptance, amazement, aggressiveness, admiration, anger, annoyance, anticipation, apprehension, boredom, awe, contempt, disapproval, disgust, distraction, ecstasy, fear, grief, interest, joy, Sadness, optimism, pensiveness, rage, remorse, loathing, love, serenity, submission, surprise, trust, terror, vigilance	Dimensional	NA
Lövheim (2012)	Anger/rage, shame/humiliation, distress/anguish, contempt/disgust enjoyment/joy, fear/terror, surprise/startle, shame/humiliation, interest/excitement	Dimensional	Cube

dimensional model represents effects in dimensional form. According to Canales and Martínez-Barco (2014) the emotions are distributed in a two-dimensional circular space: valence dimension and arousal dimension, as shown in Fig. 2. The valence dimension indicates how much PLEASANT and UNPLEASANT is an emotion. The arousal dimension differentiates ACTIVATION and DEACTIVATION states (Figs. 3, 4).
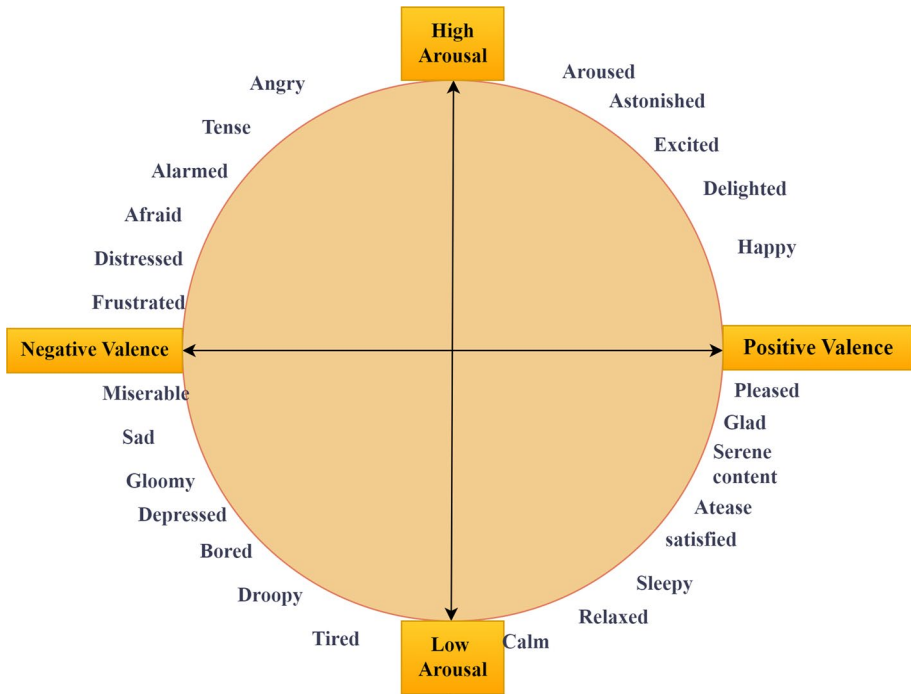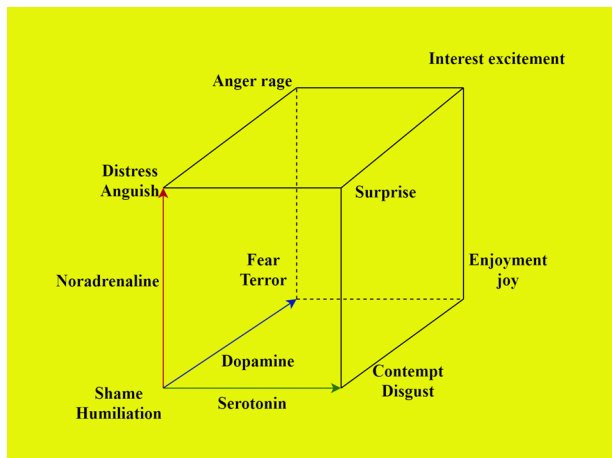

Fig. 2 Circumplex model of affect (Sailunaz et al. 2018)

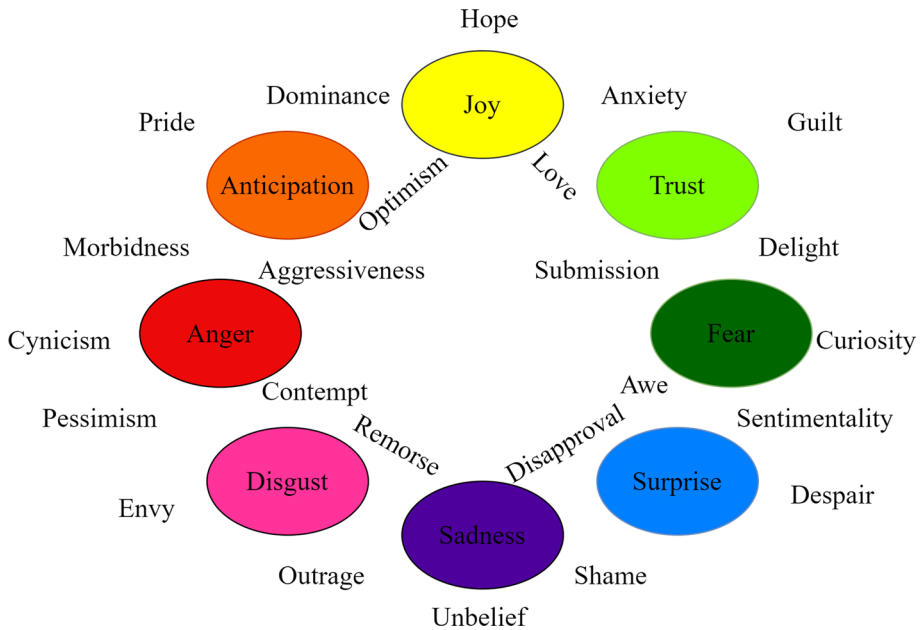**Fig. 3** Lovheim's emotion cube (Sailunaz et al. 2018)

Hope

Dominance   Joy   Anxiety

Pride   Guilt

Anticipation   *Optimism*   *Love*   Trust

Morbidness   Delight

Aggressiveness   Submission

Cynicism   Anger   Fear   Curiosity

Contempt   Awe

Pessimism   Sentimentality

Remorse   *Disapproval*

Disgust   Surprise   Despair

Envy

Sadness

Outrage   Shame

Unbelief

Fig. 4 Plutchik emotion wheel (Sailunaz et al. 2018)

3 Literature review

This section discusses the various studies related to text-based on emotion detection approaches and multiple studies on text mining techniques. Abdullah et al. (2018) is focusing on Arabic international politics and the global economy have led to the investigation of emotion detection. The author described the SEDAT by using Arabic tweets by applying CNN-LSTM. Since the SEDAT is able to determine the existence and the intensity of the emotions in Arabic tweets their real value score is between 0 and 1. From the evaluation, they are trying to gain insight into how SEDAT actually works in each language category from different languages. They split the testing dataset into MSA (Modern Standard Arabic), Egyptian, and Gulf datasets and found that the system performs best with MSA for Anger and Fear emotions, the Gulf dialect with Sadness emotions, and the Egyptian dialect performing best with Joy emotion. The author (Jamal et al. 2021) proposed an IoT-based framework using tweets for emotion detection with a hybrid approach of Term Frequency Inverse Document Frequency (TFIDF). In their work, they collect the raw tweets and filter using tokenization for storing important features without having noise. Secondly, the TFIDF is applied to estimate the importance of features both locally and globally, and then the adaptive synthetic (ADASYN) is applied to solve the problem of unbalanced classes among different emotion classes. Finally, they developed a deep learning model to predict the emotions with dynamic epoch curves. Murthy and Kumar (2021) discussed the different emotion models and their approaches (keyword-based, corpus-based, rule-based, machine learning approach, deep learning approach, and hybrid approach) and datasets that are used in emotion detection. They also discussed the different evaluation metrics that are used in emotion detection from text.

Kusal et al. (2022) discussed the TBED advancement that provides automated solutions for various fields and they present a systematic literature review of the literature published between 2005 to 2021. The author discussed the different classifiers that are used in TBED which are Machine Learning classifiers (Decision Tree (DT), K Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF), Naive Bayes (NB), Linear Regression (LR), Multinomial Naive Bayes (MNB)) and deep learning classifiers (Convolutional neural network, Recurrent Neural Network, BERT, Bi-LSTM, GRU, and pre-trained models). They concluded their discussion starting from the introduction, evolution, significance, techniques, applications, datasets, research challenges, and future scope.

Their study (Alswaidan and Menai 2020) shows that Chinese is the most dominant language after English in terms of emotion recognition in texts. In addition, there are newly published research results from other languages, including Arabic, Hindi, Indonesian, and Spanish. For emotion recognition in English, some researchers used one or more existing corpora with emotion annotations—Alm, Aman, ISEAR, SemEval-2007, SemEval-2018, and SemEval-2019, to measure the performance of their models. They concluded their study by showing that the hybrid approaches and learning-based approaches using a traditional text representation with distributed word representation outperform the other approaches in benchmark corpora. This paper also identifies the sets of features that lead to the best-performing approaches; highlights the impact of simple NLP tasks such as part-of-speech tagging and parsing on the performance of these approaches, and identifies some open issues.

Liu et al. (2022) discussed in Sect. 1 that deals with the introduction and the research objective. Section 2, discussed recent research on emotion recognition and the various methods used in this research. Sects. 3 and 4 discussed the methods and evaluation results of the automatic emotion recognition system. Finally, Sect. 5 discussed the conclusions drawn from the research results and the results of the feature research. El Hammoumi et al. (2018) discussed the important role of emotion detection in e-learning systems by using the convolutional neural network (CNN). The author proposed the system and tested it with students aged between 8 to 12 years.

Ullah et al. (2019) proposed a methodology named PDGDL (Program Dependence Graph with Deep learning) to identify different programming source codes by applying the exact coding styles' features from each programmer. The author collected the data from Google Code Jam (GCJ) and analyzed 1000 programmers' data, in which C++, Java etc. languages were used. Finally, they conclude that their proposed model defeated the existing techniques from the view of classification accuracy, recall, precision, and f-measure metrics. Bansal et al. (2021); Zhao and Wong (2023); Wang et al. (2023) also discussed SA and generative modeling, Transformer-based and lexicon-based SA.

Since there are many researchers studying emotion detection in text, Table 6 contains the summary of related research on text-based emotion detection using deep learning models. Social media uses a variety of ways to measure feelings, including diagnosing neurological problems, and anxiety in individuals, or assessing public sentiment. Changeable boundaries, language, and interpretation are major obstacles to automatic sentiment recognition principles. To quantify the inclination and provide an emotional class message, a soft classification approach is proposed. A supervised emotional framework is created and tested for text-streaming communication. Offline training and interactive classification approaches are two of the key activities. The first task is to create templates for describing emotions in text responses. The second task is to implement a two-stage framework for detecting live text transmissions for emotion monitoring. In addition, DL-EM uses an online means to assess collective emotions and identify incidents in live text sources. The deep learning model includes various stages such as preprocessing the text data, feature extraction from the text,

Table 6 Summary of related researches on text-based emotion detection using deep learning models

| References	Dataset	Method	Accuracy
El Hammoumi et al. (2018)	Arabic 15050Youtube comments	CNN + Bi-LSTM	83.46
Ullah et al. (2019)	DailyDialog, ISEAR, TEC, Tales—Emotions, EmoInt, Crowdflower, Electoral Tweets, Grounded—Emotions, Emotion cause and SemEval2016	CNN + Bi-LSTM	F1—98.8
Serte et al. (2022)	Twitter Firehose (annotated labels, Happy 109, sad 107, angry 90, and others 1920 (17.62 million tweet conversations)	LSTM	F1—80.79
Ullah et al. (2020)	SemEval-2018(Arabic tweets) contains both training and test data of Joy 952 + 448, angry 1027 + 373, sad 1030 + 370, fear 1028 + 372, and sentiment 1070 + 730	LSTM + CNN	81.8
Rajalingam et al. (2022)	Task 7 of SemEval2020, (micro-edit dataset news headline 15,095, 14,696)	BERT+ Regression	62.29
Clarizia et al. (2018)	James Pannebaker and Laura King's stream of consciousness essay dataset	CNN	62.68
Guo (2022)	Big data	DLSTA	97.22
Karna et al. (2020)	'Emotion classification' dataset with six emotional groups	LSTM, Nested LSTM, SVM	F1—4.1
Majeed et al. (2022)	Corpus of 18k sentences collected from different domains and annotated it with six other classes to concentrate on the emotional polarity of the Roman Urdu text	LSTM, KNN, Decision Tree, SVM, Adaboost, XGboost, MLP, and Random Forest	82.2
Tejaswini et al. (2022)	word2vec, Glove, fasttext	CNN, LSTM RNN	84.6
Aslan (2022)	GAMEEMO dataset	CNN	98.78
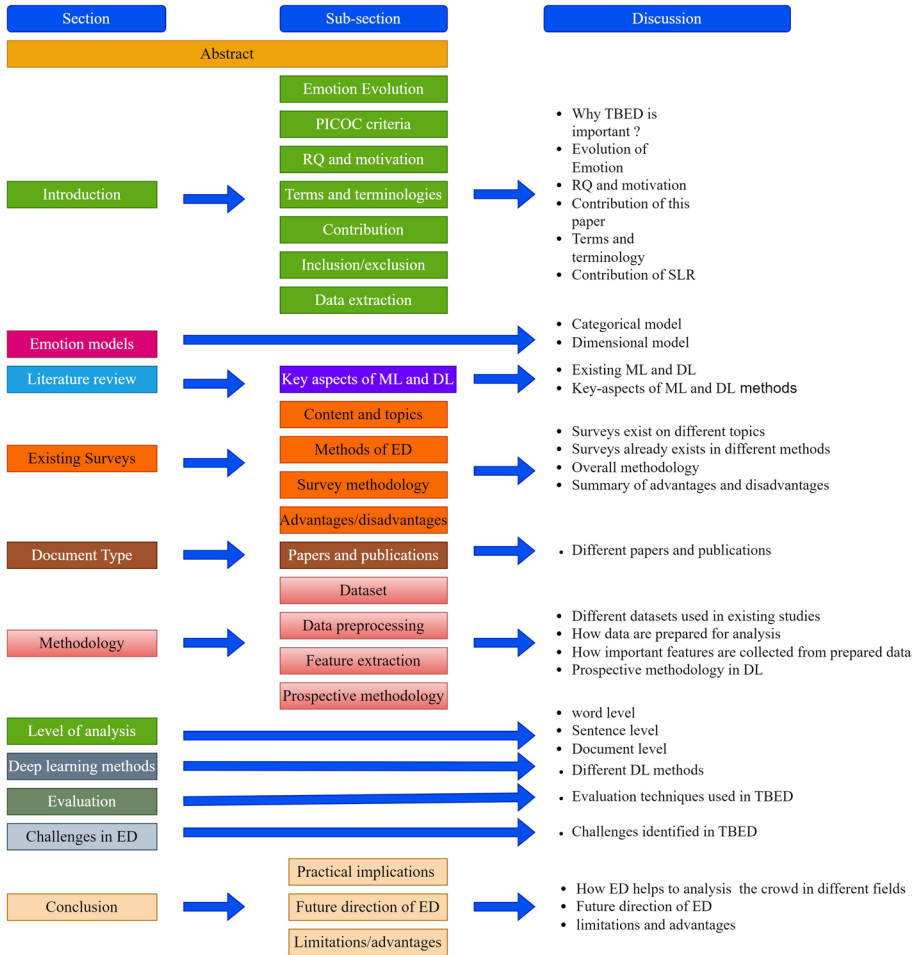
Fig. 5 Systematic literature review on emotion detection in text by using deep learning techniques

feature selection from this extracted emotion feature set, and finally a deep classifier that recognizes the different emotions of the students. In the following section, Sect. 3.1 discusses the Research gap, and Sect. 3.2.1 discusses the key aspects of Machine learning in Emotion Detection, which includes different Supervised, unsupervised, semi-supervised, and hybrid approaches. Section 3.2.2 discusses the key aspects of the Deep Learning Technique in Emotion Detection including different existing methods that were used in different research areas.

3.1 Research gap

- Studies found that the existing research aims only to identify emotion recognition in English literature, which is burdensome to generalize to cultural factors and another low-resource language such as the Assamese Language. The Deep Learning techniques

in NLP research are required to create considering the linguistic conflict, cultural norms and expression, and body language, which can precisely be recognized in cross-cultural and multilingual environments.

- Human emotions are heterogeneous, complex, and bounded with a wide range of different emotional states. In the existing literature, we have seen that many emotion recognition models aim to detect emotion based on Ekman's six basic emotion labels. Emotion Recognition in research needs to explore more emotional labels such as envy, pride, fun, depression, shame, and gratitude etc. This also includes the different models that can accurately give the divergence between the textual data in Text-Based Emotion Recognition.

- With time, a person's state of emotion can be influenced by different experiences, surroundings, events, and interactions. In the existing research on Text-Based Emotion Recognition, the researchers pay no attention to the temporal changing of emotions, only focusing on the static representations of textual data. So, we think that there is a requirement for deep learning models that can accurately give the result of dynamic emotions on textual data.

- Human emotional state is dependent on various factors such as its surroundings, circumstances, etc. So it is very difficult to detect and elucidate emotions when doing so without a large amount of data. In other words, the Deep Learning models are data-dependent. So, research needs Deep Learning models that can sort out the obscurity and grasp the contextual expressions in terms of Text-Based Emotion.

- The Assamese language is a SOV structure language. In the existing literature, we have seen that there is some research work on SA in the Assamese language. Much work has been done in Speech-based Emotion Recognition but Text-Based Emotion Recognition is in the early stage. So, the research needs a Deep Learning Approach/model to identify emotion from Assamese textual data accurately (Fig. 5).

3.2 Key aspects of machine learning and deep learning approaches in emotion detection

3.2.1 Machine learning

Machine learning, an automatic and important classification approach in Emotion detection and SA, which is performed using text features (Kohavi and Provost 1998). The author claimed that the approach can learn from the data and deal with the construction of algorithms. In the machine learning approach the whole dataset is divided into two parts for training and testing data, so it is called training dataset and testing dataset. The information used to train the model by contributing the characteristics of different instances of an item is known as the training dataset and then to check how successfully the model from the training dataset has been trained is known as the testing dataset (Nandwani and Verma 2021). Generally, the Machine learning approach is known as data-driven, since it works on labeled datasets (Anantrasirichai and Bull 2022). Dangi et al. (2021) discussed various Machine Learning techniques based on COVID-19 tweets, in which out of ten popular machine learning techniques, they achieved the highest accuracy of 96.06% in LSVM (Linear Support Vector Machine). The author also claimed that SGD (Stochastic Gradient Descent) obtained better results in all computational parameters.

Supervised, Unsupervised, Semi-supervised, and Reinforcement learning are the four different types of Machine learning classifiers. Dangi et al. (2022b) proposed a model based on five ML models Random Forest, Multinominal NB classifier, Logistic Regression, Decision tree, and Support Vector Machine, and their model is named SATD (Sentiment Analysis of Twitter social media Data). After implementing the model the author achieved an accuracy of 0.96, 0.97, 0.97, 0.98, and 0.98 in Logistic regression, SVM, Random Forest classier, Decision Tree, and Multinominal NB classifier respectively. Dangi et al. (2023) proposed a framework based on sentiment evaluation from Twitter text to detect positive and negative sentiments from the different real-time datasets, namely ARO-RRVFLN (Artificial Rabbits optimized Robust Random Vector Functional Link Network). They used new headline sentiment analysis, sentiment 140, and 4000 short stories datasets for their experiment.

1. *Supervised approach* In the supervised approach, labeled datasets are used for training and testing the supervised classifier. The supervised learning algorithm examines the training data and infers a function, which was used for mapping new examples (Mohri et al. 2018). Since a labeled dataset with emotional tags is necessary, the annotation process is considered as one of the disadvantages as it's going to be a time-consuming and monotonous task (Canales and Martínez-Barco 2014). To check the polarity of emotions whether it is positive, negative, or neutral, initially the root words are used to classify the emotions/sentiments of a text. Since, they may have different domains at different levels, so each domain may have different classifiers with different features. There are many learning algorithms in the Supervised approach:

   - *Linear classifier* Based on the linear combinations value of the characteristics, the linear classifier performs its classification. Let us consider, W is word frequency where, W = w1, w2, w3... and Vector C = c1, c2, c3... linear coefficient vector and S is scalar, the output of the Linear Predictor, also called hyperplane which separates two classes will be LP = W.C + S. The SVM and neural networks come under the Linear classifier. An important supervised learning method is the Support Vector Machine, which is used for classification tasks. The non-probabilistic classifier SVM focuses on determining the best linear separator for classification (Kaur et al. 2017; Rui et al. 2013). The author mentioned that for a given set of training data, where each data labeled belongs to one of the classes, the Support Vector Machine algorithm creates a model that allocates new data into one or two classes. To separate the classes Hyperplane is used. The neuron is the basic element of the Neural Network and has three layers: input layer, hidden layer, and output layer. Bhatti et al. (2004) discussed the neural network approach for human emotion recognition in which they introduced a feature selection method by involving the application of SFS (Sequential Forward Selection) with GRNN (General Regression Neural Network) with a consistency-based selection method. Dangi et al. (2022a) proposed a model known as COA (Chaotic coyote optimization algorithm) based on a time weight-AdaBoost support vector machine, which was used to achieve the classification precisely.
   - *Decision tree classifier*: The Decision tree classifier in which a condition is used to divide the data (Allouch et al. 2018). The decision tree classifier is a tree-like structure and is made up of nodes and branches in which each node denotes an attribute each branch denotes the rules or tests on an attribute, and each leaf node denotes a class label or output (Hasan et al. 2019). There are 3 steps for DT learning: features

choosing, decision tree generation, and pruning (Ghanbari-Adivi and Mosleh 2019). The crucial factors in using a Decision Tree are finding which attributes to employ, what criteria to utilize for splitting, and when to be terminated. During this process, examined all the attributes and different division points are examined and explored. Pruning is a process by which a tree can boost its performance. By using pruning the less important elements can be removed to reduce the complexity, resulting in the increase of predictive power by reducing overfitting (Kusal et al. 2023). Kaur and Vashisht (2012) discussed the emotional variation in adolescent age through data mining (Decision tree technique).

- *Rule-based classifier*: The data space is modeled by a set of rules in a Rule-Based Classifier. The right side presents the class label and the left-hand side presents a condition on the feature set expressed in the disjunctive normal form. Pradhan et al. (2016); Gao et al. (2015) proposed a rule-based method to detect emotion-cause components for Chinese micro-blogs. From the corpus, the emotional lexicon can be constructed automatically and manually.

- *Probabilistic classifier*: The probabilistic classifier is allowed to predict a probability function over a set of classes for input data that are given. In the probabilistic classifier, not only the most likely classes are given but also the probability function over all classes is given (Kaur et al. 2017). The Naive Bayes classifier, Bayesian network, and Maximum entropy come under the Probabilistic classifier.

    (a) *Naive Bayes classifier* Based on Bayes theorem, the Naive Bayes classifier predicts the probability that given a set of features is a part of a particular label in which to extract the feature BOW(Bag of Words) is used (Kaur et al. 2017) in which ignores the position of the word in the document (Kang et al. 2012). The author said that all the features are independent in this model. The author discussed the Senti-Lexicons and improved the Naive Bayes classifier for SA of restaurant reviews. Their experiment showed that when unigram + bigram was used as a feature with this algorithm, the gap between the positive and negative accuracy was reduced to 3.6% as compared to the original Naïve Bayes (Jeyapriya and Selvi 2015). The author proposed a system that extracts aspects of product customer reviews, from each review sentence the nouns and noun phrases are extracted. The author also claimed that to check whether the sentence was a positive or negative opinion and to identify the number of it the NB classifier used a supervised term counting-based approach.

    (b) *Bayesian network* The Bayesian network presumes that all the features are fully dependent resulting in a directed acrylic graph whose edges represent the conditional dependencies and nodes represent the random variables. In text classification this model is considered very expensive therefore it is not frequently used (Kaur et al. 2017; Pradhan et al. 2016).

    (c) *Maximum entropy* Maximum entropy is a probabilistic distribution technique broadly used in text segmentation, POS tagging, and language modeling which prefers uniform models that can satisfy some constraint (Goyal and Tiwari 2017). Maximum entropy converts the labeled feature set to vectors by using encoding. The Maximum Entropy Classifier is also known as the Conditional exponential Classifier. Duric and Song (2012) claimed that the encoded vector was used to calculate the weights of features which can be used to predict labels for each feature set.

2. *Unsupervised approach* As the name suggests "Unsupervised", the data that are used in this approach are not labeled with the classes. This classifier works with several root words for each emotion label which can be cross-referenced to the sentences, resulting in the sentences being classified to corresponding emotions. The author claimed that the supervised classification achieved more accuracy than the unsupervised one but unsupervised learning is more generalized than supervised learning (Sailunaz et al. 2018). Agrawal and An (2012) proposed a novel unsupervised context-based approach to detecting emotions from a text at sentence level analysis. The author considered Ekman's basic emotions and their methodology does not depend on any existing manually crafted affect lexicons such as Word Net-Affect. Goyal and Tiwari (2017) discussed the different unsupervised algorithms such as the Point-wise Mutual Information (PMI) algorithm, Latent Dirichlet Allocation Algorithm, and Random Walk Algorithm. They mentioned that the PMI algorithm was used to automate a system and found consecutive words and their semantic polarity using emoticons to represent positive "thumbs up" and "thumbs down" to represent negative opinions. The author also mentioned that the Latent Dirichlet Allocation Algorithm was domain-independent and lexicon-based and it's a way of automatically discovering topics that sentences contain. Random Walk Algorithm was the automatic construction of a domain-oriented sentiment lexicon in which it simulates a random walk on the graph that reflects four kinds of relationships and that are relationships in words, relationships between documents, relationships from words to documents, and relationships from documents to words.
3. *Semi-supervised approach* Since the supervised approach needs to train a classifier and it depends on more domain cost so the semi-supervised approach was introduced (Go et al. 2009; Goyal and Tiwari 2017). The author proposed a distant supervision approach that prepared the use of training datasets that were auto-generated, where emoticons were used to tag tweets as negative or positive.
4. *Hybrid approach* The lexicon-based approach has a problem of low recall in ML techniques and a problem of domain independence. To avoid both problems, a hybrid approach is introduced (Goyal and Tiwari 2017). The combination of the Machine learning approach and rule-construction into a unified model is called a Hybrid model and the hybrid approach has a higher probability of exceeding the two approaches separately. The author claimed from their survey that in recent advancements, hybrid models implementing unsupervised deep learning methods and standard rules performed better results (Acheampong et al. 2020).

3.2.2 Deep learning

In the recent advancement of the NLP research area, Deep Learning algorithms are more influential than another traditional approaches for emotion detection and SA. Without feature engineering, Deep learning algorithms can extract emotions or sentiments from text streams. There are many Deep learning algorithms namely, RNN, CNN, LSTM, BiLSTM, GRU, and Transformer Models etc. (Nandwani and Verma 2021). Since the deep learning models bring out the features and patterns by themselves, it prevents people from creating the features manually from text streams.

Jian et al. (2010) claimed the goal of SA is to recognize a text whether it is positive or negative automatically. The author used an individual model (i-model) based on an Artificial Neural Network to identify the Sentiment in text, in which the model consists of sentimental features, feature weight, and advanced knowledge. For their experiment, the author

used Cornell Movie review data. The author concluded with their result by mentioning that the accuracy level of the i-model was remarkable as compared to the SVM and Hidden-Markov models.

Arora and Kansal (2019) discussed the text normalization of unstructured data for Twitter sentiment analysis by using Deep CNN. They proposed a model named "Conv-char-Emb" which can control the problem of using less memory space for embedding and noisy data. The author used Deep CNN for embedding since it used fewer parameters in feature representation. Pasupa and Ayutthaya (2019) implemented and compared the results of different deep learning models CNN, LSTM, and Bi-LSTM in the Thai children's tales dataset. The author also claimed that word embedding helps a model to understand the semantics of each word, sentic helps to understand the sentiment of the word and POS-tagging helps the model to identify the different parts of speech of a sentence and pre-processing technique. After implementation, the author achieved a 0.817 F1 score at $p < 0.01$ from CNN model with all three features that are POS-tagging, Thai2Vec, and Sentic, which was comparatively better than among all.

Dashtipour et al. (2020) proposed a hybrid model to extract sentiment analysis in the Persian language. Alqaryouti et al. (2024) proposed a hybrid model of rule-based approach and domain lexicons for SA on government review data. The author concluded their study that their proposed model exceeded other lexicon-based models by 5%. From their result, they mentioned that the aspect extraction accuracy improved remarkably when implicit aspects were considered. Li et al. (2020b) discussed the Transformer-based context and speaker-sensitive model for emotion detection in conversations. Their proposed model namely HiTrans consisting two hierarchical transformers in which the author utilizes low-level transformers as BERT and feeds them into high-level transformers. After experimenting with the model the author claimed that HiTrans outperforms previous state-of-art models. They used three benchmark datasets that were EmoryNLP, MELD, and IEMOCAP.

Fig. 6 Distribution of private and public dataset in TBED
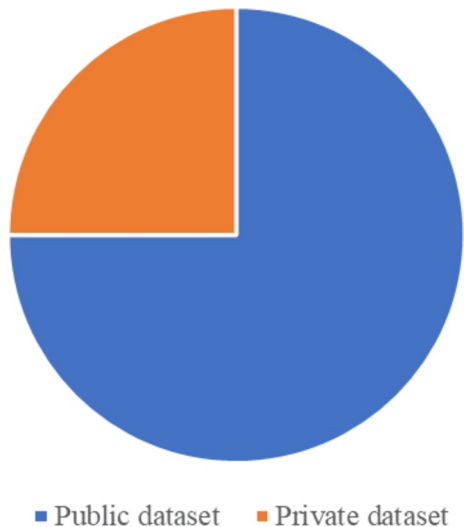
Potamias et al. (2020) proposed a neural network methodology (recurrent CNN RoB-ERTa) that is recently proposed transformer-based network architecture which was further amplified with the recurrent CNN. Their methodology outperforms state-of-art performance under all benchmark datasets. Bansal et al. (2021) proposed a framework that used a deep learning-based classification model to separate the details of tweets from others. The author mentioned that the non-situational tweets consist of grief, sorrow, anger, etc. The author also solved an informative tweet selection task and sentiment classification concurrently using MTL (Multi-Task Learning) in a deep learning framework (Fig. 6).

4 Existing surveys on text-based emotion detection

Emotion detection is a concept encompassing many tasks, such as emotion extraction, emotion classification, emotion summarization, review analysis, depression detection, anger detection, etc. In the existing surveys, the researchers mainly conducted specific analysis of the tasks, technologies, methods, approaches, and application fields involved in the TBED process.

4.1 Surveys on contents and topics of emotion detection

When research on emotion detection was still at its outset, the contents and topics of surveys mainly focused on emotion detection tasks, review analysis and application areas.

Murthy and Kumar (2021) reviewed the emotional models that are categorical approach and dimensional approach. Where they discussed the Circumplex model by Russell, Plutchik's wheel of emotion, and Emotion Space (Valence, Arousal, and Power). From the Parrot's Emotion Taxonomy, they considered the primary emotions are love, joy, surprise, sadness, and anger. The secondary emotions are affection, lust, cheerfulness, zest, contentment, pride, optimism, surprise, irritation, sadness, disappointment, shame, exasperation, rage, disgust, and suffering. They considered the tertiary emotions also. The author reviewed and collected some corpora and lexicons used in research works related to the detection of emotion using text. In their review paper, they discussed different approaches proposed in the literature for the identification of emotions from text Keyword-based approach, Corpus-based approach, Rule-based approach, Machine learning approach, and Deep learning approach. The author also focused on the different evaluation metrics that are used to measure the statistics between the good models that can be it.

Sailunaz et al. (2018) focuses their review on emotion detection from both text and speech. They also discussed keyword-based methods, lexicon-based methods, machine learning methods, and hybrid methods. The author also discussed the datasets that are used in text-based emotion detection, but they mentioned only 5 datasets, where there are only 3 public datasets (ISEAR, EmoBank, The emotion in text dataset) and others are private datasets. They also included some of the classification models that are used in emotion detection are SVM (Support Vector Machine), HMM (Hidden Markov Model), GMM (Gaussian mixture model), Hybrid Systems, and Ensemble Systems. By summarizing basic achievements in the field and highlighting possible extensions for better outcomes, the author concluded their study.

Goyal and Tiwari (2017) defined the emotion recognition from structured and unstructured text. The author defines the tasks that are useful in the recognition of emotions are

discussed that are extraction of features such as syntactic, semantic, feature selection, and linguistic resources available in the market. They also discussed the dictionary-based approach, corpus-based approach, machine learning approach, supervised approach, semi-supervised approach, unsupervised approach, and hybrid approach.

Acheampong et al. (2020)discussed some recent state-of-the-art in the emotion detection field, where the proposals are discussed in relation to the major contributions, approaches employed, datasets used, results obtained, strengths, and weaknesses. The author also discussed the datasets and their features. Some of the datasets that are mentioned in this paper are ISEAR, SemEval, EMOBANK, WASSA-2017 Emotion Intensities (EmoInt) Data, Cecilia OvesdotterAlm's Affect data, Daily Dialogue, AMAN'S emotion dataset, Grounded emotion data, Emotion Stimulus data, crowdsourcing dataset, MELD dataset, Emotion lines, Smile Dataset. Along with strengths and weaknesses, the three main methods used in the design of TBED systems have been clarified. They concluded their study by outlining current issues and potential research directions for text-based emotion detection.

Deng and Ren (2021)reviewed their study on word embedding, architecture, and training level respectively. The author also discussed the challenges and opportunities from four aspects such as the shortage of large-scale and high-quality datasets, fuzzy emotional boundaries, incomplete extractable emotional information in texts, and textual emotion recognition in dialogue. In their review, the author creates a systematic and in-depth overview of deep textual emotion recognition technologies which provides sufficient knowledge and new insights for relevant researchers to understand better the research state, remaining challenges, and future directions in textual emotion recognition.

Rabeya et al. (2017) found the results of their current experimental study on emotion detection by using different text-based data are presented in this paper. The author defined that the emotional lexicons actually affect the state of an emotion in the case of lexicon-based analysis. The main goal of their study is to extract emotions from the Bengali text at the sentence level, and they presented an emotion detection model. Their proposed model detects emotion on the basis of the sentiment of each connected sentence. To demonstrate how frequently people express their emotion in the last or final sentence, a lexicon-based backtracking approach has been developed with 77.16 accuracy. Nandwani and Verma (2021)explains the various emotion models and the process of TBED. The difficulties encountered during emotion analysis are also discussed in this paper.

4.2 Surveys on methods of emotion detection

Goyal and Tiwari (2017)reviewed the work that had been done in identifying emotion detection in text, and argued that although many techniques, methodologies, and models had been created to detect emotions in text, there are various reasons that make these methods insufficient. The author reviewed the resources that are used for detecting emotions in text, which are labeled text, emotion lexicons, and word embeddings. They also discussed the methodologies are supervised approaches and unsupervised approaches. The author discussed the open problems that are mentioned in this review paper. Some of the open problems they discussed are the complex nature of emotion detection, shortage of quality data, and inefficiency of current models. They concluded their study by saying that having a large amount of textual data with the rise of social media in the past couple of years, and therefore the availability of vast self-expression text about any major or minor event, idea, or product, points to great potential to change how entities and organizations can use their future decision-making processes.

Sailunaz et al. (2018) examines the features, drawbacks, and potential future directions of current emotion detection research initiatives, emotion models, datasets, and techniques. They also discussed emotion intensity detection, sarcasm detection, multiple emotion detection, emotion cause detection, personality or mood detection, and emotion versus individual or social parameters. The author concentrated on reviewing studies that examine text-based emotion detection where they looked into different feature sets that were employed in pre-existing methodologies. They also review the fundamental accomplishments made in the area and point out potential improvements.

Al-Omari et al. (2020)proposed a system to detect emotions using deep learning approaches, where the main input of the system is a combination of GloVe word embeddings, BERT Embeddings, and a set of psycholinguistic features (e.g. from Affective Tweets Weka-package). In their approach, they determine the English textual dialogue and classify it into four categories that are Happy, Sad, Angry, and Other. They used the public data of SemEval 2019 and pre-processing the data and in the second step, they extracted the feature vectors. The Network Architecture was built using ensembling with different sub-models such as EmoDense, EmoDet-BERT-BiLSTM, and EmoDet-BiLSTM. The proposed system combined a fully connected neural network architecture and BiLSTM neural network to obtain performance results that show substantial improvements (F1- score 0.748) over the baseline model provided by SemEval-2019/Task- 3 organizers (F1_score 0.58).

4.3 Overall survey methodology

With the increase in the popularity of emotion detection in research, more research results began to accumulate. Researchers need to systematically organize and analyze results from a few number of publications to perform literature reviews. They used different survey methodologies to conduct surveys of a few number of papers.

Cui et al. (2023)divided into six stages that are research question definition, search strategy formulation, inclusion and exclusion criteria definition, quality assessment, data extraction, and data synthesis. The author can eliminate a large number of retrieved papers by using the standard process and finally conducting further analysis and research on a small number of papers.

Yamin et al. (2023)reviewed the systematic literature review (SLR) of the emotion recognition system in Malaysia, where the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) consists of three stages, Identification, Screening, and retrieving. The author analyzed the results as document type, Year of publication, field of research, and most cited.

Kusal et al. (2023) reviewed the current literature published in TBED (Text Based Emotion Detection) between 2005 and 2021 methodically in this paper. The four main research questions were addressed in this review paper, which carefully reviewed 63 research papers from the databases of the IEEE, Science Direct, Scopus, and Web of Science. Additionally, it examines and highlights the various TBED applications in various research fields. They also summarize the various emotion models, methodologies, feature extractions, datasets, and potential research directions that have been provided.

Table 7 Summary of existing deep learning methods in the context of emotion detection, including methods, learning algorithm, datasets, key features, feature generation, model parameters, results, limitations, future work and application domain

References	Methods	Learning algorithm	Datasets	Key features	Feature generation	Model parameters	Results	Limitations	Future work	Application domain
Cahyani et al. (2022)	Deep learning	CNN-BiLSTM	Commuter Line, Transjakarta, Commuter Line and Transjakarta	Developed a system of TBED using a combination of CNN-BiLSTM Compares the use of word embeddings such as Word2Vec and GloVe in building classification models	NA	Convolutional and RELU followed by Maxpool layer that repeats up to two for feature detection. Bidirectional layer with 256 units, with dropout0. A fully connected layer with the activation function RELU is added with 256 numbers of neurons	WordVec-CNN-BiLSTM gives on Commuter Line 84.34%, Transjakarta 83.73%, Commuter Line+Transjakarta 83.88%	NA	Works on BERT and Transfer learning for better performance	NA

Table 7 (continued)

References	Methods	Learning algorithm	Datasets	Key features	Feature generation	Model parameters	Results	Limitations	Future work	Application domain
Ilyas et al. (2023)	Machine learning and Deep learning	SVM, RF, DT, NB, LR, ME, CNN, RNN, LSTM, Bi-LSTM, A-Bi-LSTM, BERT	400,000 sentences from social media (Twitter, YouTube and a local online portal Whatmobile. pk)	RU-EN-Emotion corpus developed	Term Document Matrix, Word-Vec2, GloVe	CNN ReLu activation function has been used in the output layer for all hidden layers. (raring from 0 to 1, where 1 depicts no dropout and 0 means no output from the layer)	RNN (0.753) and CNN of (0.879)	Using Twitter as a data source for scrapping text stems fromats the restriction of using 240 characters, so several variations of emotional sentences that are used on social media were excluded, hence limiting the generability of the results	Increase the quality of the corpus by adding more Emotion labels, Enhancing the dataset by further labeling the intensity of the emotions in each sentence, perform hyper-parameter tuning of deep learning techniques	Social media platform

Table 7 (continued)

References	Methods	Learning algorithm	Datasets	Key features	Feature generation	Model parameters	Results	Limitations	Future work	Application domain
Haryadi and Kusuma (2019)	Deep learning	Nested LSTM	Twitter	NA	TFIDF	Input dimension (502),(882)no of unique word, Output dimension (50), Input length (41), Nested LSTM Layer, Units (50) Dropout(0.2), Depth (2), Output Layer, Units (7), Activation (SoftMax)	99.16%	Dataset was very small and only from one domain	Planning to employ and evaluate other more sophisticated models to find out the best method for emotion detection tested in a more challenging dataset	Twitter
Sharma (2020)	Deep Learning	Bi-LSTM, Random Forest	Hindi text	Polarity of individual class (positive, negative, neutral) extracted	BoW	Sigmoid activation function (0 and 1), hyperbolic tangent activation function	NA	Not calculating the accuracy	NA	Hindi text from different domain

Table 7 (continued)

References	Methods	Learning algorithm	Datasets	Key features	Feature generation	Model parameters	Results	Limitations	Future work	Application domain
Yenter and Verma (2017)	Deep Learning	CNN-LSTM	IMDb	NA	TF-IDF	Base model: Convolution: (Branches/kernel Sizes: 3/5/7/9 Filters: 128 Kernel Regulaizer: L2(0.01)) Activation (type): ReLu max pooling (Pool size): 2 Branch dropout rate: 0.5 Batch Normalization present (no/yes): yes, LSTM units: 128, Merge Dropout rate: 0, Optimizer Type: RMSprop, Learning rate: 0.01, Learning rate decay: 0.1 Accuracy (Max): 0.895	89%	NA	Work on different types of word representations. While the embedding was used, other methods, such as BOW and TF-IDF, could be mixed in to produce different results	Online movie

Table 7 (continued)

References	Methods	Learning algorithm	Datasets	Key features	Feature generation	Model parameters	Results	Limitations	Future work	Application domain
Hofmann et al. (2020)	Cognitive appraisal theory (DL)	CNN	ISEAR	Emotion classification performs better than text-based classification, undue the assumption of the perfect appraisal predictions	NA	CNN with pre-trained GloVe (Glove840B) as a 300-dimensional embedding layer 3 with convolution filter sizes of 2, 3, and 4 with a ReLu activation function, followed by a max pooling layer of length 2 and a dropout of 0.5 followed by another dense layer	NA	They have not been able to reach a substantial improvement in a real-world setting, in which appraisal dimensions are first predicted as a basis for the emotion prediction or in the multi-task setting	Further investigate the combination of appraisal information with emotion prediction, particularly in the light of their oracle ensemble that indicated a clear improvement	Different domain

Table 7 (continued)

References	Methods	Learning algorithm	Datasets	Key features	Feature generation	Model parameters	Results	Limitations	Future work	Application domain
Sasidhar et al. (2020)	Deep Learning	CNN–BiLSTM	Hindi-English code mixed text	Created a classification model in Hinglish	Confusion matrix	Dropout 0.3 Activation ReLU (Conv 1D), Tanh (Bi-LSTM), Softmax (dense layer) Optimizer RMSprop Loss Function Categorical Cross Entropy Batch Size 50	83.21% of accuracy	NA	Extended to a finer level of emotions by creating large corpus in Indian languages code-mixed texts	Different domain

Table 7 (continued)

References	Methods	Learning algorithm	Datasets	Key features	Feature generation	Model parameters	Results	Limitations	Future work	Application domain
Awais et al. (2020)	Deep Learning	LSTM	Novel physiological sensor-based dataset	LSTM based emotion recognition system through multimodal physiological sensors with high performance, IoT based framework to enable URLLC for student learning and healthcare amid pandemics alike Covid-19 etc.	NA	Window: 4 s (800 samples) Overlapping windows: 2 s (400 samples), Total data instances:[19,050 × 800 × 8] (Train: 80%) (Test: 20%) Batch size: [512 × 1000 × 8], h:128 Lambda loss: 0.0015 Learning rate: 0.001 Optimization: Adam	F-Score: 95%	NA	Further extended with focus on end-to-end communications and visual aids to Support distance learning	Health monitoring and Distance learning in the pandemic, student engagement, emotion supports and general well-being

Table 7 (continued)

References	Methods	Learning algorithm	Datasets	Key features	Feature generation	Model parameters	Results	Limitations	Future work	Application domain
Imran et al. (2020)	Deep learning and Sentiment Analysis	DNN, LSTM + FastText, LSTM + GloVe, LSTM + GloVe Twitter, LSTM without pre-trained embedding	Sentiment140 dataset, Tweet Emotion Dataset	Analyzed reactions of citizen from differed culture to Covid19	GLoVec	LSTM without pre-trained embedding: Embedding Layer LSTM (32) with ReLU and Dense Layers 2 with Sigmoid	LSTM without pre-trained embedding: validation accuracy:83% (Highest among all)	Does not assess the performance of CNN and its variants, BERT, GloVe, and GloVe Twitter fail to capture the context, and only the English language is considered	NA	Twitter
Mustakim et al. (2022)	ML, DL, TL	LR, DT, MNB, SVM, CNN, LSTM, BiLSTM, Multilingual BERT, XLM-R	Tamil dataset	NA	TF-IDF	Input length (150), optimizer (Adam), Epoc (5) Batch size (12)	Accuracy (0.47) (XLM)	Accuracy is low	Investigate the impact of resampling on the models	Social media

4.4 Summary of advantages and disadvantages of the existing surveys

In the following, we discuss the advantages and disadvantages of the existing surveys from a number of different points of view.

4.4.1 From the point of view of the contents and topics of emotion detection

As summarized in Table 7, the researcher organized the literature and conducted depth investigations of the contents and topics of emotion detection. They reviewed the tasks of emotion detection analysis (e.g.emotion extraction, emotion detection from suicide note, emotion detection from covid-19 related tweets, depression detection, anger detection, IMDb review), the application areas of emotion detection (e.g. medical diagnosis, psychology, education, employee safety etc.), and different languages for emotion detection, such as Chinese, Arabic, Roman Urdu, German, Spanish, Tamil, Hinglish (Hindi + English), Bangla) (Nandwani and Verma 2021; Majeed et al. 2022; Tejaswini et al. 2022; Rabeya et al. 2017; Alwehaibi et al. 2022; Seyeditabari et al. 2018; Mao et al. 2022; Del Vigna12 et al. 2017; Xu et al. 2019a; Saffar et al. 2023; Sosea and Caragea 2020; Sadeghi et al. 2021; Bashir et al. 2023; Buechel et al. 2020). Since Emotion detection is an emerged topic, the knowledge is gradually integrated to it. Since we have covered a short time period for existing surveys, there have been few survey works analyzing the connection between the topics and methods (Rabeya et al. 2017).

4.4.2 From the point of view of the methods of emotion detection

Some researchers reviewed different techniques and methods of emotion detection in different tasks and areas. They analyzed and discussed emotion detection methods based on lexicons, rules, supervised, semi-supervised, unsupervised machine learning methods, as well as deep learning methods like LSTM, Bi-LSTM, BERT, ALBERT, RoBERT, MLP, GRU and other hybrid approaches (Sailunaz et al. 2018; Karna et al. 2020; Goyal and Tiwari 2017; Acheampong et al. 2020; Al-Omari et al. 2020; Yenter and Verma 2017; Cahyani et al. 2022; Sasidhar et al. 2020; Nugroho and Bachtiar 2021; Hassan et al. 2021; Samy et al. 2018; Majeed et al. 2020; Eckel 2009; Polignano et al. 2019). These researchers also mention the advantages and limitations of methods. As summarized in Table 7, even though existing surveys analyze the techniques and methods of emotion detection, which provides good insights into TBED.

4.4.3 From the point of view of the overall survey methodology

The survey methods that are used mainly include the content analysis method and systematic literature review (SLR) method. Cui et al. (2023); Yamin et al. (2023); Kusal et al. (2023) used the SLR method to analyze the content effectively. However, informatics methods have not been used to study the evolution of research methodologies and emotion detection over time. There are not many surveys that use community detection and keyword co-occurrence analysis to examine the relationship between research methods and topics and how they have changed over time. This paper presents a survey on the research methods, topics, and their evolution over time in order to fill any gaps in the existing surveys. It finds the techniques, subjects, and development of emotion detection in this field
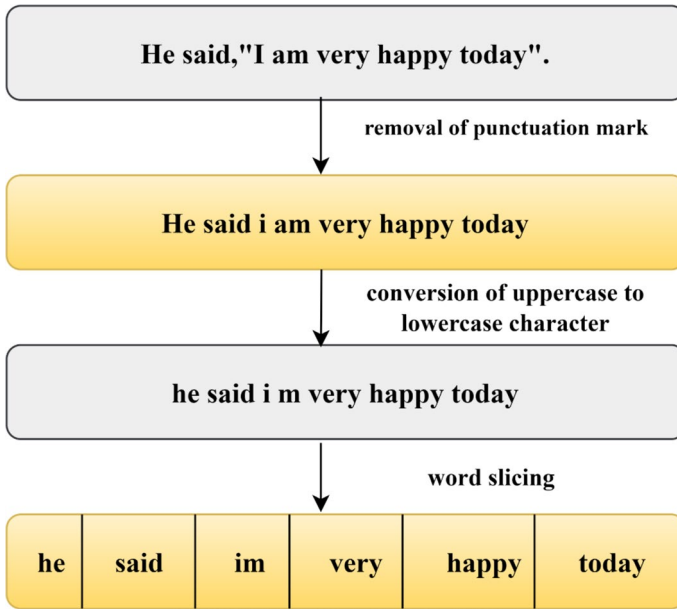
Fig. 7 Three stages of text data preprocessing

from 2017 to 2023 by fusing keyword co-occurrence analysis informatics analysis, and informatics analysis tools (Fig. 7).

5 Document type

The data were classified based on three categories i.e. journal article, conference proceedings article, and dissertation. These kinds of research works are selected for the quality control process normally occurring procedure prior to publication. The process used by the subject matter experts to evaluate the task (Alwehaibi et al. 2022). We have collected some papers that are from journal articles and conference proceedings, where journal papers have a better structure than conference papers.precise (Yamin et al. 2023; Eckel 2009). Figure 8 contains the document type, number of publications, and percentage. The majority of the publications are journal articles; 117 publications or 51.76% while 95 publications or 42.03% are conference proceedings articles; 13 articles from the workshop or 5.75%; and 1 publication or 0.44% from the project category suggesting there is a potential for postgraduate research area (Table 8).

5.1 Paper and publications

We have collected many papers that discuss emotion detection from text. Among all, we have been selecting 226 papers that were published between 2013 to 2023. Figure 13 shows the graph of the development of the paper and publications from year to year for 10 years, from the graph it appears that research on emotion detection is still relevant.

Table 8 Surveys with their advantages and disadvantages

| Existing surveys	Advantages	Disadvantages
From the point of view of the contents and topics of emotion detection	The existing surveys conducted in-depth investigations of the contents and topics of emotion detection. They reviewed the tasks of emotion detection e.g. emotion extraction, emotion detection from suicide notes, emotion detection from covid-19 related tweets, depression detection, anger detection, IMDb review), the application areas of emotion detection (e.g. medical diagnosis, psychology, education, employee safety etc.), and different languages for emotion detection, such as Chinese, Arabic, Roman Urdu, German, Spanish, Tamil, Hinglish (Hindi + English), Bangla). The existing surveys provided useful insights	The existing surveys cover a short time period, and there has been a review dedicated to the evolution of research and the topic of emotion detection. There have also been few survey works analyzing the connection between topics and methods
From the point of view of the method of emotion detection	The existing surveys analyzed different methods of emotion detection. Emotion detection methods based on lexicons, rules, supervised, semi-supervised, and unsupervised machine learning methods, as well as deep learning methods like LSTM, Bi-LSTM, BERT, ALBERT, Ro BERT, MLP, GRU, and other hybrid approaches	Even if we are analyzing the different methods of emotion detection, there have been few review works on TBED. The connection between topics and methods of emotion detection has not been studied
From the point of view of the overall survey methodology	The survey methods used have mainly been the content analysis method and systematic literature review (SLR) method. Kusal et al. (2023), Cui et al. (2023) and Yamin et al. (2023) used the SLR method to analyze the content effectively. However, informatics methods have not been used to study the evolution of research methodologies and emotion detection over time. There aren't many surveys that use community detection and keyword co-occurrence analysis to examine the relationship between research methods and topics and how they have changed over time. This paper presents a survey on the research methods, topics, and their evolution over time in order to fill any gaps in the existing surveys	Since there is a large amount of literature to be surveyed, the use of each content and topic analysis requires more time. There have been few survey methodology works except SLR (Systematic Literature Review) in TBED (Text Based Emotion Detection)
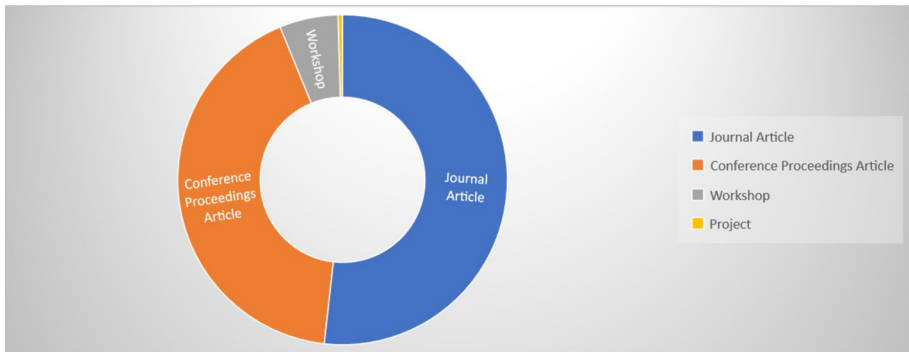
Fig. 8 Document type for selected studies (journal, conference, workshop, and project)

6 Methodology

The process of emotion recognition involves various stages, such as collecting data, pre-processing text, developing feature extraction models, and evaluating. Data collection is the first step to collecting a labeled data set of emotional expressions. This data set should include examples of different emotions such as joy, sadness, anger, etc. The data set can consist of pictures, videos, or audio recordings, depending on the modality of the input.

- *Preprocessing* The collected data must be pre-processed to ensure consistency and compatibility with deep learning models. Preprocessing steps may include resizing images, normalizing pixel values, extracting audio features, or other relevant steps specific to the data type.
- *Model architecture* The next step is to design the architecture of the deep learning model. Convolutional Neural Networks (CNNs) are commonly used for image-based emotion recognition, while Recurrent Neural Networks (RNNs) or variants such as Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRU) are suitable for sequential data such as audio or video. These models can capture spatial and temporal dependencies in the data.
- *Training* Once the model architecture is defined, the next step is to train the model on the labeled dataset. This involves presenting the input data to the model and adjusting its internal parameters (weights and biases) to minimize the difference between the predicted emotions and the actual labels. Training typically uses optimization algorithms such as stochastic gradient descent (SGD) or its variants.
- *Data augmentation* Data augmentation techniques can be used to improve the generalization ability of the model and reduce overfitting. In these techniques, additional training samples are created by applying transformations such as rotations, translations, or adding noise to the input data.
- *Evaluation of the model* After training, the model must be evaluated against a separate validation dataset. Common evaluation metrics for emotion recognition include accuracy, precision, recognition, and F1 score. These metrics measure how well the model performs in classifying emotions.
- *Hyperparameter tuning* Deep learning models often have several hyperparameters that need to be tuned to achieve optimal performance. These hyperparameters include learn-

Fig. 9 Basics steps to perform emotion detection in text
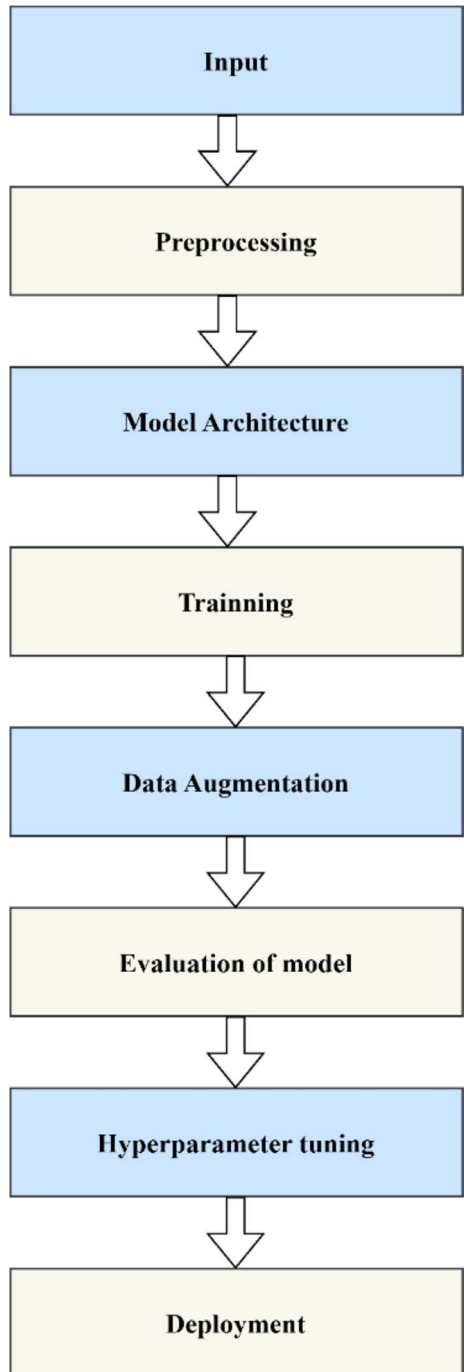
ing rate, stack size, number of layers, number of hidden units, etc. Tuning of hyperparameters is usually done using techniques such as grid search or random search.

- *Deployment* Once the model is trained and evaluated, it can be used for emotion recognition tasks. To do this, the model can be integrated into a larger system or application, such as a real-time emotion recognition system in video surveillance or a chatbot that responds based on user emotions (Fig. 9).

6.1 Dataset

There are different existing and customized datasets are used in TBED according to the types of experiments of different researchers (Sailunaz and Alhajj 2019). Text-based emotion recognition by using Deep learning is data-dependent and relies on a large amount of data. The data that are available for the general public through Google Cloud's public dataset program are called public datasets and the datasets that are customized are called private datasets. This section discussed different publicly available datasets that are used in Emotion detection and sentiment analysis. The data consists of different data sets such as ISEAR, WASSA, and Emotion-stimulus, (having text and emotion as attributes), SemEval, Stanford sentiment treebank (SST) etc. Researchers collect data mainly from tweets, Facebook posts, Instagram, reviews, feedback etc. Data that are collected from social media or e-commerce sites are usually unstructured so it must be processed to make it structured. A dimensional model named valence, arousal dominance model (VAD) is used in the EmoBank data set collected from the news, blogs, letters etc. (Nandwani and Verma 2021). Table 10 contains some of the datasets, data size, respective emotions, and domains that are used in emotion detection from text. The distribution of private and public datasets in text-based emotion detection is represented in Fig. 6.

6.1.1 Datasets that are used in existing studies on text-based emotion recognition

1. EmoBank dataset is collected from Essays, news headlines, blogs, fiction, letters, and travel guides and has six emotion labels that are Anger, disgust, fear, joy, sadness, and surprise. EmoBank dataset contains 10,548 sentences (Buechel et al. 2020).
2. Sem Eval 17 This dataset is collected from Twitter, News headlines, Google news, and other major newspapers and has six emotion labels i.e. Anger, sad, disgust, joy, fear, surprise. It contains a total of 1250 sentences (Ahmad et al. 2020).
3. Crowd flower dataset was collected from different posts on Twitter and consists of various emotions like Boredom, rage, anger, empty, fun, enthusiasm, happiness, love, hate, neutral, sadness, relief, worry, and surprise. Crowd flower dataset contains 40,000 sentences (Seyeditabari et al. 2019).
4. MELD dataset is collected from Dialogues of television shows and consists of various emotions like Anger, disgust, fear, joy, sadness, surprise, neutral, and dataset contains 13,000 sentences (Poria et al. 2018; Sharma et al. 2023; Ghosh et al. 2022; Kharat et al. 2021; Chudasama et al. 2022; Mozhdehi and Moghadam 2023).
5. IMDB dataset is collected from a Movie review. This dataset consists of 2 different emotion labels Positive and negative. The dataset contains a total data size of 25,000 sentences (Sikhi et al. 2022).
6. The GoEmotions dataset is collected from Reddit comments. This dataset consists of different emotion labels like Admiration, amusement, approval, annoyance, anger,

curiosity, confusion, caring, desire, disgust, disapproval, disappointment, embarrassment, excitement, fear, joy, grief, gratitude. The dataset contains a total data size of 58,000 sentences (Demszky et al. 2020; Alvarez-Gonzalez et al. 2021; Ngo et al. 2022; Kane et al. 2022; Kamath et al. 2022; Mahima et al. 2021).

7. ISEAR (The International Survey on Emotion Antecedents and Reactions) dataset. The dataset is constructed by the Swiss National Centre of Competence in Research. ISEAR is lead by Scherer and Wallbott. The dataset consists of seven different emotion labels i.e. joy, sadness, fear, anger, guilt, disgust, and shame. The data are collected from different Cross-cultural studies and contain a total data size of 7665 sentences (Seal et al. 2020; Mac Kim et al. 2010; Adoma et al. 2020b; Poria et al. 2013; Cortiz 2021; Danisman and Alpkocak 2008; Acheampong et al. 2021a).

8. Cecilia Ovesdotter Alm's Affect data. This dataset is collected from different stories and consists of various emotions like Fear, anger, disgust, sadness, happiness, and surprise. Cecilia Ovesdotter Alm's Affect dataset contains 15,000 sentences (Alm et al. 2005; Agrawal and An 2012).

9. Aman emotion dataset collects data from Blogs. This dataset has different emotion labels like No emotion, mixed emotion, Anger, disgust, fear, joy, sadness, and surprise. The dataset contains a total data size of 1466 sentences (Aman and Szpakowicz 2007; Hosseini 2017).

10. Smile dataset, collects data from Twitter. The dataset has different emotion labels which are Anger, disgust, happiness, surprise, and sadness. The dataset contains a total data size of 3085 sentences (Ebling et al. 2018).

11. DailyDialog dataset, collects data from Regular people conversations. This dataset have different emotion labels like Neutral, anger, disgust, fear, happiness, sadness, surprise and it contains 13,118 sentences (Li et al. 2017; Dave and Khare 2021; Lee et al. 2023; Liu et al. 2023).

12. NLPCC dataset, collects data from a Chinese question-answering forum Zhihu. This dataset has different emotion labels like Anger, fear, sadness, surprise, and happiness. The dataset contains a total data size of 7928 sentences (Qiu et al. 2016).

13. EmoInt (WASSA-2017 Emotion Intensities) dataset, collects data from Twitter and has different emotion labels like Anger, fear, joy, sad. The dataset contains a total data size of 7100 sentences (Köper et al. 2017; Lakomkin et al. 2018; Madisetty and Desarkar 2017; Kulshreshtha et al. 2018).

14. Ren-CECps dataset collects data from Blog posts. This dataset has different (Quan and Ren 2010b, 2009) emotion labels like Anger, joy, surprise, anxiety, hate, love, sorrow, expect,and sentiment with intensity. The dataset contains a total data size of 1487 sentences.

15. EEC dataset has different emotion labels like Anger, fear, sadness, surprise, happiness, disgust, love, optimism, pessimism, trust, anticipation, and sentiment with intensity. The dataset contains a total data size of 8640 sentences (Ayub and Wagner 2021).

16. Emotic (Emotions in Context Dataset), Kosti et al. (2017) collects data from people in real environments. This dataset has different emotion labels like Arousal and Dominance. The dataset contains 18,313 images with 23,788 annotated people.

17. Emory NLP dataset, collects data from 97 episodes, 897 scenes, and different utterances. This dataset has different emotion labels like Joy, sad, neutral, mad, scared, powerful, and peaceful. The dataset contains a total data size of 12,600 sentences (He et al. 2021).

18. SEMAINE dataset, Poria et al. (2019); McKeown et al. (2010) collects data from Dialogues. This dataset has different emotion labels like Poppy (happy), Obadiah

Table 9 Summary of some datasets that are used in emotion detection

References	Dataset	Domain	Emotion label	Datasize
Quan and Ren (2009, 2010b)	Ren-CECps	Blog post	Anger, joy, surprise, anxiety, hate, love, sorrow, expect	1487 sentences
Ayub and Wagner (2021)	EEC	NA	Anger, fear, sad, surprise, happiness, disgust, love, optimism, pessimism, trust, anticipation, sentiment with intensity	8640 sentences
Kosti et al. (2017)	Emotion context	People in real environment	Anger, sad, happiness, others	30,000 sentences
He et al. (2021)	Emory NLP	NA	Joy, sad, neutral, mad, scared, powerful, peaceful	12,600 sentences
Duric and Song (2012)	Emotion Lines	NA	Anger, disgust, fear, sad, surprise, neutral, happiness	29,245 labeled utterness
Poria et al. (2019) and McKeown et al. (2010)	SEMAINE	Dialogues	NA	240 sentences
Tripathi et al. (2018), Yu and Kim (2020) and Tripathi and Beigi (2018)	IEMOCAP	different speakers	Anger, disgust, sad, surprise, dimensional, happiness, excitement, frustration, other, neutral	5500 sentences
Nandwani and Verma (2021) and Jia et al. (2020)	SST (Stanford Sentiment Treebank)	Movie reviews	Positive, very positive, negative, very negative, neutral	11,855 sentences
Charoensuk (2018)	Thai Fairy Tales	Children tales	Positive, negative, neutral	1964 sentences
Hasan et al. (2019, 2014)	EmoTex	Twitter	NA	134,100 sentences
Asif et al. (2023), Mishra et al. (2021, b)	Dens	Online narratives and projects	Joy, sadness, anger, fear, anticipation, surprise, love, disgust, neutral	NA
Varshney et al. (2024)	Grounded emotions	Twitter	Happy, sad	2557 sentences
Pasupa and Ayutthaya (2019) and Dashtipour et al. (2020)	Aman	Blog post	Ekman with neutral class	1890 Sentences
Alqaryouti et al. (2024) and Li et al. (2020b)	Alm	185 children stories	Fearfull, surprise, happy, sad, angry - disgust	1207 Sentences

Table 9 (continued)

References	Dataset	Domain	Emotion label	Datasize
Sailunaz et al. (2018)	The Emotion in text dataset	Twitter	Annotated with Anger, Boredom, Empty, Enthusiasm, Fun, Happiness, Hate, Love, Relief, Sadness, Surprise, Worry, Neutral	40,000 tweets
Potamias et al. (2020)	EmoLex	Unigram	Anger, Anticipation, Disgust, Fear, Joy, Sadness, Surprise, Trust	14,182
Ahmad et al. (2020)	SemEval-17	Twitter, News headlines, Google News	Anger,joy,sad,fear	1250 sentences
Kang et al. (2012)	EmoBank	News, blogs, letters	VAD	10,548 sentences
Alswaidan and Menai (2020)	Neviarouskaya Dataset	Stories and Blogs	interest, joy, surprise, sadness, anger, disgust, contempt, fear, shame, and guilt	(10,000 sentences and 700 sentences)
Cahyani et al. (2022)	ISEAR (International Survey on Emotion Antecedents and Reactions)	Incident reports	Guilt, joy, fear, disgust, sadness and shameless	7500 sentences
Ilyas et al. (2023)	Text Affect	Google News	Ekman	1250 sentence
Seyeditabari et al. (2019)	Crowd flower	Twitter	Boredom, rage, anger, empty, fun, enthusiasm, happiness, love, hate, neutral, sadness, relief, worry, surprise	40,000 sentences
Poria et al. (2018), Sharma et al. (2023), Ghosh et al. (2022), Kharat et al. (2021), Chudasama et al. (2022) and Mozhdehi and Moghadam (2023)	MELD	Dialogues from television show	Anger, disgust, fear, joy, sadness, surprise, neutral, positive, negative	13,000 sentences
Sikhi et al. (2022)	IMDB	Movie review	Positive, negative	25,000 sentences

Table 9 (continued)

References	Dataset	Domain	Emotion label	Datasize
Demszky et al. (2020), Alvarez-Gonzalez et al. (2021), Ngo et al. (2022), Kane et al. (2022), Kamath et al. (2022) and Mahima et al. (2021)	Go Emotions	Reddit comment	Admiration, amusement, approval, annoyance, anger, curiosity, confusion, caring, desire, disgust, disapproval, disappointment, embarrassment, excitement, fear, joy, grief, gratitude	58,000 sentences
Ebling et al. (2018)	Smile dataset	Tweet	Anger, disgust, happiness, surprise, sadness	3085 sentences
Li et al. (2017), Dave and Khare (2021), Lee et al. (2023) and Liu et al. (2023)	Daily Dialog	Regular people conversations	Neutral, anger, disgust, fear, happiness, sadness, surprise	13,118 sentences
Qiu et al. (2016)	NLPCC-2018	NA	Anger, fear, sad, surprise, happiness	7928 sentences
Köper et al. (2017), Lakomkin et al. (2018), Madisetty and Desarkar (2017) and Kulshreshtha et al. (2018)	EmoInt dataset (WASSA)	NA	Anger, fear, joy, sad, sentiment with intensity	7100 sentences

(gloomy), Spike (angry) and Prudence (pragmatic). The dataset contains a total data size of 240 sentences

19. IEMOCAP dataset, Tripathi et al. (2018); Yu and Kim (2020); Tripathi and Beigi (2018) collects data from different speakers and have different emotion labels like Anger, disgust, sad, surprise, dimensional, happiness, excitement, frustration, other, neutral. The dataset contains a total data size of 5500 sentences.

20. SST (Stanford Sentiment Treebank) (Nandwani and Verma 2021; Jia et al. 2020) dataset, collects data from Movie reviews. This dataset has different emotion labels like Positive, very positive, negative, very negative, and neutral. The dataset contains a total data size of 11,855 sentences.

21. SemEval dataset, collects data from Twitter, News headlines, Google news, and other major newspapers. This dataset has different emotion labels like Anger, sad, disgust, joy, fear, and surprise. The dataset contains a total data size of 1250 sentences (Chatterjee et al. 2019b; Abdullah and Shaikh 2018; George et al. 2018).

22. Thai fairy tales dataset, Charoensuk (2018) collects data from Children tales. This dataset has three different emotion labels Positive, negative, and neutral. The dataset contains a total data size of 1964 sentences.

23. SS-Tweet dataset (Ahuja et al. 2019; Zhang et al. 2018b; Lynch et al. 2020) collects data from Twitter. This dataset has two different emotion labels i.e. Positive strength, negative strength. The dataset contains a total data size of 4242 sentences.

24. EmoTex dataset (Hasan et al. 2014, 2019) collects data from Twitter and have 134,100 sentences.

25. Text Affect dataset collects data from Google News. This dataset has different emotion labels like Anger, sadness, joy, disgust, surprise, and fear. The dataset contains a total data size of 1250 sentences (Canales and Martínez-Barco 2014; Acheampong et al. 2020; Seyeditabari et al. 2018).

26. Neviarouskaya dataset (Neviarouskaya et al. 2010; Neviarouskaya and Aono 2013) collects data from Blogs, stories. This dataset has emotion labels of interest, joy, surprise, sadness, anger, disgust, contempt, fear, shame, and guilt. The dataset contains a total data size of Dataset1: 1000 sentences and Dataset 2: 700 sentences

27. Dens (Dataset on Emotions using Naturalistic Stimuli) (Mishraa et al. 2021, b) collects data from Online narratives and projects. This dataset has a different emotion labels like Joy, sadness, anger, fear, anticipation, surprise, love, disgust, neutral.

28. Grounded Emotion dataset, Varshney et al. (2024)collects data from Twitter. This dataset has two different emotion labels i.e. Happy, sad. The dataset contains a total data size of 2557 sentences

29. Amazon Alexa dataset (Pandey et al. 2019), collects data from Alexa reviews. The dataset contains a total data size of 2970 sentences (Table 9).

6.2 Data preprocessing

Data preprocessing is a critical step in data cleaning. If the researcher collected the data from social media posts, reviews, blogs etc. There must be a possibility of having unstructured data. In text pre-processing unwanted information is removed from the original text which involves three steps that are removal of punctuation marks, conversion of lowercase characters, and word slicing. Tokenization is a process of breaking down a paragraph into

words or a document into paragraphs or one sentence into words called tokens (Chaffar and Inkpen 2011). Let us consider the sentence "Assam is a beautiful place." After tokenization, the sentence will become "Assam," is, a, "beautiful", "place". The words "is", and "a" will be considered as the stop words so need to be removed from the original text document. The unnecessary word removal is beneficial for finding various aspects from a sentence that are generally described by nouns or noun phrases while emotions are conveyed by adjectives (Sun et al. 2017). Stemming and lemmatization are two crucial steps in preprocessing. In stemming, words are converted to their root form by shortening the suffixes. For example, the words "argues" and 'argue" become "argue". This process reduces the unwanted computation of sentences (Kratzwald et al. 2018; Akilandeswari and Jothi 2018). Lemmatization is a method of grouping different inflected forms of the same word. It's used in computational linguistics, natural language processing (NLP), and chatbots. Lemmatization involves morphological analysis to remove infectious endings from a token or word and transform it into the base word lemma (Ghanbari-Adivi and Mosleh 2019). The authors concluded that removing numbers and lemmatization improved accuracy while removing punctuation did not affect accuracy. In Fig. 7, we are taking a sentence for data preprocessing. Where in the 1st step the punctuation marks are removed, then all uppercase characters are converted into lowercase characters and lastly the sentence is sliced into words called word slicing. Then the sentence/data were ready for feature extraction. Table 10 contains the counting of words from Fig. 7 after word slicing which are ready for feature extraction.

6.3 Feature extraction

Before analyzing the data feature extraction is the major step. The TF-IDF method is widely used in various natural language processing tasks, including emotion classification. By extracting features by using TF-IDF, the model can identify words that carry more weight and contribute significantly to the overall sentiment or emotion expressed in the text. In emotion classification, this method helps in identifying keywords or phrases that are indicative of specific emotions. By assigning higher TF-IDF values to these words, the model can better understand and classify the emotions present in the text (Liu et al. 2022). In the process of word vectorization, a document is first broken down into sentences and each sentence is then further broken down into individual words. This step is crucial in capturing the semantic meaning of the text. Once the text is divided into words, the next step involves building a feature map or matrix. This feature map represents the numerical vectors for each word in the document. The feature map is constructed based on various algorithms and methods, such as the popular Word2Vec or GloVe models. The feature map or matrix allows the machine to understand the relationships and similarities between words so that similar words are represented by vectors that are close to each other in the feature map. This helps the machine to perform tasks such as emotion detection, and retrieval of information. In conclusion, word vectorization or word embedding is a fundamental technique used by machines to understand and process text data. By converting text into numerical vectors, computers can effectively analyze and interpret textual information, leading to various applications in natural language processing and machine learning (Nandwani and Verma 2021).

6.4 Prospective methodology for emotion detection using deep learning, including methods and algorithms

For solving the emotion detection in texts we have discussed the perspective of deep learning methodology in this section. The deep learning methods are considered supervised and semi-supervised, however, they come under unsupervised learning (Kusal et al. 2023). The remaining of this section is organized as follows. Sect. 6.4.1 presents CNN, Sect. 6.4.2 presents RNN, Sect. 6.4.3 presents LSTM, Sect. 6.4.4 presents GRU, Sect. 6.4.5 presents Transformer based model which includes BERT, RoBERTa, ALBERT, DistilBERT, XLNet, ULMFiT, and GPT.

6.4.1 Convolutional neural network (CNN)

CNN, a widely used method in Computer Vision, and NLP applications. The first Deep neural network CNN, with multiple convolutional layers was proposed by Poria et al. (2016). A small chunks of text were applied that correspond to the word in its given context. Zhang et al. (2018a) It is a convolutional layer of a deep feed-forward neural network that feeds data into a fully connected neural network layer. Kusal et al. (2023) mentioned that, the Convolutional in CNN extracts the features by filtering data that are given as input, and multiple filters are combined to get outputs. By using sub-sampling or pooling, layers feature resolution is reduced to enhance the CNN's ability to handle noise and distortion. Different pooling algorithms were used to decrease the number of results while preserving important characteristics and the most commonly used pooling technique was max pooling, which selects the largest factor in the pooling window. Fully connected layers are responsible for carrying out the classification task (Kusal et al. 2023). Jaiswal and Nandi (2020) built a model that can predict human emotion from a real-time image and their model was based on Convolutional Neural Network with 74% of accuracy and tested robustly on 8 different datasets, such as Chicago Face Database, Fer2013, JAFFE Dataset, FEI face dataset, TFEID, CK and CK+, IMFDB, and custom dataset with different angles, faces, backgrounds and age groups.

6.4.2 Recurrent neural network (RNN)

RNN is widely used by researchers for text-mining and text-classifications (Sutskever et al. 2011). RNN is one type of neural network that is particularly suited for modeling sequential information. RNN was configured to recognize the sequential data properties and then anticipate the scenarios using patterns. The author (Abdelgwad et al. 2022) mentioned that the main function of RNN over feed-forward neural networks was its ability to process inputs of any length and remember all information all the time, which was helpful in any time-series prediction. The activation was based on the previous step each time and used recurrent hidden units (Bengio et al. 1994; Pascanu et al. 2013). The RNN (Britz 2015) was composed of feedback loops where the connections between neurons established a directed cycle. Across the network, when gradient decent error was backpropagated, RNN was prone to vanishing the gradient and exploding gradient difficulties. So, the utilization of GRU and LSTM is primarily favored in RNN.

6.4.3 Long short-term memory (LSTM)

Graves and Schmidhuber (2005) mentioned in their study that Hochreiter and Schmidhuber (1997) and Gers et al. (2002) invented LSTM that has been motivated by the flow of error analysis. Robinson (1991) discussed that the layer of LSTM consists of many recurrently connected blocks which are known as memory blocks. Each block contains one or more recurrently connected memory cells. Kusal et al. (2023) mentioned that the input of LSTM was preprocessed to sort for the embedding matrix. There were 200 cells in the next layer of LSTM and in the final layer, there were 128 text-classification and fully connected cells present. The final layer reduced the 128-height vector to a one-height output vector if there were only positive and negative classes to forecast by using the sigmoid activation function. Riza and Charibaldi (2021) discussed TBED using LSTM and Fast-Text by collecting the datasets from Twitter, where they considered 6 emotion classes (happiness, sadness, fear, disgust, anger, surprise). The author compared word embedding with WordVec2, GloVe, and FastText. Awais et al. (2020) discussed emotion detection by using physiological signals which were LSTM-based and mentioned that it is an integrated framework that may enable physiological signals of wireless communication to data processing nodes where LSTM-based emotion detection was performed. Bi-LSTM has two hidden layers (Kusal et al. 2023). Rayhan et al. (2020) mentioned that since BiLSTM processes information in two ways the model forms context for every character within a given text. They combined CNN with Bi-LSTM to identify emotions from Bangla text. Polignano et al. (2019) proposed a classification method based on Bi-LSTM, CNN, and Self-attention which demonstrates its effectiveness on different datasets (ISEAR, SemEval 2018, SemEval 2019). Sumanathilaka et al. (2021) also discussed TBED using Bi-LSTM, where they used three datasets and models that were evaluated from classification accuracy.

6.4.4 Gated recurrent unit (GRU)

Abdelgwad et al. (2022) and Hochreiter and Schmidhuber (1997) mentioned that a powerful and simple alternative of LSTM was GRU, a family of RNNs, which has been proposed to solve the vanishing gradient problem (Kusal et al. 2023). The author (Cho et al. 2014) mentioned that to solve the vanishing gradient problem GRU used the strategy of reset and update. The model decided the amount of earlier data from previous time steps that must be utilized in the future step by using the update gate and the amount of data discarded decided by the reset gate. Liu et al. (2020) discussed TBED using GRU with attention mechanism and emoticon emotion. Aslam et al. (2022) discussed the SA and Emotion Detection by using the Ensemble LSTM-GRU Model on Cryptocurrency-Related Tweets, in which two recurrent neural network applications including GRU and LSTM. Yan et al. (2021) proposed a deep learning model named CNN-BiGRU-AT for identifying the emotional distribution of student text, in which the model combines the convolutional gated recurrent unit with the attention mechanism. Lu and Zhang (2021) proposed a model named AT-BiGRU which was the combination of BiGRU and attention mechanism, which overcomes the problem of complex calculation and cost of high space text of most SA models.

6.4.5 Transformer based model

The recent developments in NLP have illustrated that the transformer-based architecture is more advanced in solving several classification problems irrespective of language discrepancy (Puranik et al. 2021; Li et al. 2021; Sharif et al. 2021). The Deep learning transformer-based models used the process of self-attention, weighing the importance of each portion of the input data differently (Kusal et al. 2023). Acheampong et al. (2021b) discussed the Transformer models for TBED and other NLP tasks, their discussion includes the GPT and its variants, Cross-lingual language model (XLM), Transformer XL, and Bidirectional Encoder Representations from Transformers (BERT). Al-Rfou et al. (2019) mentioned that the transformer models were originally designed for machine translation and used for language modeling, for implementing NLP research areas such as document summarization, question answering, and text classification. They used 64 transformer-based layers and casual attention to foresee the next character in fixed-length input. The model segments input data and learns from each segment separately and is limited by its fixed-sized input length of 512 characters.

Following are the different transformer-based models:

- Transformer-XL Al-Rfou et al. (2019) proposed the the vanilla transformer. To resolve the limitations of the vanilla transformer, Transformer-XL was proposed. Dai et al. (2019) proposed an architecture of fixed-length context of Transformer-XL and accomplished the relative positional encoding and segment-level recurrence. Acheampong et al. (2021b) mentioned that the Transformer-XL was better for longer-term dependencies than RNN and Acheampong et al. (2021b) 's transformer, which was also efficient for short-term dependencies. For audio analysis tasks and language modeling, the model achieved SOTA results and also can be used in different application areas such as; machine translation and audio analysis.
- BERT (Bidirectional Encoder Representations from Transformers) Devlin et al. (2018) said that BERT was a language of transformation methodology launched by Google. BERT obtains substantial explanation of text by considering left to right context equally, which is used to solve different NLP problems and to train the different general-purpose language models. There are two steps involved in BERT that are Pre-training and Fine-tuning steps. In the pre-training phase, BERT trained the unlabeled data. After being initialized with the pre-trained parameters BERT model is then fine-tuned for specific NLP tasks. The cost of fine-tuning of BERT model is considerably lower in terms of system resources (Kusal et al. 2023). Li et al. (2019c) discussed the Aspect-based SA using exploiting BERT. They considered BERT as an Embedding Layer, where they mentioned that the embedding layer of BERT takes a sentence as input and calculates the token-level representations using the information from the whole sentence. The author experiments BERT + Linear, BERT + GRU, BERT + SAN, BERT + TFM, BERT + CRF, from where they achieved highest 61.12 F1 for BERT + GRU. Karimi et al. (2021) proposed a novel architecture named BERT Adversarial Training (BAT) to identify the Aspect Extraction and Aspect Sentiment Classification. Xu et al. (2019b) creates a dataset named RRC (ReviewRC), based on a popular benchmark for SA. The author then takes a look at a novel post-training approach on the BERT model to enhance the act of fine-tuning of BERT for RRC.

- RoBERTa (Robustly optimized BERT technique) RoBERTa was first introduced by Liu et al. (2019); Kusal et al. (2023) a large dataset and tuned hyperparameters for pre-training and a detailed replica of BERT. Kusal et al. (2023)The pre-training masking method was changed from static to dynamic in RoBERTa. The static masking was conducted only once at the time of preprocessing and the dynamic masking was conducted many times during preprocessing step. This model was trained on five different English datasets and that were BookCorpus, the Stories dataset, Wikipedia, the Open Web data, and the CC-News data. Liao et al. (2021) discussed the aspect-category SA in text-stream by using RoBERTa. They proposed a multi-task aspect-category SA model based on RoBERTa and achieved that the proposed model exceeded other models for comparison in aspect-category sentiment analysis. Tan et al. (2022)They proposed a hybrid deep learning model that combines the strength of the sequence model and transformer model while suppressing the obstruction of the sequence model. Their proposed model achieved the F1-score of 93%, 90%, and 91% on the IMDb dataset, Sentiment140 dataset, and Twitter US Airline Sentiment dataset respectively.

- *DistilBERT* Tang et al. (2019) proposed distillation in a neural network in which the author makes an effort to increase the speed of models. DistilBERT uses the early form of BERT by reducing the layers by a factor of 2 and uses the dynamic masking method for better inference. The model was on two datasets as BERT, BooksCorpus dataset, and English Wikipedia. Sanh et al. (2019) introduced a general-purpose pre-trained version of BERT, DistilBERT, 40% smaller, 60% faster, that remains 97% of language understanding ability. Dogra et al. (2021) discussed the DistilBERT for sentiment classification of Banking Financial News, where they found that DistilBERT can transfer basic semantic understanding to further domain which can lead to better accuracy than baseline TF-IDF. The random Forest with DistilBERT gives 78% accuracy.

- *XLNet* In XLNet both Permutation Language Model and Transformer-XL were used and a language model which can be both auto-encoding and auto-regressive (Yang et al. 2019). Before applying the positional encoding and recurrence, the model learned bidirectional context using the Permutation Language Model, so that it is trained on completely different combinations within a sentence. XLNet is unsupervised learning and also addresses the limitations of BERT and GPT (Topal et al. 2021; Adoma et al. 2020a). The author discussed the effectiveness of BERT, RoBERTa, DistilBERT, and XLNet in identifying emotions from text streams. XLNet is pretrained and fine-tuned on English Wikipedia, Clueweb 2012–13, Common-Crawl, BooksCorpus, and Giga 5.

- *ULMFiT (Universal language model fine-tuning for text-classification)*was developed in the year 2018 and is the first "pure" transfer learning application among all. Howard and Ruder (2018) claimed that it can be applied in any task of NLP and the techniques that were key for fine-tuning a language model. The author mentioned that ULMFit consists of the following steps that are, General-domain Language Model pre-training, target task Language Model fine-tuning, and target task classifier fine-tuning. The model was experimented on different datasets TERC-6, Yelp-bi, IMDb dataset, DBpedia, and AG's news (Kusal et al. 2023). Nithya et al. (2022) discussed the Deep learning-based analysis on Code-Mixed Tamil text for Sentiment classification with Pre-Trained ULMFit in which they focused on finding sentiment in Tamil code mixed words, where they found that the deep learning-based Bi-LSTM model with ULMFit embedding gives better results in code-mixed language than other existing models.

- *GPT (Generative pre-trained Transformer series)* Radford et al. (2019) designed GPT as an open AI without recurrent layers and purely attention-based. For standard language modeling purposes pre-training is achieved by integrating the learned position

embeddings with byte-pair encoded tokens and including these embeddings into a cross-layer transformer decoder framework. The author also mentioned that the GPT was formed by the transformer decoder with 12 transformer layers and 12 attention heads, which are primarily used for representing text. These are fine-tuned and trained by pre-processing unlabeled large datasets, such that the BookCorpus dataset is on the restricted supervised dataset. In 2019, Brown et al. (2020) Radford released an enhanced version of GPT, open AI GPT2. Liu et al. (2021); Topal et al. (2021) mentioned that the GPT-3 trained with 175 billion parameters to process human-like text, to provide the largest language model by a large distance, which is an autoregressive language model. In various NLG tasks, GPT-3 gives better performance without fine-tuning or gradient updates. Kheiri and Karimi (2023) evaluated various natural language understanding and generation standards, where the retrieval-based prompt selection method usually overcomes the random baseline. They observed notable gains such as table-to-text generation, 41.9% on the ToTTo dataset, and 45.5% on the NQ dataset as open-domain question answering. Ouyang et al. (2023) discussed the GPT by examining the different GPT methodologies in Sentiment analysis. They also compare the methodologies based on GPT with other current, intuitive models which are previously used with the same dataset. The author concluded their study by saying that the GPT models have expansive benefits in sentiment analysis, which was vitally important for taking action and making sense of enormous user-generated data. Suhaeni and Yong (2023) discussed the GPT-based sentiment analysis on Adversarial Review data generation and detection on the Amazon.com review dataset and a fine-tuned GPT model was implemented. Additionally, they used Sentiment analysis to relate the data novelty and discussed the model's behaviors and some variants of DSA were developed. Rahaman et al. (2023) took two initiatives on the class imbalance in SA through GPT-3-generated synthetic sentences and those were: (1) By fine-tuning the Davinci base model from GPT-3 to generate synthetic review generation and (2) Sentiment classification on balanced and imbalanced dataset by using nine models. Where the author achieved the highest accuracy on Multinominal Naïve Bayes, registering 75.12% on balanced datasets. Conneau and Lample (2019) discussed the how latest version of the GPT-4 and 5 can help with significant advancement in AI-driven NLP tools. The author also claimed that GPT-4 improves the model's training data and speed, which can be evaluated as its overall performance. Additionally, the author also claims that in translating languages, answering questions, and resolving what people think about things, GPT-4 was better than GPT-3.

- *XLM (Cross-lingual language model)* Acheampong et al. (2021b) mentioned that with monolingual data, BERT-based model achieved excellent results in NLP research areas. Lample et al. (2018) mentioned that XLM's or cross-lingual language model enhanced the BERT-based model for translation and classification by pre-training XLM's for NLP which provides MLM extension in BERT by using the model translation method proposed by Kao et al. (2009). Additionally, they have mentioned (Acheampong et al. 2021b) the significance of XLM such that parallel text input is possible with XLM, the model reflects on the bilingual classification task, and pre-trained XLM gives SOTA results.

Table 10 Summary of method, level of analysis, and domain

References	Year	Method	Level of analysis	Domain
Majeed et al. (2022)	2021	LSTM, CNN, KNN, SVM, Adaboost, XGboost, Random Forest, MLP, Decision Tree	Word	Different domains
Tejaswini et al. (2022)	2021	LSTM, CNN	Word	Social media
Dheeraj and Ramakrishnudu (2021)	2020	CNN	Word	Story text
Haryadi and Kusuma (2019)	2019	Nested LSTM, LSTM	Word	Twitter
Nugroho and Bachtiar (2021)	2021	BERT	Word	Indonesian Tweet
Hassan et al. (2021)	2021	BERT	Word	Tweet

7 Level of analysis

There are various levels of analysis in emotion detection, i.e. word level, sentence level, and document level analysis. The keyword level analysis is the first step of TBED where the emotions are extracted from the keyword. It is possible to create emotion labels and related words used in-depth to determine whether a sentence contains any feelings (Ashish et al. 2016). The keyword spotting technique for emotion recognition where a text document is taken as input and output is generated as an emotion class (sad, happy, anger, surprise etc.) (Shirsat et al. 2019). At the sentence-level sentiment analysis, documents or paragraphs are broken down into sentences, and each sentence's polarity is identified (Shivhare and Khethawat 2012). At the document level analysis, the emotions are detected from the entire document. Sailunaz and Alhajj (2019) discussed how the keyword spotting technique works for a given set of strings and substrings. The authors described lexical similarity methods, learning-based methods, and hybrid methods. Various limitations were also considered, including ambiguity in keyword definitions, inability to recognize non-keyword phrases, lack of linguistic information, and difficulty in determining sentiment indices. Murthy and Kumar (2021) reviewed the keyword-based approach as it exploits the knowledge of key features that are combined with emotion labels using a lexicon such as Word-Net Affect and SentiwordNet, where the rules are applied and sentence structures are exploited. Additionally, keyword spotting and emotion intensity are evaluated including with Negation checks. So, it can determine the emotion label for each sentence. Asghar et al. (2017) created generalized and using the latter data where the Twitter user reviews that tailored based on their actions. The preprocessing of the text is based on the keyword-based approach. Seal et al. (2020) proposed a sentence-level emotion detection method, created 25 emotion classes based on keyword analysis, emoticons, short words, keyword negation, a set of proverbs, etc., and achieved 80% accuracy. They also discussed sentence-level emotion recognition from text based on semantic rules. The authors proposed an efficient emotion detection method from sentence-level text by searching for direct emotion words from a predefined emotion keyword database. This method analyzes emotional and phrasal verbs, takes negative words into account, and outperforms the new approach. In their experimental results, they used the C programming language and Python to implement the emotion detection algorithm and achieved 65% accuracy. Cheng et al. (2020) used a sentence-level sentiment detection framework with rule-based classification. Their research focus is on integrating cognitive-based emotion theory and sentiment-analysis-based computational techniques to detect and classify emotions from text. Furthermore, they seek to improve the performance of state-of-the-art techniques by incorporating additional emotion-related signals. Lu et al. (2021) simultaneously created a constraint diagram that encodes the relative polarity between unstructured features to improve the parameter estimation process and proposed a regularization-based framework to achieve better classification performance. Their approach can reveal unique semantic information from unstructured text data. It also justifies the underlying equivalence relationship with the standard Bayesian classifier, which is ideally optimal when the sampling distribution is known. The authors then show that his new method offers better generalization power due to its reduced solution search space and attractive asymptotic consistency. Figure 8 includes an analysis of different levels of text-based emotion recognition collected from 2017 to 2023 and Table 10 contains the deep learning methods level of analysis and respective domain of dataset used.
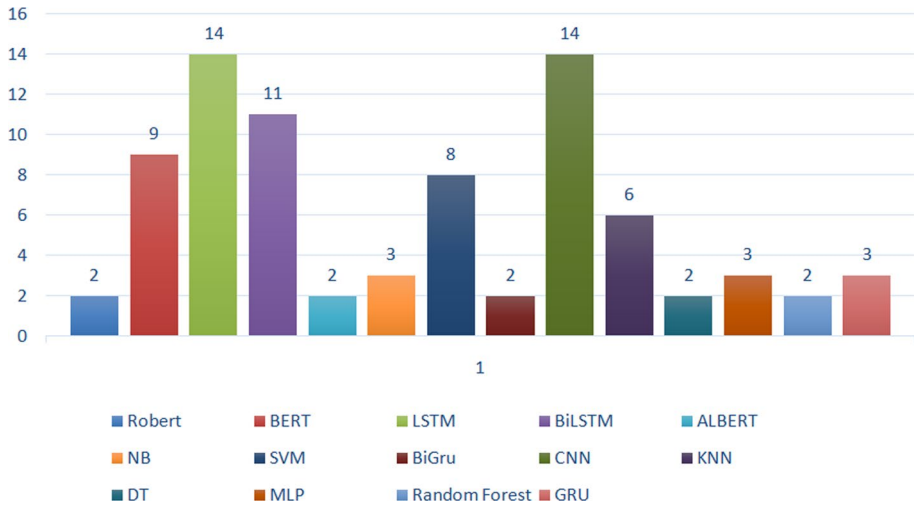
Fig. 10 Distribution of deep learning methods used in TBED

8 Deep learning methods

The method of emotion detection in research that is used there are various kinds of techniques. The distribution methods from 2018 to 2023 for this research are described in Fig. 10. Based on literature studies, the most widely used methods in emotion detection over the past 6 years are LSTM and CNN. The CNN and LSTM are the most favorite because, in the last 6 years of research, both CNN and LSTM are often used to extract or determine the final emotion from the words or sentences. CNN has been widely used and has recently proven to be a powerful tool for emotion detection. It consists basically of a convolution layer to extract features that is followed by a max pooling layer to condense the features output from the convolutional layer. Finally, a fully connected layer applies activation and assigns a prediction (Alwehaibi et al. 2022). Kim (2014) used a two-dimensional convolution neural network to output a two-dimensional array, namely convolution layer, by cross-correlation operation of a two-dimensional input array and two-dimensional convolutional kernel array. After experimenting with several rounds of convolution layer and pooling layer, they are able to extract emotional information from the text.

In Yenter and Verma (2017) used a novel approach through the use of combining kernel from multiple branches of CNN with LSTM layers. After that, they produce a model with the highest reported accuracy on the Internet Movie Database (IMDb) review. In the proposed method they utilize a CNN and an LSTM on the word-level classification of the IMDb review dataset. The method combines versions of the networks from Rahman et al. (2016) and Wadhawan and Aggarwal (2021); the novelty of the proposed network is that it has combined kernels through several branches that accept the data and perform convolution. The output of the CNN branches is fed into an LSTM before being concatenated and sent to a fully connected layer to produce a single, final output. The network is trained and tested in mini-batches between 16 and 128 embeddings. The first layer of the network gives the input scores as a sequence of indices and embeds each word into a vector of a

given size e where a vector of 500 word indices embedded with 32 becomes 500 vectors of length 32. The embedding layer consists of a matrix of trainable weights that generate the vectors for each word index by matrix multiplication. So, during the training, the embedding layer improves upon the embeddings of each word. In convolution, the kernel layer during 1-dimensional convolution is of shape kernel size c specific to that branch. For example, when c = 3, the layer views 3 words at a time and, therefore establishes the sense of 3-word combinations. In the Activation stage, they use a rectified linear unit (ReLU) activation is applied for each branch of CNN layers output. This layer replaces any value having a negative value with zero (0). In the max pooling layer, it converts each kernel size of input into a single output of the maximum observed number. The result is a reduced, scaled-down version of the input. The purpose of this layer is to reduce overfitting while allowing further processing. Similar to the CNN layer, the one-dimensional max-pooling kernel shape is adjusted to the width of the data, so that the kernel size p parameter implies a kernel shape of p times the data width. After max-pooling, each branch goes through a dropout layer that randomly sets some of the inputs to 0. The dropout is applied to a specific d portion of the inputs. The next level for each branch is batch normalization; this level simply normalizes the distribution for each batch after elimination. The purpose of batch normalization is to reduce the internal covariate shift and thus achieve faster convergence. The output of this layer has the same form as the input. This layer is used to prevent overfitting and to generalize the network so that it does not focus on certain portions of the inputs. The shape of the output corresponds to the shape of the input.

Rayhan et al. (2020) used word embeddings are given as input, from which features are extracted and final classification is performed. The embedding layer serves as the first layer used to transfer the word vector representations of selected words of the considered tweet into the model structure. Four convolutional layers connected in parallel receive the output of the embedding layer, followed by a global max-pooling layer to which dropout is applied. Three dense fully connected layers follow, with the last layer responsible for classification. The application of Dropout resulted in better convergence and a smaller difference between training and validation accuracies. One of the most widely used methods in emotion detection is Long - Short Term Memory (LSTM).

Sharma et al. (2023) using the LSTM approaches for in Hindi Text Emotion Recognition. Firstly, they preprocess the emotion dataset which includes tokenization, stop word removal, and lemmatization. After that, the embedding layer is built and fed into one or more LSTM layers. The output is fed into a dense neural network (DNN) with units equal to the number of emotion labels and a sigmoid activation function to perform the classification. Karna et al. (2020) investigates the effectiveness of the deep learning-based Long Short Term Memory mechanism for the identification of textual emotions, which was carried out on the "Emotion Classification" dataset with six emotional groups. Their experimental results proved that LSTM-based text emotion classification provides relatively higher accuracy compared to the existing methods. In their experiment, they provide a matrix with uncertain properties that 1 is for Rage; 2 is for Worry; 3 is for Happiness; 4 is for Affection; 5 is for Sadness; 6 is for Astonished, and 7 is for Kindness, where the LSTM achieves the overall accuracy 94.15%. They are also implementing the nested LSTM and its overall accuracy reaches 92.34%. Jamal et al. (2021) proposed a hybrid approach based on deep learning models TFIDF and S-LSTM. Raw tweets are pre-processed into two different datasets using tokenization techniques to clean up noisy data. TFIDF is used to obtain local and global weight values for each feature. This indicates how important each feature is whether or not it is important for the next step. The length of tweets varies depending

Table 11 Summary of different evaluation metrics, problem domain, used datasets, dataset size, methods/techniques used, accuracy

References	Methods	Datasets	Dataset size	Problem domain	Evaluation metrics/ Feature generation	Performance measure
Tripto and Ali (2018)	Multi-head attention with bidirectional long short-term memory and convolutional neural network (MHA-BCNN)	Webmd and Healthtap online medical healthcare platform	Webmd:2086 Health-tap:5328 questions	Health care	Glove	97.8% of accuracy
Hogenboom et al. (2021)	LSTM and CNN	YouTube	NA	YouTube	Word2Vec	59.2% accuracy
Chowanda et al. (2021)	Graph-based WSD	Historical stock prices of companies listed in NASDAQ-100 and news articles	16,685 news articles	News articles	Precision	0.533 Precision
Jayakrishnan et al. (2018)	Generalised linear model, NB, fast-large margin etc.	Affective Tweets	NA	Tweet	Sentistrength, N-Gram, TF-IDF, Accuracy	97.8% with Adam optimizer
Allouch et al. (2018)	Multi-Layer Neural Network, SVM, Naive Bayes, Decision Tree, and Tree Bagger	From interviewers, article comments, Twitter	9500 sentences	Interviewers, article comments, Twitter	Precision, Recall, F1-Score, Accuracy	SVM method, with 80% recall and precision of more than 75%, and the Multi-Layer Neural Network and the Tree Bagger, with precision and recall of more than 75%
Jose and Chooralil (2015)	WSD with SVM classification	Malayalam Novels	NA	Novels	Accuracy	91.8% of accuracy
Xu et al. (2020)	WSD with WordNet and SentiWordNet classifier	Twitter	NA	Twitter	Accuracy	78.6% of accuracy
Sumanth and Inkpen (2015)	SVM, CNN, RNN	Moud, IEMOCAP	NA	NA	Accuracy	96.75% of accuracy

Table 11 (continued)

References	Methods	Datasets	Dataset size	Problem domain	Evaluation metrics/Feature generation	Performance measure
Cai and Hao (2018)	Senti Wordnet, WSD with RF and SVC	Twitter and SMS, SemEval 2013	3768 Sentences	Social media	Accuracy	58.55% of accuracy
Batbaatar et al. (2019)	Bi-LSTM	Weibo microblogging site, NLPCC 2013	10,000 Weibo texts, the test set contains 4000 Weibo texts	Weibo microblogging	Accuracy	6% increase in macro-average F1 score and an increase of 8% compared with an AVE-BI-LSTM
Tzacheva et al. (2020)	SENN	ISEAR, EmoInt, Tweets etc.	NA	NA	Word2Vec, GloVe, EWE	98.8% accuracy with GloVe + EWE and SENN
Shrivastava et al. (2019)	SVM	Micro-blogging platform	NA	Twitter	Accuracy	88% of accuracy
Mozafari and Tahayori (2019)	CNN, RF, LSTM	TV shows transcript	NA	TV show	Accuracy	80.41% (3 emotion labels) 77.54% (7 emotion labels)
Grover and Verma (2016)	VSM	ISEAR	7666 sentences	Different domain	F1-score, Precision	0.53% Precision
Asghar et al. (2017)	SVM, RF, NB	ISEAR	7666 sentences	Different domain	Accuracy	43.24% on NAVA (Noun, Adverb, Verb, Adjective) text with NB
Chaffar and Inkpen (2011)	NB, DT, SVM	Multiple datasets	NA	Different domain	BOW, N-grams, WordN-etAffects	Accuracy 81.16% on Aman's dataset and 71.69% on Global dataset
Kratzwald et al. (2018)	Sent2Affect, Customer embedding GloVe, LSTM	Election Tweet ISEAR Headlines General tweets, Literary tales	NA	NA	F1-score, GloVe	68.8% on Literary dataset with pre-trained Bi-LSTM

Table 11 (continued)

References	Methods	Datasets	Dataset size	Problem domain	Evaluation metrics/ Feature generation	Performance measure
Perikos and Hatzily-geroudis (2016)	SVM, NB	HC corpora	NA	Punjabi websites of news channels, social networking, blogs	NA	NA
Mashal and Asnani (2017)	Naïve Bayes and a Maximum Entropy learner	ISEAR and the Effective texts datasets	NA	Different domain	NA	NA
Alotaibi (2019)	WordNet	Twitter dataset	140 tweets	Twitter	NA	NA
Chatterjee et al. (2019a)	Logistic regression, SVM, K-Nearest Neighbour (KNN)	ISEAR	7666 sentences	Different domain	Precision, Recall, F-Score	Precision 86%, Recall 84%, F-score 85%
Huang et al. (2019a)	LSTM	SemEval	NA	Twitter, News headlines	NA	NA
Malte and Ratadiya (2019)	BERT	EmotionLines dataset	29,245 utterness	NA	F1 Score	F1 score of 0.815 and 0.885 for the friends
Suhasini and Srinivasu (2020)	Bi-directional transformer	BERT	Facebook multilingual text	NA	Facebook	F1 score 0.4521 for Hindi text, and 0.5520 for English text
LeCompte and Chen (2017)	NB,KNN	Twitter	NA	Twitter	Accuracy	72.06% NB, 55.50% in KNN
Wikarsa and Thahir (2015)	MNB, SVM	Emoji data	54,061	Twitter	Accuracy	Classification accuracy with 0.63785
Nida et al. (2019)	NB	Twitter messages	105 Tweets	Twitter	Accuracy	83%
Jain et al. (2017)	SVM Classifier	Customized datasets	NA	News headlines	NA	NA
Lee and Wang (2015)	SVM	RSS (Rich Site Summary) feeds from Twitter	NA	Political election, healthcare, and sports	NA	NA

Table 11 (continued)

References	Methods	Datasets	Dataset size	Problem domain	Evaluation metrics/ Feature generation	Performance measure
Ye et al. (2023)	Multiview learning	Weibo texts	NA	SNS website from China	F1 score	0.486
Ghanbari-Adivi and Mosleh (2019)	Ensemble classifier, Parzen Tree Estimator (TPE)	OANC, Crowd-Flower, ISEAR	NA	Different domain	Docto Vec	99.49%

on the words used, which can lead to class imbalance issues. When training functions with different categories, there is a risk of class imbalance. The ADASYN method is used to balance each class and oversample some classes. LSTM-based deep learning models take weighted features as input, and these inputs are built on emotional groups such as emotions. B. Anger, sadness, happiness, surprise, etc. are identified. LSTM-based deep learning models are designed to take weight features as inputs and classify these inputs in terms of sentiment classes. Anger, sadness, happiness, surprise, etc. Configured a stack definition of LSTM layers that can interact with each other by processing tweets. Epoch curves are visualized to show each data point's accuracy, effectiveness, and loss. To demonstrate the effectiveness of the proposed method, we extract precision and loss values representing significance. The proposed approach outperformed most state-of-the-art methods with a classification accuracy of 97.5%.

In Rayhan et al. (2020), the authors used bidirectional long-short-term memory with attentional layers in emotion recognition to improve prediction accuracy. They used his ISEAR dataset with his LSTM in ANN, CNN, and CNN phases, bidirectional learning with forward and backward passes of Bi-LSTM used unidirectional learning I have found it to be more effective than his LSTM. Accuracy is 59.81, 65.8, and 65.9 for KNN, CNN, and LSTM with CNN phase (Cahyani et al. 2022) study two experimental types of emotion detection based on text, that is Word2Vec-CNN-BiLSTM, and GloVe-CNN-BiLSTM. In the first and second scenarios, they classified emotion into five types of data namely happiness, anger, sadness, fear, and surprise. The accuracy for Word2Vec-CNN-BiLSTM, on the computer line, transjakarta, and computer line + transjakarta data are 84.34%, 83.73%, and 83.88% respectively. In their second scenario, the result shows that Word2Vec-CNN-BiLSTM has better accuracy than GloVe-CNN-BiLSTM. The accuracy results are 90.78%, 87.51%, and 90.31% respectively on the the computer line, transjakarta, and computer line + transjakarta data.

Park et al. (2020) developed two models to predict the meaning of Bangla texts. Their two predictive models allow prediction of six identical sentiment labels আনন্দ,বিষন্নতা, ভয়, রাগ, ভালবাসা, আযর). To achieve this, they rely on a bidirectional gated repetition unit he called BiGRU and a combination of the most classical deep learning algorithms, CNNs, and bidirectional long-term memory he called CNN-BiLSTM. They selected 7,214 sentences containing six of his emotions and six identical terms from the dataset for the experiment. CNN - BiLSTM found that they performed slightly better compared to BiGRU (64.96%) with an accuracy of 66.62%. They concluded that CNN-BiLSTM shows promising results in emotion detection of long Bangla phrases. Sasidhar et al. (2020) created a dataset of 12,000 Hindi and English codes. Texts from various sources were mixed and annotated with emotions of 'happy', 'sad', and 'angry'. They used a pre-trained bilingual model to generate feature vectors and employed a deep neural network as the classification model. They found that CNN-BiLSTM outperformed other experimental models with a classification accuracy of 83.21%. Methods used for emotion detection in the last 6 years include BERT, ALBERT, Robert, GMFCC, DMFCC, NB, SVM, CRF, DT, MLP, random forest, graph-based, GRU, GRNN, DLSTA, RNN, CGRU, mBERT etc. To find out the important points in emotion detection from the text of research from selective studies in 8 years, it can be seen in Fig. 13. Table 10 and explains the Summary of methods, level of analysis, and domain, and Table 11 contains a summary of related research on text-based emotion detection using deep learning models.
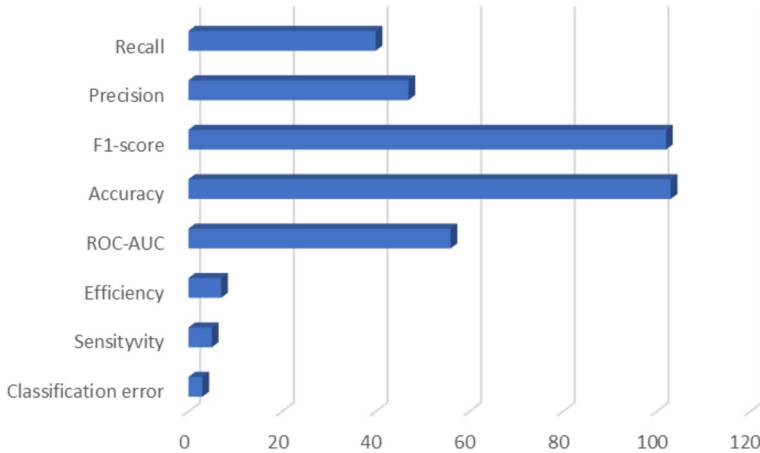
Fig. 11 Distribution of evaluation methods used in TBED

9 Evaluation

Evaluation is a process of identifying the process to measure the performance of a specific method. The most common metrics that are used to measure the method in emotion detection are Kappa coefficient, Jaccard Accuracy, Precision, Recall, F-score, Pearson correlation, 10-fold cross-validation, and Chi-square.

Ilyas et al. (2023) developed a large RU-EN (Roman Urdu-English) emotion detection corpus in which 20,000 sentences are annotated as Emotion—sentences. In their experiment for deep learning techniques, where they use the evaluation method that is Precision, Recall, and F1 score and the classifiers are CNN, RNN, LSTM, Bi-LSTM, At-Bi-LSTM (Attention-based-Bi-directional LSTM) and BERT. From the literature studies of the last 6 years, most evaluation approaches taken in emotion detection are accurate according to Fig. 11. Researchers often use the F1 measure, Recall, and Precision also to evaluate a method.

Rachman et al. (2016) use F1, recall, accuracy, and precision for their experiment. Where they take the four sets of parameters i.e. happy, angry, disgusted, and sad emotions. Filipe et al. (2020) takes categorical label emotion based on six basic emotions of
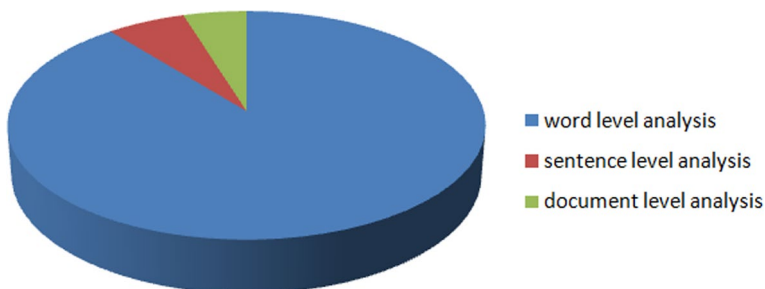
Fig. 12 Level of analysis in different papers and publications that we have collected

Ekman. Based on their experimental result, they used the evaluation methods F1, Recall, and Precision and found that the Corpus-Based Emotion (CBE) is able to improve the accuracy in the detection of emotions. The F-measure using WNA+ANEW was 0.50 and F-Measure using CBE with expanding is 0.61.

Majeed et al. (2020) discussed the different machine learning algorithms is like KNN, Decision Tree, SVM, and Random Forest for their corpus. They use three evaluation measures i.e. F1 score, Precision, and Recall where they perform 20 experiments to find out the best algorithm in machine learning and the performance of each classifier is measured. The results show that SVM gives higher accuracy, Precision, and Recall as compared to the Decision Tree, KNN, and Random Forest.

Strapparava and Mihalcea (2007) texted six different classification approaches that combine adjacent pairs of basic emotions: joy, trust, fear, surprise, sadness, disgust, anger, and anticipation to emotion identification in Portuguese tweets. They achieved a better performance that can be obtained both by improving a lexicon and by directly translating tweets into English and then applying an existing English lexicon. Their classifier succeeded in assigning at least one emotion to 118,046 tweets from Brazil, 6435 tweets from Portugal, 148 tweets from PALOP countries, and another 3459 tweets to other countries. The Portuguese and Brazilians show the same tendency for all emotions. Anger and disgust are the least the least pre-dominant emotions in all cases. Sadness is the most dominant emotion for Portugal and Brazil, but Sadness is surpassed by joy in PALOP countries and by anticipation in other countries (Fig. 12).

10 Challenges that were identified in existing studies in text-based emotion detection

Text-based emotion recognition brings some open challenges for NLP research. The challenge may be the complex nature of the emotional expression. In this section, existing challenges are discussed from the different aspects: Deng and Ren (2021) mentioned that there are four aspects of challenges and potential opportunities of text-based emotion recognition. Those are:

1. Shortage of high-quality datasets: Dimensional and discrete emotion models are broadly used in TBED as we have mentioned above. Deng and Ren (2021) claimed that incompatibilities of different databases, since there were no uniform annotation standards. Mohammad and Bravo-Marquez (2017) mentioned that SemEval-2007 was annotated with Ekman's six basic emotions such as anger, disgust, happiness, sadness, fear, and surprise. The author WASSA-2017 annotated with four emotions anger, fear, joy, sadness (Quan and Ren 2010a) and Ren-CECps annotated with eight emotions i.e. Anger, joy, surprise, anxiety, hate, love, sorrow, expect (Krawczyk 2016). Since all the datasets have different emotion labels, cross-corpus emotion resources may not be compatible. The author also mentioned that creating a link between the annotation scheme of different emotion corpus and building an emotion detection benchmark will incorporate existing emotion corpus and a more substantial corpus will cover all emotion categories. Buda et al. (2018) said that for a large amount of data, the emotion classifier is likely to identify emotion labels, which was always unavoidable in TBED tasks. Moreover, Deng and Ren 2021 gives suggestions for eliminating data imbalances by saying that

the way to eliminate data imbalances was to adjust data distribution by using two sampling techniques, over-sampling, and under-sampling, in which the oversampling often reproduced the data in minority classes to get overall relative balance distribution. Under-sampling removes the data majority classes to balance the data distribution in each emotion class. Chawla (2010) and Liew et al. (2016) mentioned that over-sampling can reduce the problem of data imbalance in TBED tasks. In case a large amount of training data were present, the model itself already be generalized and over-sampling may not work for a small amount of training data, over-sampling may lead to model overfitting resulting increase in the training time and computation cost. To achieve better or high-quality performance, high-quality datasets are necessary (Deng and Ren 2021; Lingren et al. 2014). There must be uniform emotional annotation, which gives compatibility between different corpora. The imbalanced distribution of data between each emotion label also affects the performance of the model. In short, the text-based emotion recognition models are data-dependent. Deng and Ren (2020) shows that it provides a large-scale emotional database to counteract and reduce the challenges of emotion annotation.

2. Fuzzy emotional boundaries: The nature of human emotional expression is not easy to understand. Even some symbols also have emotions, for example, "Whoa!", exclamation marks have their own emotions may be shock or surprise. Deng and Ren (2021) mentioned that due to the people's current feelings and own experiences. The emotions could be different even in the same expressions. Liu and Chen (2015) mentioned that fuzzy emotional boundaries and human subjective feelings are challenges for TBED. Buitinck et al. (2015) defined multi-label emotions as the different basic emotions, which can provide complex emotions more comprehensively and reduce the limitations. Nowadays, ML-EC (Multi-label emotion recognition) attracted more awareness and can recognize all possible emotions in textual expression. Tsoumakas and Vlahavas (2007), Jabreel and Moreno (2019) and Swain et al. (2018) claimed that to identify different emotion labels multi-label classification is used and MLC task can be transformed into more sophisticated learning scenarios through commonly used problem transformation algorithms and methods. Swain et al. (2018) transforms the ML-EC task into one binary classification problem by regenerating the input training data. Additionally, they mentioned that the Classifier chains (CC) transforms the MLC into a chain of binary classification problems to capture the label correlations.

3. Incomplete emotional information in text: Deng and Ren (2021) mentioned that achieving accurate emotion from humans only through textual information is a challenging task. For example, a person expressed in textual data is Joy, but this may also be the forced laughter, the inner emotion may be sad. Mohammad and Bravo-Marquez (2017) manually labeled SemEval-2007 for a surprise from 0.36 to 0.68 for sadness by using Pearson Correlation and other emotions were approximately 0.50. From which the author concluded their study that emotional information in text stream is incomplete and it is difficult for humans to recognize text stream emotion recognition tasks. The textual data are generally associated with other modalities like audio (Ko 2018), visual (García-Martínez et al. 2019; Schuller 2018), body language and psychological signals (Xu et al. 2018; Cabrera-Quiros et al. 2019) and multi-modal information predicts the overall performance of emotion recognition. The popularity of wearable devices is increasing nowadays and resulting people are more aware of recognizing emotions through physiological signals (Li et al. 2019a). Hence, Huang et al. (2019b) mentioned that the activity of emotions is directly proportional to the activity level of physiological signals (Li et al. 2019b; Chatterjee et al. 2019b).

4. Textual emotion recognition in dialogue: Since the open dialogue data is increasing with respect to the increasing users of social media, most of the text-based emotion recognition is conducted based on the assumption that the sentence is separated from expresses static emotion and the context, whereas the text-based emotion is highly correlated with contextual information and dynamic in nature for making Text-based emotion recognition task more challenging (Deng and Ren 2021). The performance was not adequate for directly applying the sentence-level deep text-based emotion recognition to the dialogue, and each utterance is highly dependent on the contextual signal. The contextual utterances and the current utterance are essential to the model for determining the dialogue's overall emotional tendency in text-based emotional dialogue. The context utterance and current utterance are merged and all tokens are compressed into a word sequence by flattened context modeling. Then, the word sequence is fed into neural networks for final prediction (Agrawal and Suri 2019; Yang et al. 2011). The dynamic emotion modeling aims to track and explore the overall tone and the contextual information in dialogue. It is non deterministic pattern. The emotion state transition matrix is the basis for emotional interactive learning in traditional methods (Teoh and Cho 2011). Generally, the Markov model is used to present the dynamic emotions in continuous dialogue (Adikari et al. 2019).

11 Conclusion

Emotion detection is an interesting research topic in the Sentiment analysis that helps produce concise information. The idea of this paper is to present the latest research and progress made in this field with deep learning methods. The paper introduces the concept of emotion detection, emotion models, evolution, some important datasets, some deep learning methods, and limitations that can be seen for TBED. To recognize emotions, remove any content that can't be used to identify emotions, such as hash-tagged content, URLs, emails, etc. is the first step in recognizing emotions. Then we converted the words that were misspelled, such as acronyms or unintentional message content. Then we find the group to which it belongs using the classifier, then look for polarity, valence, etc. to understand feelings. It's easy to identify emotions by using negative words, modalities, adjectives, and emoticons. We also discovered that the semantic underpinnings of every context, in addition to the words themselves, are what distinguish an emotion. With the deep learning method, it is proven that it can provide a more structured, broad, diverse review, datasets, pre-processing, features, approach techniques, problems, and methods of evaluation that are available as a guide for future work. The important thing that is considered interesting from the review that has been done is the results of the analysis is that deep learning techniques are still the topic of current favorite trends. This is because there are still many things that are challenging for researchers to do. The machine-learning approach is also a favorite technique because of its automatic machine-learning performance. We hope that through this review paper, the reader will have a better understanding of the research done on this topic, features extracted, methodologies used, and results reported by various researchers.
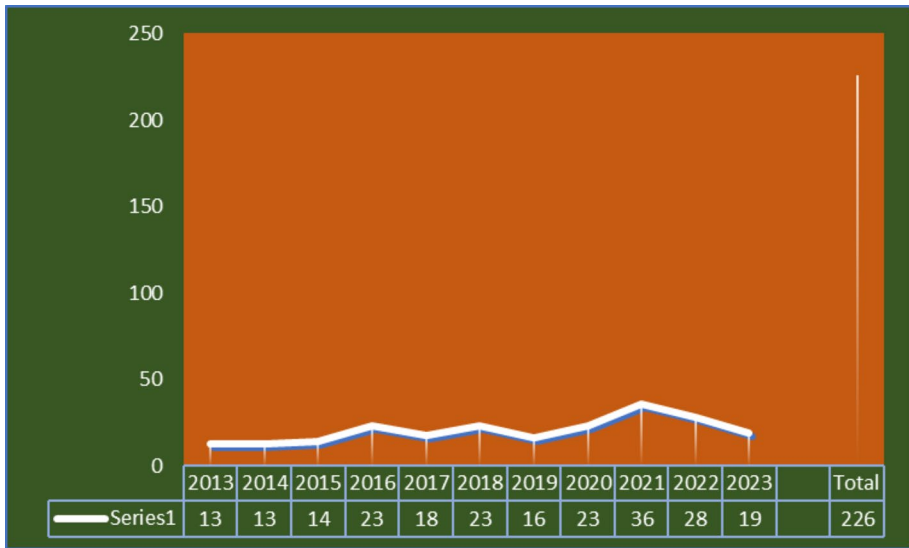
Fig. 13 Distribution over past 10 years for selected studies in text-based emotion detection

11.1 Practical implications and technical directions of emotion detection

Applications of emotion detection include a wide range of platforms for e-commerce sites, social media, public opinion, customer service, the medical field, and psychology. Emotion detection has undergone many related tasks such as emotion identification, depression analysis, anger detection, opinion mining, etc. (Fig. 13).

This analysis can be used to organize user reviews for a product or brand, find public opinion on a specific topic from social media, and look into customer satisfaction with products. Emotion detection can be used to extract the emotions of humans connected to user-generated content as long as it is involved. Emotion detection can help machines to become more powerful, and intelligent, better to compare the thoughts and opinions, which helps the businessman, psychologist, and policymakers etc. In emotion detection emotion dictionaries, statistical machine learning models, neural network-based deep learning models, and pre-trained models are used for better performance. Emotion detection is an important component of understanding natural language processing as well as sentiment analysis, in both business and academics to pay close attention to it. Currently, machines can provide much less accurate sentiment analysis than humans do. The emotion detection analysis will be enhanced by these, along with rules for affective reasoning to supplement interpretable information (Table 12).

11.2 Future directions of the related studies for overcoming the limitations and challenges

When considering the future direction, we think that the portability of the domain is probably the most important research direction (Michelucci 2022). The methods that are used in text-based emotion detection have observed a huge evolution over the past few years. The major difficulties in TBED have been addressed by AI-based technologies such as

Table 12 The summary of some challenges from the existing literature are represented in the table (Kusal et al. 2023)

References	Application domain	Dataset used	Methods used	Challenges
Basile et al. (2019)	Industrial Chat-bots	Service dialog dataset	Markov chains, word embedding	Performance of the system can be improved
Hien et al. (2018)	Chatbot interaction in Indian chat room	SemEval-2019	Bi-LSTM with deep attention mechanism, BERT, RF, SVM, LR	Distribution of the data are unbalanced, accuracy was not up to the mark Social media
Xue et al. (2018)	Education	Student data were utilized	Text classification, Named Entity Recognition	No knowledge structure between the response generated and user inputs
Wang et al. (2019)	Call center	Call center customer support agent	Bi-LSTM with attention mechanism	Less labeled data used
Thomas (2016)	Healthcare	Health care clinic data was used	SVM	Suggested activities are not clinically validated
Kumari et al. (2021)	E-business	FAQ from the e-business domain	Artificial Intelligence Markup Language (AIML) and Latent Semantic Analysis (LSA)	Only designed for FAQs, where general questions are not considered
Becker et al. (2022)	Behavioral analysis	Twitter and My space	Word embeddings, Stacked denoising autoencoder	Not considered word order in message which can further improve the robustness
Song et al. (2017)	Epidemics	Twitter	SVM, NB, RF, K-means clustering	Need to improve the performance of the classifier for sarcasm, Mislabeled data
Sun et al. (2015)	Crime detection	YouTube	Labeled latent Dirichlet allocation (L-LDA)	Need to explore other application domains besides YouTube, High-level features must be learned
Wang et al. (2017)	Recommendation System	Del.icio.us website	Bi-Clustering Algorithm, SVM	Performance must be improved, Place information, time-series information and dynamic variations among users are not considered for recommendation

Table 12 (continued)

References	Application domain	Dataset used	Methods used	Challenges
Kanchinadam et al. (2021)	Business Intelligence	Movie review dataset	Linear threshold and independent cascade diffusion models	Accuracy must be improved, need to observe the dynamic customers from social media, not consider the demographic characteristics of the most valuable customers
Bai et al. (2021)	Amazon	Customer care product review	GRU, CNN, GNN	Active learning approaches can improve the system
Chaturvedi et al. (2020)	Amazon	Customer product review	PoS tagging, Bootstrapping technique, Expectation maximization semi-supervised algorithm, feature ranking approach	The model does not directly link to the emotions, Limited domain knowledge and needs to explore more e-commerce sites
Mirtalaie et al. (2018)	Amazon	Product review	Domain Adaption using genetic programming, Semi-supervised convolutional Deep Belief Network (CDBN)	Variance is high during the search
Topal and Ozsoyoglu (2016)	Digital Camera review	Pros and cons of digital camera review	Syntactic rules with polarity determination	Dependent on language, can't analyze the movie review
Xiang et al. (2017)	Movie review	IMDb movie review	K-means clustering	Very few amount of movie were considered for study
Truşcă and Frasincar (2023)	Tourism and hospitality	Yelp, Expedia, TripAdvisor	Latent Dirichlet Allocation, Elbow method, NB, TF-IDF	Destination specific data used, Less specific case discussed

autoencoder, different learning styles, data augmentation (GANs), transfer learning, explainable AI, and adversarial machine learning (AML).

1. *Auto-encoders*: Auto-encoders are a type of neural network that are typically used in supervised settings where the feature vectors are transferred into an arbitrary lower or high dimensional space during the encoding stage. The purpose of the learning is an "informative" representation of the data that may be used in different applications by reconstructing a set of input observations well enough (Sagha et al. 2017). The autoencoder can deal with missing data situations, improve system performance in noisy environments, and can be trained unsupervised. Another form of autoencoder is Denoising autoencoders, which have been employed in domain adaption, training a model on one labeled corpus with a different context (Kusal et al. 2023). Becker et al. (2022) used autoencoder to build an LSTM model and trained with non-aggressive comments. The data they have been collected from Facebook and Twitter with Bilingual (English and Hindi) sentences and developed a deep-learning model to identify different levels of aggression (direct, indirect, and no aggression). Kusal et al. (2023), Deldjoo et al. (2021) used SDA's (Stacked Denoising Autoencoders) to perform sentiment analysis on a multitude to domains and languages.

2. *Adversarial Machine Learning (AML)* NLP is an emerging research area, where deep learning and machine learning models have achieved success day by day. Another emerging research area is AML, which brings together the best machine learning practices, computer security, and robust statistics. Alsmadi (2021) The goal of their survey was two-fold the first one was to present recent advances on AML for the security of RS (i.e. attacking and defense recommendation models) and the second one was to show another successful application of AML in GAN's (Generative Adversarial Networks) for generative applications. Alshemali and Kalita (2020) discussed the different types and attempts of adversarial attacks where they focused on studying aspects and research trends in AML specifically in text generation and text analysis. Zhang et al. (2020) and Bayer et al. (2023) discussed the different adversarial attacks on deep learning techniques in NLP.

3. *Data Augmentation*: In NLP, it required vast training data to improve the performance of the model to complete a task like text classification. It is a time-consuming and expensive process in order to generate a large amount of training data. To solution of this scenario may be data augmentation. Data augmentation is a machine learning technique that artificially generates the additional synthetic training data through label preserving the transformation. Shi et al. (2020) mentioned that the data augmentation can be achieved by the thesaurus, back translation, word-embedding, replacing the synonyms, text-generation, etc. The author proposed a novel text-generation method to improve classifiers by artificially created training data and also used the GPT-2 355 million parameter model. They also presented and evaluated the text generation method for long and short text, where they achieved promising results when evaluating short as well as long text. Stanton and Irissappane (2019) developed an effective data augmentation algorithm called AUG-BERT for text-classification in labeled sentences.

GAN is one of the trending data augmentation techniques for dealing with imbalanced classes. Kusal et al. (2023) defined GAN as a new framework that sets two neural networks against one other to create new, synthetic data that may look like actual data. Croce et al. (2020) proposed spamGAN, a generative adversarial network that works in a limited set of labeled data as well as unlabeled data for opinion spam detection. The author claimed that the spanGAN improves the state-of-art GAN-based techniques for

text classification, where they experiment with the TripAdvisor dataset. The spamGAN is not only for detecting span opinion but also generating reviews that have reasonable perplexity. Li et al. (2020a) proposed a GAN-BERT algorithm that works on fine-tuning of BERT-based architectures with unlabeled data in GAN setting. Their experimental results show that it still achieved good performance in sentence classification tasks after the requirement for annotated examples that were reduced to 50–100 examples. Fang and Zhan (2015) proposed a deep multi-modal generative and fusion framework for multi-modal classification in imbalanced data. Their work consists of two modules, DMGAN and DMHFN, i.e. Deep Multi-modal Generative Adversarial Network and Deep Multi-modal Hybrid Fusion Network. Here, the DMGAN is used to class imbalanced problems and DMHFN identifies the integration of different information sources and fine-grained interactions for multi-modal classification.

4. *Transfer learning*: Kusal et al. (2023) discussed transfer learning in text-based emotion recognition. The author mentioned that collecting huge amounts of data and labeling each emotion is a time-consuming process, the transfer learning allows us to reduce the amount of labeled data we need and take advantage of previously trained models to achieve better performance. The author also claimed that the source domain knowledge can be transferred to separate but related target domains using transfer learning, whereas it tries to increase the learning of the target task using useful information from the source domain when given a source and a target (Ahmad et al. 2020; Kusal et al. 2023). Omara et al. (2019) proposed a system by using Transfer learning for emotion recognition via cross-lingual embedding. Yasaswini et al. (2021) proposed a deep CNN modal for Arabic sentiment analysis based on character level representation. Their proposed model was used for applying Transfer learning between emotion recognition and sentiment analysis domains in the Arabic language. Hazarika et al. (2021) discussed the Transfer learning for offensive language detection in Dravidian Languages. They also discussed the challenges with various transfer learning-based models to classify the given sentences that are Malayalam, Tamil, and Kannada languages with six emotion labels. Zucco et al. (2018) discussed the conversational transfer learning for emotion recognition. They also proposed an approach, named TL-ERC (Transfer Learning-Emotion Recognition in Conversations), where they pre-train a hierarchical dialogue model on multi-turn conversations and then transfer its parameters to a targeted conversational emotion classifier.

5. *Explainable Artificial Intelligence (XAI)* Kusal et al. (2023) claimed that machine learning and deep learning models have drawbacks for not being humanly interpretable, where lack of transparency is the major problem of both models. Very less amount of work has been done in this research area. Turcan et al. (2021) discussed the sentiment analysis approaches and already existing explainable systems, additionally they reviewed a critical area of explainable sentiment analysis approaches and discussed the insights of using explainable sentiment analysis in the medical field. Farruque et al. (2021) and Yin (2020) also used the Explainable Zero-shot model to Detect Depression Symptoms from text streams.

6. *Different learning styles* There are different learning styles depending on how algorithms use multiple layers to extract increasingly higher-level features from the raw data. The following learning styles are going to be advanced trends in NLP research areas.

- *Meta-learning* To solve a novel problem, it collects a large number of tasks, and each task is treated as a training example, after training a model to adapt all those training tasks meta-learning is used. Kusal et al. (2023) said that Meta-learning is a

crucial task for both task-specific and task-agnostic scenarios. The major objectives they mentioned in their research were to expedite the learning process and improve quantitative model performance. Wu and Guo (2020) discussed the problem of impulsive recognition of the text classification task using meta-learning, where they mentioned a total of 73 meta-features that were tested in 81 data sets that are related to different five types of tasks.

- *Co-learning* According to Kusal et al. (2023) co-learning is based on co-training and was first introduced by Mitchell and Blum in 1998. Where the Co-learning proposed that multiple classifiers can learn from different features of data. When the unavailability of labeled data or lack of training data situation occurs, co-learning is used to improve the generalization capacity of classifiers with learned features, which gives a better classification model in text classification. The first step is for all classifiers to be trained on a small labeled dataset in a completely supervised method and then each unlabeled data is tested as a trained classifier (Kusal et al. 2023). This automatically labeled data, by each classifier, is then used to retrain the other classifiers in a semi-supervised model. Helwe et al. (2019) proposed a unique technique for multi-domain text categorization named dual adversarial co-learning. De and Guo (2015) also discussed Assessing Arabic weblog credibility via deep co-learning, where they showed a lack of training data in the respective web blogs co-learning approach performed well with other baseline models.

- *Multi-task learning* As the name suggests the Multi-task learning is simultaneously trained on many tasks rather than separately on each task. The goal is to create a generalized model that can handle a variety of tasks at the same time (Kusal et al. 2023). The author mentioned that shared representation is used to learn several linked tasks. There are two mechanisms in Multi-task learning, these are Kusal et al. (2023):

  – Soft-sharing mechanism that applies a task-specific layer to a different task
  – Hard-sharing mechanism that utilizes a shared feature space to extract the features for different tasks.

  Zhang et al. (2021) created a framework for classifying and finding aspect elements in a unified model by using Bi-LSTM and CNN network in a cascade for collaborative learning from which they can extract terms that are related to aspects and classify aspect sentiment in a multitasking framework. Hu et al. (2021) proposed a Multi-task learning framework using several application domains based on a hard-sharing mechanism for sentiment analysis and

- *Zero-shot learning* Zero-shot learning is also called Few-shot learning which is the most used machine learning classification model. Kusal et al. (2023) mentioned that, at the time of testing if there is no labeled example, it is known as zero-shot learning, when there is only one labeled example, it is named one-shot learning, and when there are a few labeled instances at the time of testing is called few-shot learning. Thakkar et al. (2022) proposed a multi-label few-shot learning system for Aspect Category Detection (ACD) in sentiment analysis. Kusal et al. 2023 and Chen et al. (2019) used zero-shot and few-shot learning in news articles to perform cross-lingual sentiment analysis, where they used single-task and multi-task environments on Slovene and Croatian datasets.

- *Reinforcement learning* Kusal et al. (2023) discussed reinforcement learning and defined it as a machine learning technique where an agent is trained in an interactive environment through reward and punishment mechanisms using feedback from

its own actions and experiences. The agent has been punished for the wrong actions and rewarded for the correct actions. The author also mentioned that most of the reinforcement learning used in NLP research areas are Question–answer, Text-summarization, machine translation, dialogue systems etc. Chai et al. (2020) proposed a framework named WS-LSTM (Word-level Sentiment Analysis), where they try to achieve sentiment tendency for each word in a sentence. Su et al. (2021) proposed a model to perform the most salient texts with respect to the emotion label, which can be regarded as a hard version of attention for better performance. The author also observed the significant performance boosts over a wide range of text-classification tasks which includes single-labeled, multi-labeled, and multi-aspect classification on sentiment analysis.

- *Self-supervised learning* As the name suggests it is a self-supervision process, where the models learn themselves by utilizing one part to anticipate the other of the data and create accurate annotations, called pretext tasks. The self-supervised learning is used where the dataset is too large and labeling it will be time-consuming, in which the data labeling works as an automated process, where the human intervention is eliminated. Su et al. (2021) proposed an approach that automatically trains attention supervision information for neural aspect-based sentiment analysis. Su et al. (2021) proposed a framework named SSentiA (Self-supervised Sentiment Analyzer), which is a hybrid method of self-supervised learning for sentiment classification from unlabeled data that combines a lexicon-based strategy with a machine learning classifier.

11.3 Limitations of our work

There are some limitations of our work:

- We only collected the papers that are written in English, Persian, Indonesian, Arabic, Urdu, and Bangla languages and searched from different journals and publications. We believe that there are papers in other languages or other databases that also include emotion detection from text.
- The keywords we chose to search on the internet were mainly "emotion detection", "emotion detection from the text", "anger detection", and "emotion extraction". There may be many papers related to our research topic that may not have the above keywords. To investigate the developments in emotion detection research, future studies could replicate this work by employing more precise keywords and using different literature databases (Itani et al. 2017).
- We selected the main high-frequency keywords for emotion detection analysis, and some low-frequency keywords may have been ignored.
- Another limitation of this article may be theoretic concepts, and in the literature, moreover there is a scope to establish a bridge between the "sentiment analysis" and "emotion recognition". So that we develop some hybrid definitions from the best of the two domains.
- The normalization procedure that may deployed in this article, particularly in the deep learning algorithm is a more conventional one. In future work, we can analyze the changes in each keyword in detail from the perspective of time and more comprehensive analysis results.

Author contributions In the initial step we have collected the research papers from the different languages with different deep learning methods. Then we summarize the literature review process. We also focused on the technical aspects of the article. They delved into the details of various machine learning and natural language processing techniques employed in emotion detection, highlighting their strengths and limitations and the ethical and social implications section of the article. In this article, we also discussed the potential impacts of emotion detection in text on privacy, bias, and societal implications, providing critical insights into the broader context of the technology. We mentioned the different emotional models and its classifications. We had discussed about the related work that are done in text based emotion analysis, about the existing surveys on contents and topics, surveys on methods of emotion detection, and overall methodologies of text based emotion detection, the different document type whether it is from journal, conference paper, workshop paper or dissertation. We also discussed about the datasets, data preprocessing and feature extraction. The word level, sentence level and document level analysis is also discussed in this paper . The different deep learning techniques are reviewed in its different methodologies. The different evaluation techniques/metrices with results and conclusions are discussed in this paper also.The author Tulika chutia is a research scholar under the supervision of Nomi Baruah and supervised the entire project, provided guidance throughout the writing process, and ensured that the review article met the highest standards in terms of accuracy and academic rigor.

Declarations

References

Abdelgwad MM, Soliman THA, Taloba AI, Farghaly MF (2022) Arabic aspect based sentiment analysis using bidirectional gru based models. J King Saud Univ Comput Inf Sci 34(9):6652–6662

Abdullah M, Shaikh S (2018) Teamuncc at semeval-2018 task 1: Emotion detection in english and arabic tweets using deep learning. In: Proceedings of the 12th International Workshop on Semantic Evaluation, pp. 350–357

Abdullah M, Hadzikadicy M, Shaikhz S (2018) SEDAT: sentiment and emotion detection in arabic text using CNN-LSTM deep learning. In: 2018 17th IEEE international conference on machine learning and applications (ICMLA). IEEE, pp 835–840

Acheampong FA, Wenyu C, Nunoo-Mensah H (2020) Text-based emotion detection: advances, challenges, and opportunities. Eng Rep 2(7):12189

Acheampong FA, Nunoo-Mensah H, Chen W (2021a) Recognizing emotions from texts using an ensemble of transformer-based language models. In: 2021 18th International computer conference on wavelet active media technology and information processing (ICCWAMTIP). IEEE, pp 161–164

Acheampong FA, Nunoo-Mensah H, Chen W (2021b) Transformer models for text-based emotion detection: a review of BERT-based approaches. Artif Intell Rev 54:5789–5829

Adikari A, De Silva D, Alahakoon D, Yu X (2019) A cognitive model for emotion awareness in industrial Chatbots. In: 2019 IEEE 17th international conference on industrial informatics (INDIN), vol 1. IEEE, pp 183–186

Adoma AF, Henry N-M, Chen W (2020a) Comparative analyses of Bert, Roberta, Distilbert, and Xlnet for text-based emotion recognition. In: 2020 17th International computer conference on wavelet active media technology and information processing (ICCWAMTIP). IEEE, pp 117–121

Adoma AF, Henry N-M, Chen W, Andre NR (2020b) Recognizing emotions from texts using a BERT-based approach. In: 2020 17th International computer conference on wavelet active media technology and information processing (ICCWAMTIP). IEEE, pp 62–66

Agrawal A, An A (2012) Unsupervised emotion detection from text using semantic and syntactic relations. In: 2012 IEEE/WIC/ACM international conferences on web intelligence and intelligent agent technology, vol 1. IEEE, pp 346–353

Agrawal P, Suri A (2019) NELEC at SemEval-2019 task 3: think twice before going deep. arXiv preprint. arXiv:1904.03223

Ahmad Z, Jindal R, Ekbal A, Bhattachharyya P (2020) Borrow from rich cousin: transfer learning for emotion detection using cross lingual embedding. Expert Syst Appl 139:112851

Ahuja R, Chug A, Kohli S, Gupta S, Ahuja P (2019) The impact of features extraction on the sentiment analysis. Procedia Comput Sci 152:341–348

Akilandeswari J, Jothi G (2018) Sentiment classification of tweets with non-language features. Procedia Comput Sci 143:426–433

Allouch M, Azaria A, Azoulay R, Ben-Izchak E, Zwilling M, Zachor DA (2018) Automatic detection of insulting sentences in conversation. IEEE, pp 1–4

Alm CO, Roth D, Sproat R (2005) Emotions from text: machine learning for text-based emotion prediction. In: Proceedings of human language technology conference and conference on empirical methods in natural language processing, pp 579–586

Al-Omari H, Abdullah MA, Shaikh S (2020) EmoDet2: emotion detection in English textual dialogue using BERT and BiLSTM models. In: 2020 11th International conference on information and communication systems (ICICS). IEEE, pp 226–232

Alotaibi FM (2019) Classifying text-based emotions using logistic regression. VAWKUM Trans Comput Sci 7(1):31–37

Alqaryouti O, Siyam N, Abdel Monem A, Shaalan K (2024) Aspect-based sentiment analysis using smart government review data. Appl Comput Inf 20(1/2):142–161

Al-Rfou R, Choe D, Constant N, Guo M, Jones L (2019) Character-level language modeling with deeper self-attention. In: Proceedings of the AAAI conference on artificial intelligence, vol 33, pp 3159–3166

Alshemali B, Kalita J (2020) Improving the reliability of deep neural networks in NLP: a review. Knowl Based Syst 191:105210

Alsmadi I (2021) Adversarial machine learning in text analysis and generation. arXiv preprint. arXiv:2101.08675

Alswaidan N, Menai MEB (2020) A survey of state-of-the-art approaches for emotion recognition in text. Knowl Inf Syst 62:2937–2987

Alvarez-Gonzalez N, Kaltenbrunner A, Gómez V (2021) Uncovering the limits of text-based emotion detection. arXiv preprint. arXiv:2109.01900

Alwehaibi A, Bikdash M, Albogmi M, Roy K (2022) A study of the performance of embedding methods for Arabic short-text sentiment analysis using deep learning approaches. J King Saud Univ Comput Inf Sci 34(8):6140–6149

Aman S, Szpakowicz S (2007) Identifying expressions of emotion in text. In: International conference on text, speech and dialogue. Springer, Berlin, pp 196–205

Anantrasirichai N, Bull D (2022) Artificial intelligence in the creative industries: a review. Artif Intell Rev 55:589–656

Arora M, Kansal V (2019) Character level embedding with deep convolutional neural network for text normalization of unstructured data for twitter sentiment analysis. Soc Netw Anal Min 9:1–14

Asghar MZ, Khan A, Bibi A, Kundi FM, Ahmad H (2017) Sentence-level emotion detection framework using rule-based classification. Cogn Comput 9:868–894

Ashish V, Somashekar R, Sundeep Kumar K (2016) Keyword based emotion word ontology approach for detecting emotion class from text. Int J Sci Res (IJSR) 5(5):1636–1639

Asif M, Mishra S, Vinodbhai MT, Tiwary US (2023) Emotion recognition using temporally localized emotional events in EEG with naturalistic context: DENS# dataset. IEEE Access. https://doi.org/10.1109/ACCESS.2023.3266804

Aslam N, Rustam F, Lee E, Washington PB, Ashraf I (2022) Sentiment analysis and emotion detection on cryptocurrency related tweets using ensemble LSTM-GRU model. IEEE Access 10:39313–39324

Aslan M (2022) CNN based efficient approach for emotion recognition. J King Saud Univ Comput Inf Sci 34(9):7335–7346

Awais M, Raza M, Singh N, Bashir K, Manzoor U, Islam SU, Rodrigues JJ (2020) LSTM-based emotion detection using physiological signals: IoT framework for healthcare and distance learning in Covid-19, vol 8. IEEE, pp 16863–16871

Ayub A, Wagner AR (2021) EEC: learning to encode and regenerate images for continual learning. arXiv preprint. arXiv:2101.04904

Bai X, Huerta O, Unver E, Allen J, Clayton JE (2021) A parametric product design framework for the development of mass customized head/face (eyewear) products. Appl Sci 11(12):5382

Bansal D, Grover R, Saini N, Saha S (2021) GenSumm: a joint framework for multi-task Tweet classification and summarization using sentiment analysis and generative modelling. IEEE Trans Affect Comput. https://doi.org/10.1109/TAFFC.2021.3131516

Bashir MF, Javed AR, Arshad MU, Gadekallu TR, Shahzad W, Beg MO (2023) Context-aware emotion detection from low-resource Urdu language using deep neural network, vol 22. ACM, New York, pp 1–30

Basile A, Franco-Salvador M, Pawar N, Štajner S, Chinea-Ríos M, Benajiba Y (2019) Symantoresearch at SemEval-2019 task 3: combined neural models for emotion classification in human-chatbot conversations. In: Proceedings of the 13th international workshop on semantic evaluation, pp 330–334

Batbaatar E, Li M, Ryu KH (2019) Semantic-emotion neural network for emotion recognition from text. IEEE Access 7:111866–111878

Bayer M, Kaufhold M-A, Buchhold B, Keller M, Dallmeyer J, Reuter C (2023) Data augmentation in natural language processing: a novel text generation approach for long and short text classifiers. Int J Mach Learn Cybern 14(1):135–150

Becker LA, Penagos H, Flores FJ, Manoach DS, Wilson MA, Varela C (2022) Eszopiclone and zolpidem produce opposite effects on hippocampal ripple density. Front Pharmacol 12:792148

Bell C (1824) Essays on the anatomy and philosophy of expression. J. Murray, London

Bengio Y, Simard P, Frasconi P (1994) Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Netw 5(2):157–166

Bhatti MW, Wang Y, Guan L (2004) A neural network approach for human emotion recognition in speech. In: Proceedings of the 2004 international symposium on circuits and systems (ISCAS '04), vol 2. IEEE, p 181

Britz D (2015) Recurrent neural networks tutorial, part 1—introduction to RNNs. http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns

Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, Neelakantan A, Shyam P, Sastry G, Askell A et al (2020) Language models are few-shot learners. Adv Neural Inf Process Syst 33:1877–1901

Buda M, Maki A, Mazurowski MA (2018) A systematic study of the class imbalance problem in convolutional neural networks. Neural Netw 106:249–259

Buechel S, Rücker S, Hahn U (2020) Learning and evaluating emotion lexicons for 91 languages. arXiv preprint. arXiv:2005.05672

Buitinck L, Van Amerongen J, Tan E, Rijke M (2015) Multi-emotion detection in user-generated reviews. In: Advances in information retrieval: 37th European conference on IR research (ECIR 2015), Vienna, Austria, March 29–April 2, 2015, proceedings, vol 37. Springer, Cham, pp 43–48

Cabrera-Quiros L, Gedik E, Hung H (2019) Multimodal self-assessed personality estimation during crowded mingle scenarios using wearables devices and cameras. IEEE Trans Affect Comput 13(1):46–59

Cahyani DE, Wibawa AP, Prasetya DD, Gumilar L, Akhbar F, Triyulinar ER (2022) Text-based emotion detection using CNN-BILSTM. In: 2022 4th International conference on cybernetics and intelligent system (ICORIS) IEEE, pp 1–6

Cai X, Hao Z (2018) Multi-view and attention-based bi-lstm for weibo emotion recognition. In: 2018 International conference on network, communication, computer engineering (NCCE 2018). Atlantis Press, Amsterdam, pp 772–779

Calvo RA, Mac Kim S (2013) Emotions in text: dimensional and categorical models. Comput Intell 29(3):527–543

Canales L, Martínez-Barco P (2014) Emotion detection from text: a survey. In: Proceedings of the workshop on natural language processing in the 5th information systems research working days (JISIC), pp 37–43

Chaffar S, Inkpen D (2011) Using a heterogeneous dataset for emotion analysis in text. In: Advances in artificial intelligence: 24th Canadian conference on artificial intelligence, Canadian AI 2011, St. John's, Canada, 25–27 May 2011. Proceedings, vol 24. Springer, Berlin, pp 62–67

Chai D, Wu W, Han Q, Wu F, Li J (2020) Description based text classification with reinforcement learning. In: International conference on machine learning, pp 1371–1382. PMLR

Charoensuk J (2018) Human emotion recognition in thai short text. Dissertation, National Institute of Development Administration

Chatterjee A, Gupta U, Chinnakotla MK, Srikanth R, Galley M, Agrawal P (2019a) Understanding emotions in text using deep learning and big data. Comput Hum Behav 93:309–317

Chatterjee A, Narahari KN, Joshi M, Agrawal P (2019b) Semeval-2019 task 3: emocontext contextual emotion detection in text. In: Proceedings of the 13th international workshop on semantic evaluation, pp 39–48

Chaturvedi I, Cambria E, Cavallari S, Welsch RE (2020) Genetic programming for domain adaptation in product reviews. In: 2020 IEEE congress on evolutionary computation (CEC). IEEE, pp 1–8

Chawla NV (2010) Data mining for imbalanced datasets: an overview. In: Data mining and knowledge discovery handbook. Springer, Boston, pp 875–886

Chen R, Zhou Y, Zhang L, Duan X (2019) Word-level sentiment analysis with reinforcement learning. IOP Conf Ser Mater Sci Eng 490:062063

Cheng R, Zhang J, Hu P (2020) Document-level emotion detection using graph-based margin regularization. Neurocomputing 398:55–63

Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv preprint. arXiv:1406.1078

Chowanda A, Sutoyo R, Tanachutiwat S et al (2021) Exploring text-based emotions recognition machine learning techniques on social media conversation. Procedia Comput Sci 179:821–828

Chudasama V, Kar P, Gudmalwar A, Shah N, Wasnik P, Onoe N (2022) M2FNet: multi-modal fusion network for emotion recognition in conversation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 4652–4661

Clarizia F, Colace F, De Santo M, Lombardi M, Pascale F, Pietrosanto A (2018) E-learning and sentiment analysis: a case study. In: Proceedings of the 6th international conference on information and education technology, pp 111–118

Conneau A, Lample G (2019) Cross-lingual language model pretraining. In: Proceedings of the conference on neural information processing systems, vol 32 Vancouver

Cortiz D (2021) Exploring transformers in emotion recognition: a comparison of BERT, DistilBERT, RoBERTa, XLNet, and ELECTRA. arXiv preprint. arXiv:2104.02041

Croce D, Castellucci G, Basili R (2020) GAN-BERT: generative adversarial learning for robust text classification with a bunch of labeled examples. In: Proceedings of the 58th annual meeting of the Association for Computational Linguistics

Cui J, Wang Z, Ho S-B, Cambria E (2023) Survey on sentiment analysis: evolution of research methods and topics. Artif Intell Rev 56:8469–8510

Dai Z, Yang Z, Yang Y, Carbonell J, Le QV, Salakhutdinov R (2019) Transformer-XL: attentive language models beyond a fixed-length context. arXiv preprint. arXiv:1901.02860

Dangi D, Dixit DK, Bhagat A, Nair R, Verma N (2021) Analyzing the sentiments by classifying the tweets based on Covid-19 using machine learning classifiers. IEEE, pp 1–6

Dangi D, Bhagat A, Dixit DK (2022a) Sentiment analysis of social media data based on chaotic coyote optimization algorithm based time weight-adaboost support vector machine approach. Concurr Comput Pract Exp 34(3):6581

Dangi D, Dixit DK, Bhagat A (2022b) Sentiment analysis of COVID-19 social media data through machine learning. Multimed Tools Appl 81:42261–42283

Dangi D, Chandel ST, Dixit DK, Sharma S, Bhagat A (2023) An efficient model for sentiment analysis using artificial rabbits optimized vector functional link network. Expert Syst Appl 225(1):119849

Danisman T, Alpkocak A (2008) Feeler: emotion classification of text using vector space model. In: AISB 2008 convention communication, interaction and social intelligence, vol 1, p 53

Darwin C, Prodger P (1998) The expression of the emotions in man and animals. Oxford University Press, Oxford

Dashtipour K, Gogate M, Li J, Jiang F, Kong B, Hussain A (2020) A hybrid Persian sentiment analysis framework: integrating dependency grammar based rules and deep neural networks, vol 380. Elsevier, Amsterdam, pp 1–10

Dave C, Khare M (2021) Emotion detection in conversation using class weights. In: 2021 8th international conference on soft computing and machine intelligence, ISCMI 2021. IEEE, pp 231–236

De A, Guo C (2015) An adaptive vector quantization approach for image segmentation based on som network. Neurocomputing 149:48–58

Del Vigna12 F, Cimino23 A, Dell'Orletta F, Petrocchi M, Tesconi M (2017) Hate me, hate me not: hate speech detection on Facebook. In: Proceedings of the first italian conference on cybersecurity (ITASEC17), pp 86–95

Deldjoo Y, Noia TD, Merra FA (2021) A survey on adversarial recommender systems: from attack/defense strategies to generative adversarial networks. ACM Comput Surv (CSUR) 54(2):1–38

Demszky D, Movshovitz-Attias D, Ko J, Cowen A, Nemade G, Ravi S (2020) Goemotions: a dataset of fine-grained emotions. arXiv preprint. arXiv:2005.00547

Deng J, Ren F (2020) Multi-label emotion detection via emotion-specified feature extraction and emotion correlation learning. IEEE Trans Affect Comput 14:475–486

Deng J, Ren F (2021) A survey of textual emotion recognition and its challenges. IEEE Trans Affect Comput. https://doi.org/10.1109/TAFFC.2021.3053275

Devlin J, Chang M-W, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint. arXiv:1810.04805

Dheeraj K, Ramakrishnudu T (2021) Negative emotions detection on online mental-health related patients texts using the deep learning with mha-bcnn model. Expert Syst Appl 182:115265

Dogra V, Singh A, Verma S, Kavita Jhanjhi N, Talib M (2021) Analyzing DistilBERT for sentiment classification of banking financial news. In: Intelligent computing and innovation on data science: proceedings of ICTIDS 2021. Springer, Singapore, pp 501–510

Duric A, Song F (2012) Feature selection for sentiment analysis based on content and syntax models. Decis Support Syst 53(4):704–711

Ebling S, Camgöz NC, Braem PB, Tissi K, Sidler-Miserez S, Stoll S, Hadfield S, Haug T, Bowden R, Tornay S et al (2018) Smile swiss German sign language dataset. In: Proceedings of the 11th international conference on language resources and evaluation (LREC 2018). The European Language Resources Association (ELRA), Paris

Eckel EJ (2009) The emerging engineering scholar: a citation analysis of theses and dissertations at Western Michigan University. Issues Sci Technol Librariansh. https://doi.org/10.5062/F4HD7SKP

Ekman P (1992) An argument for basic emotions. Cogn Emotion 6(3–4):169–200

Ekman P et al (1999) Basic emotions. In: Handbook of cognition and emotion, vol 98. Springer, New York, pp. 45–60

El Hammoumi O, Benmarrakchi F, Ouherrou N, El Kafi J, El Hore A (2018) Emotion recognition in e-learning systems. In: 2018 6th international conference on multimedia computing and systems (ICMCS). IEEE, pp 1–6

Fang X, Zhan J (2015) Sentiment analysis using product review data. J Big Data 2(1):1–14

Farruque N, Goebel R, Zaïane OR, Sivapalan S (2021) Explainable zero-shot modelling of clinical depression symptoms from text. In: 2021 20th IEEE international conference on machine learning and applications (iCMLA). IEEE, pp 1472–1477

Filipe S, Batista F, Ribeiro R (2020) Different lexicon-based approaches to emotion identification in Portuguese tweets. In: 9th symposium on languages, applications and technologies (SLATE 2020), pp 12:1–12:8

Gao K, Xu H, Wang J (2015) A rule-based approach to emotion cause detection for chinese micro-blogs. Expert Syst Appl 42(9):4517–4528

García-Martínez B, Martinez-Rodrigo A, Alcaraz R, Fernández-Caballero A (2019) A review on nonlinear methods using electroencephalographic recordings for emotion recognition. IEEE Trans Affect Comput 12(3):801–820

George A, HB BG, Soman K, et al (2018) Teamcen at Semeval-2018 task 1: global vectors representation in emotion detection. In: Proceedings of the 12th international workshop on semantic evaluation, pp 334–338

Gers FA, Schraudolph NN, Schmidhuber J (2002) Learning precise timing with LSTM recurrent networks. J Mach Learn Res 3(Aug), 115–143

Ghanbari-Adivi F, Mosleh M (2019) Text emotion detection in social networks using a novel ensemble classifier based on parzen tree estimator (TPE). Neural Comput Appl 31(12):8971–8983

Ghosh S, Ramaneswaran S, Tyagi U, Srivastava H, Lepcha S, Sakshi S, Manocha D (2022) M-MELD: a multilingual multi-party dataset for emotion recognition in conversations. arXiv preprint. arXiv:2203.16799

Go A, Bhayani R, Huang L (2009) Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford 1, p 12

Goyal S, Tiwari N (2017) Emotion recognition: a literature survey. Int J Technol Res Eng 4:1502–1524

Graves A, Schmidhuber J (2005) Framewise phoneme classification with bidirectional lstm and other neural network architectures. Neural Netw 18(5–6):602–610

Grover S, Verma A (2016) Design for emotion detection of punjabi text using hybrid approach. In: 2016 International conference on inventive computation technologies (ICICT), vol 2. IEEE, pp 1–6

Guo J (2022) Deep learning approach to text analysis for human emotion detection from big data. J Intell Syst 31(1):113–126

Haryadi D, Kusuma GP (2019) Emotion detection in text using nested long short-term memory. Int J Adv Comput Sci Appl 10(6):351–357. https://doi.org/10.14569/ijacsa.2019.0100645

Hasan M, Rundensteiner E, Agu E (2014) Emotex: detecting emotions in Twitter messages. In: 2014 ASE BIGDATA/SOCIALCOM/CYBERSECURITY conference, Stanford University, 27–31 May 2014

Hasan M, Rundensteiner E, Agu E (2019) Automatic emotion detection in text streams by analyzing twitter data, vol. 7, pp. 35–51. Springer, ???

Hassan S, Shaar S, Darwish K (2021) Cross-lingual emotion detection. arXiv preprint. arXiv:2106.06017

Hazarika D, Poria S, Zimmermann R, Mihalcea R (2021) Conversational transfer learning for emotion recognition. Inf Fusion 65:1–12

He H, Xu L, Choi JD (2021) ELIT: emory language and information toolkit. arXiv preprint. arXiv:2109.03903

Helwe C, Elbassuoni S, Al Zaatari A, El-Hajj W (2019) Assessing arabic weblog credibility via deep co-learning. In: Proceedings of the 4th Arabic natural language processing workshop, pp 130–136

Hien HT, Cuong P-N, Nam LNH, Nhung HLTK, Thang LD (2018) Intelligent assistants in higher-education environments: the FIT-EBot, a chatbot for administrative and learning support. In: Proceedings of the 9th international symposium on information and communication technology, pp 69–76

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

Hofmann J, Troiano E, Sassenberg K, Klinger R (2020) Appraisal theories for emotion classification in text. In: Proceedings of the 28th international conference on computational linguistics, pp 125–138

Hogenboom A, Brojba-Micu A, Frasincar F (2021) The impact of word sense disambiguation on stock price prediction. Expert Syst Appl 184:115568

Hosseini AS (2017) Sentence-level emotion mining based on combination of adaptive meta-level features and sentence syntactic features, vol 65. Elsevier, Amsterdam, pp 361–374

Howard J, Ruder S (2018) Universal language model fine-tuning for text classification. arXiv preprint. arXiv:1801.06146

Huang Y-H, Lee S-R, Ma M-Y, Chen Y-H, Yu Y-W, Chen Y-S (2019a) EmotionX-IDEA: Emotion BERT—an affectional model for conversation. arXiv preprint. arXiv:1908.06264

Huang H, Xie Q, Pan J, He Y, Wen Z, Yu R, Li Y (2019b) An EEG-based brain computer interface for emotion recognition and its application in patients with disorder of consciousness. IEEE Trans Affect Comput 12(4):832–842

Hu M, Zhao S, Guo H, Xue C, Gao H, Gao T, Cheng R, Su Z (2021) Multi-label few-shot learning for aspect category detection. arXiv preprint. arXiv:2105.14174

Ilyas A, Shahzad K, Kamran Malik M (2023) Emotion detection in code-mixed Roman Urdu-English text, vol 22. ACM New York, pp 1–28

Imran AS, Daudpota SM, Kastrati Z, Batra R (2020) Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets. IEEE Access 8:181074–181090

Itani M, Roast C, Al-Khayatt S (2017) Developing resources for sentiment analysis of informal arabic text in social media. Procedia Computer Science 117:129–136

Jabreel M, Moreno A (2019) A deep learning-based approach for multi-label emotion classification in tweets. Appl Sci 9(6):1123

Jain VK, Kumar S, Fernandes SL (2017) Extraction of emotions from multilingual text using intelligent text processing and computational linguistics. J Comput Sci 21:316–326

Jaiswal S, Nandi GC (2020) Robust real-time emotion detection system using CNN architecture. Neural Comput Appl 32(15):11253–11262

Jamal N, Xianqiao C, Al-Turjman F, Ullah F (2021) A deep learning-based approach for emotions classification in big corpus of imbalanced tweets. Trans Asian Low-Resource Lang Inf Process 20(3):1–16

Jayakrishnan R, Gopal GN, Santhikrishna M (2018) Multi-class emotion detection and annotation in Malayalam novels. In: 2018 International conference on computer communication and informatics (ICCCI). IEEE, pp 1–5

Jeyapriya A, Selvi CK (2015) Extracting aspects and mining opinions in product reviews using supervised learning algorithm. In: 2015 2nd International conference on electronics and communication systems (ICECS). IEEE, pp 548–552

Jia Z, Lin Y, Cai X, Chen H, Gou H, Wang J (2020) Sst-emotionnet: Spatial-spectral-temporal based attention 3D dense network for eeg emotion recognition. In: Proceedings of the 28th ACM international conference on multimedia, pp 2909–2917

Jian Z, Chen X, Wang H (2010) Sentiment classification using the theory of anns, vol 17. Elsevier, Amsterdam, pp 58–62

Jose R, Chooralil VS (2015) Prediction of election result by enhanced sentiment analysis on Twitter data using word sense disambiguation. In: 2015 International conference on control communication & computing India (ICCC). IEEE, pp 638–641

Kamath R, Ghoshal A, Eswaran S, Honnavalli P (2022) An enhanced context-based emotion detection model using RoBERTa. In: 2022 IEEE International conference on electronics, computing and communication technologies (CONECCT). IEEE, pp 1–6

Kanchinadam T, Meng Z, Bockhorst J, Singh V, Fung G (2021) Graph neural networks to predict customer satisfaction following interactions with a corporate call center. arXiv preprint. arXiv:2102.00420

Kane A, Patankar S, Khose S, Kirtane N (2022) Transformer based ensemble for emotion detection. arXiv preprint. arXiv:2203.11899

Kang H, Yoo SJ, Han D (2012) Senti-lexicon and improved naïve Bayes algorithms for sentiment analysis of restaurant reviews. Expert Syst Appl 39(5):6000–6010

Kao EC-C, Liu C-C, Yang T-H, Hsieh C-T, Soo V-W (2009) Towards text-based emotion detection a survey and possible improvements. In: 2009 International conference on information management and engineering. IEEE, pp 70–74

Karimi A, Rossi L, Prati A (2021) Adversarial training for aspect-based sentiment analysis with bert. In: 2020 25th International conference on pattern recognition (ICPR). IEEE, pp 8797–8803

Karna M, Juliet DS, Joy RC (2020) Deep learning based text emotion recognition for chatbot applications. In: 2020 4th International conference on trends in electronics and informatics (ICOEI)(48184). IEEE, pp 988–993

Kaur H, Mangat V et al (2017) A survey of sentiment analysis techniques. In: 2017 International conference on I-SMAC (IoT in Social, Mobile, Analytics and cloud) (I-SMAC). IEEE, pp 921–925

Kaur J, Vashisht S (2012) Analysis and identifying variation in human emotion through data mining. Int J Comput Appl Technol 3:1963–1996

Kharat A, Patel A, Bhatt D, Parikh N, Rathore H (2021) Emotion recognition using multimodalities. In: Hybrid intelligent systems: 20th international conference on hybrid intelligent systems (HIS 2020), 14–16 December 2020. Springer, Cham, pp 309–319

Kheiri K, Karimi H (2023) Sentimentgpt: Exploiting gpt for advanced sentiment analysis and its departure from current machine learning. arXiv preprint. arXiv:2307.10234

Kim Y (2014) Convolutional neural networks for sentence classification. arXiv preprint. arXiv:1408.5882

Ko BC (2018) A brief review of facial emotion recognition based on visual information. Sensors 18(2):401

Kohavi R, Provost F (1998) Glossary of terms. Mach Learn 30(2):271–274

Köper M, Kim E, Klinger R (2017) IMS at EmoInt-2017: emotion intensity prediction with affective norms, automatically extended resources and deep learning. In: Proceedings of the 8th workshop on computational approaches to subjectivity, sentiment and social media analysis, pp 50–57

Kosti R, Alvarez JM, Recasens A, Lapedriza A (2017) Emotic: Emotions in context dataset. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 61–69

Kratzwald B, Ilić S, Kraus M, Feuerriegel S, Prendinger H (2018) Deep learning for affective computing: text-based emotion recognition in decision support. Decis Support Syst 115:24–35

Krawczyk B (2016) Learning from imbalanced data: open challenges and future directions. Prog Artif Intell 5(4):221–232

Kulshreshtha D, Goel P, Singh AK (2018) How emotional are you? Neural architectures for emotion intensity prediction in microblogs. In: Proceedings of the 27th international conference on computational linguistics, pp 2914–2926

Kumari K, Singh JP, Dwivedi YK, Rana NP (2021) Bilingual cyber-aggression detection on social media using lstm autoencoder. Soft Comput 25:8999–9012

Kusal S, Patil S, Choudrie J, Kotecha K, Vora D, Pappas I (2022) A review on text-based emotion detection—techniques, applications, datasets, and future directions. arXiv preprint. arXiv:2205.03235

Kusal S, Patil S, Choudrie J, Kotecha K, Vora D, Pappas I (2023) A systematic review of applications of natural language processing and future challenges with special emphasis in text-based emotion detection. Artif Intell Rev 56(12):1–87

Lakomkin E, Bothe C, Wermter S (2018) GradAscent at EmoInt-2017: character-and word-level recurrent neural network models for tweet emotion intensity detection. arXiv preprint. arXiv:1803.11509

Lample G, Ott M, Conneau A, Denoyer L, Ranzato M (2018) Phrase-based & neural unsupervised machine translation. arXiv preprint. arXiv:1804.07755

LeCompte T, Chen J (2017) Sentiment analysis of tweets including emoji data. In: 2017 International conference on computational science and computational intelligence (CSCI). IEEE, pp 793–798

Lee SYM, Wang Z (2015) Multi-view learning for emotion detection in code-switching texts. In: 2015 International conference on asian language processing (IALP). IEEE, pp 90–93

Lee K, Park K, Kim D (2023) Dailytalk: spoken dialogue dataset for conversational text-to-speech. In: ICASSP 2023-2023 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 1–5

Li Y, Su H, Shen X, Li W, Cao Z, Niu S (2017) Dailydialog: a manually labelled multi-turn dialogue dataset. arXiv preprint. arXiv:1710.03957

Li Z, Wu X, Xu X, Wang H, Guo Z, Zhan Z, Yao L (2019a) The recognition of multiple anxiety levels based on electroencephalograph. IEEE Trans Affect Comput 13(1):519–529

Li Y, Zheng W, Wang L, Zong Y, Cui Z (2019b) From regional to global brain: a novel hierarchical spatial-temporal neural network model for EEG emotion recognition. IEEE Trans Affect Comput 13(2):568–578

Li X, Bing L, Zhang W, Lam W (2019c) Exploiting Bert for end-to-end aspect-based sentiment analysis. arXiv preprint. arXiv:1910.00883

Li Q, Yu G, Wang J, Liu Y (2020a) A deep multimodal generative and fusion framework for class-imbalanced multimodal data. Multimedia Tools Appl 79:25023–25050

Li J, Ji D, Li F, Zhang M, Liu Y (2020b) HiTrans: a transformer-based context-and speaker-sensitive model for emotion detection in conversations. In: Proceedings of the 28th international conference on computational linguistics, pp 4190–4200

Li X, Xia Y, Long X, Li Z, Li S (2021) Exploring text-transformers in AAAI 2021 shared task of COVID-19 fake news detection in English. In: International workshop on combating online hostile posts in regional languages during emergency situation. Springer, Cham, pp 106–115

Liao W, Zeng B, Yin X, Wei P (2021) An improved aspect-category sentiment analysis model for text sentiment analysis based on RoBERTa. Appl Intell 51:3522–3533

Liew JSY, Turtle HR, Liddy ED (2016) EmoTweet-28: a fine-grained emotion corpus for sentiment analysis. In: Proceedings of the 10th international conference on language resources and evaluation (LREC'16), pp 1149–1156

Lingren T, Deleger L, Molnar K, Zhai H, Meinzen-Derr J, Kaiser M, Stoutenborough L, Li Q, Solti I (2014) Evaluating the impact of pre-annotation on annotation speed and potential bias: natural language processing gold standard development for clinical named entity recognition in clinical trial announcements. J Am Med Inform Assoc 21(3):406–413

Liu SM, Chen J-H (2015) A multi-label classification based approach for sentiment classification. Expert Syst Appl 42(3):1083–1093

Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V (2019) RoBERTa: a robustly optimized bert pretraining approach. arXiv preprint. arXiv:1907.11692

Liu T, Du Y, Zhou Q (2020) Text emotion recognition using GRU neural network with attention mechanism and emoticon emotions. In: Proceedings of the 2020 2nd international conference on robotics, intelligent control and artificial intelligence, pp 278–282

Liu J, Shen D, Zhang Y, Dolan B, Carin L, Chen W (2021) What makes good in-context examples for GPT-3? arXiv preprint. arXiv:2101.06804

Liu C, Kirubakaran S et al (2022) Deep learning approach for emotion recognition analysis in text streams. Int J Technol Hum Interact (IJTHI) (IJTHI) 18(2):1–21

Liu Z, Nie P, Cai J, Wang H, Niu Z-Y, Zhang P, Sachan M, Peng K (2023) XDailyDialog: a multilingual parallel dialogue corpus. In: Proceedings of the 61st annual meeting of the Association for Computational Linguistics (vol 1: long papers), pp 12240–12253

Lövheim H (2012) A new three-dimensional model for emotions and monoamine neurotransmitters. Med Hypotheses 78(2):341–348

Lu X, Zhang H (2021) Sentiment analysis method of network text based on improved at-bigru model. Sci Program 2021:1–11

Lu R, Li Y, Yan Y (2021) Construction of text emotion classification model based on convolutional neural network. In: 2021 International conference on intelligent computing, automation and applications (ICAA). IEEE, pp 271–274

Lynch C, O'Leary C, Smith G, Bain R, Kehoe J, Vakaloudis A, Linger R (2020) A review of open-source machine learning algorithms for twitter text sentiment analysis and image classification. In: 2020 International joint conference on neural networks (IJCNN). IEEE, pp 1–9

Mac Kim S, Valitutti A, Calvo RA (2010) Evaluation of unsupervised emotion models to textual affect recognition. In: Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text, pp 62–70

Madisetty S, Desarkar MS (2017) Nsemo at emoint-2017: an ensemble to predict emotion intensity in tweets. In: Proceedings of the 8th workshop on computational approaches to subjectivity, sentiment and social media analysis, pp 219–224

Mahima M, Patel NC, Ravichandran S, Aishwarya N, Maradithaya S (2021) A text-based hybrid approach for multiple emotion detection using contextual and semantic analysis. In: 2021 International conference on innovative computing, intelligent communication and smart electrical systems (ICSES). IEEE, pp 1–6

Majeed A, Mujtaba H, Beg MO (2020) Emotion detection in roman urdu text using machine learning. In: Proceedings of the 35th IEEE/ACM international conference on automated software engineering, pp 125–130

Majeed A, Beg MO, Arshad U, Mujtaba H (2022) Deep-EmoRU: mining emotions from roman urdu text using deep learning ensemble. Multimedia Tools Appl 81(30):43163–43188

Malte A, Ratadiya P (2019) Multilingual cyber abuse detection using advanced transformer architecture. In: TENCON 2019-2019 IEEE Region 10 conference (TENCON). IEEE, pp 784–789

Mao R, Liu Q, He K, Li W, Cambria E (2022) The biases of pre-trained language models: An empirical study on prompt-based sentiment analysis and emotion detection. IEEE, ???

Mashal SX, Asnani K (2017) Emotion intensity detection for social media data. In: 2017 International Conference on computing methodologies and communication (ICCMC). IEEE, pp 155–158

McKeown G, Valstar MF, Cowie R, Pantic M (2010) The semaine corpus of emotionally coloured character interactions. In: 2010 IEEE International Conference on Multimedia and Expo, pp. 1079–1084. IEEE

Michelucci U (2022) An introduction to autoencoders. arXiv preprint. arXiv:2201.03898

Mirtalaie MA, Hussain OK, Chang E, Hussain FK (2018) Extracting sentiment knowledge from PROS/CONS product reviews: discovering features along with the polarity strength of their associated opinions. Expert Syst Appl 114:267–288

Mishra S, Asif M, Srinivasan N, Tiwary US (2021) Dataset on emotion with naturalistic stimuli (DENS) on Indian samples. bioRxiv 2021-08

Mishraa S, Asifa M, Tiwaya US (2021) Dataset on emotions using naturalistic stimuli (DENS). bioRxiv. https://doi.org/10.1101/2021.08.04.455041

Mohammad SM, Bravo-Marquez F (2017) Wassa-2017 shared task on emotion intensity. arXiv preprint. arXiv:1708.03700

Mohri M, Rostamizadeh A, Talwalkar A (2018) Foundations of machine learning. MIT, Cambridge

Mozafari F, Tahayori H (2019) Emotion detection by using similarity techniques. In: 2019 7th Iranian joint congress on fuzzy and intelligent systems (CFIS). IEEE, pp 1–5

Mozhdehi MH, Moghadam AME (2023) Textual emotion detection utilizing a transfer learning approach. J Supercomput 79:13075–13089

Murthy AR, Kumar KA (2021) A review of different approaches for detecting emotion from text. IOP Conf Ser Mater Sci Eng 1110:012009

Mustakim N, Rabu R, Mursalin GM, Hossain E, Sharif O, Hoque MM (2022) CUET-NLP@TamilNLP-ACL2022: multi-class textual emotion detection from social media using transformer. In: Proceedings of the 2nd workshop on speech and language technologies for dravidian languages, pp 199–206

Nandwani P, Verma R (2021) A review on sentiment analysis and emotion detection from text. Soc Netw Anal Min 11(1):81

Neviarouskaya A, Aono M (2013) Extracting causes of emotions from text. In: Proceedings of the 6th international joint conference on natural language processing, pp 932–936

Neviarouskaya A, Prendinger H, Ishizuka M (2010) Recognition of fine-grained emotions from text: an approach based on the compositionality principle. In: Modeling machine emotions for realizing intelligence: foundations and applications. Springer, Berlin, pp 179–207

Ngo A, Candri A, Ferdinan T, Kocoń J, Korczynski W (2022) Studemo: A non-aggregated review dataset for personalized emotion recognition. In: Proceedings of the 1st workshop on perspectivist approaches to NLP@ LREC2022, pp 46–55

Nida H, Mahira K, Mudasir M, Mudasir Ahmed M, Mohsin M (2019) Automatic emotion classifier. In: Progress in advanced computing and intelligent engineering: proceedings of ICACIE 2017, vol 1. Springer, Singapore, pp 565–572

Nithya K, Sathyapriya S, Sulochana M, Thaarini S, Dhivyaa C (2022) Deep learning based analysis on code-mixed tamil text for sentiment classification with pre-trained ULMFiT. In: 2022 6th International conference on computing methodologies and communication (ICCMC). IEEE, pp 1112–1116

Nugroho KS, Bachtiar FA (2021) Text-based emotion recognition in Indonesian Tweet using bert. In: 2021 4th International seminar on research of information technology and intelligent systems (ISRITI). IEEE, pp 570–574

Oatley K, Johnson-Laird PN (1987) Towards a cognitive theory of emotions. Cogn Emot 1(1):29–50

Omara E, Mosa M, Ismail N (2019) Emotion analysis in arabic language applying transfer learning. In: 2019 15th International computer engineering conference (ICENCO). IEEE, pp 204–209

Ortiz-Ospina E, Roser M (2023) The rise of social media. Our world in data. https://ourworldindata.org/rise-of-social-media. Accessed 4 Apr 2023

Ouyang T, Nguyen-Son H-Q, Nguyen HH, Echizen I, Seo Y (2023) Quality assurance of a GPT-based sentiment analysis system: adversarial review data generation and detection. arXiv preprint. arXiv:2310.05312

Pandey AC, Seth SR, Varshney M (2019) Sarcasm detection of Amazon Alexa sample set. In: Advances in signal processing and communication: select proceedings of ICSC 2018. Springer, Singapore, pp 559–564

Park S-H, Bae B-C, Cheong Y-G (2020) Emotion recognition from text stories using an emotion embedding model. In: 2020 IEEE international conference on big data and smart computing (BigComp). IEEE, pp 579–583

Pascanu R, Mikolov T, Bengio Y (2013) On the difficulty of training recurrent neural networks. In: International conference on machine learning, pp 1310–1318. PMLR

Pasupa K, Ayutthaya TSN (2019) Thai sentiment analysis with deep learning techniques: a comparative study based on word embedding, POS-tag, and sentic features, vol 50. Elsevier, Amsterdam, p 101615

Perikos I, Hatzilygeroudis I (2016) Recognizing emotions in text using ensemble of classifiers. Eng Appl Artif Intell 51:191–201

Plutchik R (1980) Emotion. A psychoevolutionary synthesis. Harper and Row, New York

Polignano M, Basile P, Gemmis M, Semeraro G (2019) 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, pp 63–68

Poria S, Gelbukh A, Hussain A, Howard N, Das D, Bandyopadhyay S (2013) Enhanced senticnet with affective labels for concept-based opinion mining. IEEE Intell Syst 28(2):31–38

Poria S, Cambria E, Gelbukh A (2016) Aspect extraction for opinion mining with a deep convolutional neural network. Knowl Based Syst 108:42–49

Poria S, Hazarika D, Majumder N, Naik G, Cambria E, Mihalcea R (2018) MELD: a multimodal multiparty dataset for emotion recognition in conversations. arXiv preprint. arXiv:1810.02508

Poria S, Majumder N, Mihalcea R, Hovy E (2019) Emotion recognition in conversation: research challenges, datasets, and recent advances. IEEE Access 7:100943–100953

Potamias RA, Siolas G, Stafylopatis A-G (2020) A transformer-based approach to irony and sarcasm detection, vol 32. Springer, Berlin, pp 17309–17320

Pradhan VM, Vala J, Balani P (2016) A survey on sentiment analysis algorithms for opinion mining. Int J Comput Appl 133(9):7–11

Puranik K, Hande A, Priyadharshini R, Thavareesan S, Chakravarthi BR (2021) IIIT_DWD@LT-EDI-EACL2021: hope speech detection: there is always hope in transformers. arXiv preprint. arXiv:2104.09066

Qiu X, Qian P, Shi Z (2016) Overview of the NLPCC-ICCPOL 2016 shared task: Chinese word segmentation for micro-blog texts. In: Natural language understanding and intelligent applications: 5th CCF conference on natural language processing and Chinese computing, NLPCC 2016, and 24th International conference on computer processing of oriental languages, ICCPOL 2016, Kunming, China, 2–6 December 2016, proceedings, vol 24. Springer, Cham, pp 901–906

Quan C, Ren F (2009) Recognizing sentence emotions based on polynomial kernel method using REN-CECPS. In: 2009 International conference on natural language processing and knowledge engineering. IEEE, pp 1–7

Quan C, Ren F (2010a) A blog emotion corpus for emotional expression analysis in Chinese. Comput Speech Lang 24(4):726–749

Quan C, Ren F (2010b) Automatic annotation of word emotion in sentences based on ren-cecps. In: LREC

Rabeya T, Ferdous S, Ali HS, Chakraborty NR (2017) A survey on emotion detection: a lexicon based backtracking approach for detecting emotion from Bengali text. IEEE, pp 1–7

Rachman FH, Sarno R, Fatichah C (2016) CBE: corpus-based of emotion for emotion detection in text document. In: 2016 3rd International conference on information technology, computer, and electrical engineering (ICITACEE). IEEE, pp 331–335

Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I et al (2019) Language models are unsupervised multitask learners. Open AI Blog 1(8):9

Rahaman MS, Ahsan MT, Anjum N, Terano HJR, Rahman MM (2023) From ChatGPT-3 to GPT-4: a significant advancement in AI-driven NLP tools. J Eng Emerg Technol 2(1):1–11

Rahman L, Mohammed N, Al Azad AK (2016) A new LSTM model by introducing biological cell state. In: 2016 3rd International conference on electrical engineering and information communication technology (ICEEICT). IEEE, pp 1–6

Rajalingam B, Al-Turjman F, Santhoshkumar R, Rajesh M (2022) Intelligent multimodal medical image fusion with deep guided filtering. Multimedia Syst 28(4):1449–1463

Rayhan MM, Al Musabe T, Islam MA (2020) Multilabel emotion detection from Bangla text using bigru and CNN-BILSTM. In: 2020 23rd International conference on computer and information technology (ICCIT). IEEE, pp 1–6

Riza MA, Charibaldi N (2021) Emotion detection in Twitter social media using long short-term memory (LSTM) and fast text. Int J Artif Intell Robot (IJAIR) 3(1):15–26

Robinson T (1991) Several improvements to a recurrent error propagation network phone recognition system. Cambridge University Engineering Department Technical Report. https://doi.org/10.13140/RG.2.2.15418.90567

Rui H, Liu Y, Whinston A (2013) Whose and what chatter matters? the effect of tweets on movie sales. Decis Support Syst 55(4):863–870

Sadeghi SS, Khotanlou H, Rasekh Mahand M (2021) Automatic persian text emotion detection using cognitive linguistic and deep learning. J AI Data Min 9(2):169–179

Saffar AH, Mann TK, Ofoghi B (2023) Textual emotion detection in health: Advances and applications. J Biomed Inform 137:104258

Sagha H, Cummins N, Schuller B (2017) Stacked denoising autoencoders for sentiment analysis: a review. Wiley Interdiscip Rev Data Min Knowl Discov 7(5):1212

Sailunaz K, Alhajj R (2019) Emotion and sentiment analysis from twitter text. J Comput Sci 36:101003

Sailunaz K, Dhaliwal M, Rokne J, Alhajj R (2018) Emotion detection from text and speech: a survey. Soc Netw Anal Min 8:1–26

Samy AE, El-Beltagy SR, Hassanien E (2018) A context integrated model for multi-label emotion detection. Procedia Comput Sci 142:61–71

Sangeetha K, Prabha D (2021) Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM. J Ambient Intell Humaniz Comput 12:4117–4126

Sanh V, Debut L, Chaumond J, Wolf T (2019) Distilbert, a distilled version of Bert: smaller, faster, cheaper and lighter. arXiv preprint. arXiv:1910.01108

Sasidhar TT, Premjith B, Soman K (2020) Emotion detection in hinglish (hindi+ english) code-mixed social media text. Procedia Comput Sci 171:1346–1352

Schuller BW (2018) Speech emotion recognition: two decades in a nutshell, benchmarks, and ongoing trends. Commun ACM 61(5):90–99

Seal D, Roy UK, Basak R (2020) Sentence-level emotion detection from text based on semantic rules. In: Information and communication technology for sustainable development: proceedings of ICT4SD 2018. Springer, Singapore, pp 423–430

Serte S, Serener A, Al-Turjman F (2022) Deep learning in medical imaging: a brief review. Trans Emerg Telecommun Technol 33(10):4080

Seyeditabari A, Tabari N, Gholizadeh S, Zadrozny W (2019) Emotion detection in text: focusing on latent representation. arXiv preprint. arXiv:1907.09369

Seyeditabari A, Tabari N, Zadrozny W (2018) Emotion detection in text: a review. arXiv preprint. arXiv:1806.00674

Sharif O, Hossain E, Hoque MM (2021) NLP-CUET@DravidianLangTech-EACL2021: offensive language detection from multilingual code-mixed text using transformers. arXiv preprint. arXiv:2103.00455

Sharma A (2020) Hindi text emotion recognition based on deep learning. IOSR J Mob Comput Appl (IOSR-JMCA) 7(3):24–29

Sharma A, Sharma K, Kumar A (2023) Real-time emotional health detection using fine-tuned transfer networks with multimodal fusion. Neural Comput Appl 35(31):22935–22948

Shaver P, Schwartz J, Kirson D, O'connor C (1987) Emotion knowledge: further exploration of a prototype approach. J Pers Soc Psychol 52(6):1061

Shi L, Liu D, Liu G, Meng K (2020) AUG-BERT: an efficient data augmentation algorithm for text classification. In: Communications, Signal Processing, and Systems: Proceedings of the 8th international conference on communications, signal processing, and systems. Springer, Singapore, pp 2191–2198

Shirsat VS, Jagdale RS, Deshmukh SN (2019) Sentence level sentiment identification and calculation from news articles using machine learning techniques. In: Computing, communication and signal processing: proceedings of ICCASP 2018. Springer, Singapore, pp 371–376

Shivhare SN, Khethawat S (2012) Emotion detection from text. arXiv preprint. arXiv:1205.4944

Shrivastava K, Kumar S, Jain DK (2019) An effective approach for emotion detection in multimedia text data using sequence based convolutional neural network. Multimedia Tools Appl 78:29607–29639

Sikhi Y, Devi SA, Jasti SK, Ram MS (2022) Sentimental analysis through speech and text for IMDB dataset. In: 2022 4th International conference on smart systems and inventive technology (ICSSIT). IEEE, pp 1519–1522

Song D, Oh EY, Rice M (2017) Interacting with a conversational agent system for educational purposes in online courses. In: 2017 10th International conference on human system interactions (HSI). IEEE, pp 78–82

Sosea T, Caragea C (2020) CancerEmo: a dataset for fine-grained emotion detection. In: Proceedings of the 2020 conference on empirical methods in natural language processing, pp 8892–8904

Stanton G, Irissappane AA (2019) GANS for semi-supervised opinion spam detection. arXiv preprint. arXiv:1903.08289

Strapparava C, Mihalcea R (2007) Semeval-2007 task 14: affective text. In: Proceedings of the 4th international workshop on semantic evaluations (SemEval-2007), pp 70–74

Su J, Tang J, Jiang H, Lu Z, Ge Y, Song L, Xiong D, Sun L, Luo J (2021) Enhanced aspect-based sentiment analysis models with progressive self-supervised attention learning. Artif Intell 296:103477

Suhaeni C, Yong H-S (2023) Mitigating class imbalance in sentiment analysis through GPT-3-generated synthetic sentences. Appl Sci 13(17):9766

Suhasini M, Srinivasu B (2020) Emotion detection framework for twitter data using supervised classifiers. In: Data engineering and communication technology: proceedings of 3rd ICDECT-2K19. Springer, Singapore, pp 565–576

Sumanathilaka T, Selvarai V, Raj U, Raiu VP, Prakash J (2021) Emotion detection using bi-directional LSTM with an effective text pre-processing method. In: 2021 12th International conference on computing communication and networking technologies (ICCCNT). IEEE, pp 1–4

Sumanth C, Inkpen D (2015) How much does word sense disambiguation help in sentiment analysis of micropost data? In: Proceedings of the 6th workshop on computational approaches to subjectivity, sentiment and social media analysis, pp 115–121

Sun Z, Han L, Huang W, Wang X, Zeng X, Wang M, Yan H (2015) Recommender systems based on social networks. J Syst Softw 99:109–119

Sun S, Luo C, Chen J (2017) A review of natural language processing techniques for opinion mining systems. Inf Fusion 36:10–25

Sutskever I, Martens J, Hinton GE (2011) Generating text with recurrent neural networks. In: Proceedings of the 28th International conference on machine learning (ICML-11), pp 1017–1024

Swain M, Routray A, Kabisatpathy P (2018) Databases, features and classifiers for speech emotion recognition: a review. Int J Speech Technol 21:93–120

Tan KL, Lee CP, Anbananthen KSM, Lim KM (2022) ROBERTA-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network. IEEE Access 10:21517–21525

Tang R, Lu Y, Liu L, Mou L, Vechtomova O, Lin J (2019) Distilling task-specific knowledge from Bert into simple neural networks. arXiv preprint arXiv:1903.12136

Tejaswini V, Babu KS, Sahoo B (2022) Depression detection from social media text analysis using natural language processing techniques and hybrid deep learning model. ACM Transactions on Asian and Low-Resource Language Information Processing

Teoh T-T, Cho S-Y (2011) Human emotional states modeling by hidden Markov model. In: 2011 7th International conference on natural computation

Thakkar G, Preradovic NM, Tadic M (2022) Multi-task learning for cross-lingual sentiment analysis. arXiv preprint. arXiv:2212.07160

Thomas N (2016) An e-business Chatbot using AIML and LSA. In: 2016 International conference on advances in computing, communications and informatics (ICACCI). IEEE, pp 2740–2742

Topal K, Ozsoyoglu G (2016) Movie review analysis: emotion analysis of IMDB movie reviews. In: 2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM). IEEE, pp 1170–1176

Topal MO, Bas A, Heerden I (2021) Exploring transformers in natural language generation: GPT, Bert, and XLNET. arXiv preprint. arXiv:2102.08036

Tripathi S, Beigi H (2018) Multi-modal emotion recognition on iemocap with neural networks. arXiv preprint. arXiv:1804.05788

Tripathi S, Tripathi S, Beigi H (2018) Multi-modal emotion recognition on IEMOCAP dataset using deep learning. arXiv preprint. arXiv:1804.05788

Tripto NI, Ali ME (2018) Detecting multilabel sentiment and emotions from bangla youtube comments. In: 2018 International conference on Bangla speech and language processing (ICBSLP). IEEE, pp 1–6

Truşcă MM, Frasincar F (2023) Survey on aspect detection for aspect-based sentiment analysis. Artif Intell Rev 56(5):3797–3846

Tsoumakas G, Vlahavas I (2007) RAndom k-labELsets: an ensemble method for multilabel classification. In: European conference on machine learning. Springer, Berlin, pp 406–417

Turcan E, Muresan S, McKeown K (2021) Emotion-infused models for explainable psychological stress detection. In: Proceedings of the 2021 conference of the North American chapter of the Association for Computational Linguistics: human language technologies, pp 2895–2909

Tyng CM, Amin HU, Saad MN, Malik AS (2017) The influences of emotion on learning and memory. Front Psychol 8:1454

Tzacheva A, Ranganathan J, Mylavarapu SY (2020) Actionable pattern discovery for tweet emotions. In: Advances in artificial intelligence, software and systems engineering: proceedings of the AHFE 2019

international conference on human factors in artificial intelligence and social computing, the AHFE international conference on human factors, software, service and systems engineering, and the AHFE international conference of human factors in energy, 24–28 July 2019, Washington DC, USA, vol 10. Springer, Cham, pp 46–57

Ullah F, Wang J, Jabbar S, Al-Turjman F, Alazab M (2019) Source code authorship attribution using hybrid approach of program dependence graph and deep learning model. IEEE Access 7:141987–141999

Ullah F, Jabbar S, Al-Turjman F (2020) Programmers' de-anonymization using a hybrid approach of abstract syntax tree and deep learning. Technol Forecast Soc Chang 159:120186

Varshney D, Ekbal A, Cambria E (2024) Emotion-and-knowledge grounded response generation in an open-domain dialogue setting. Knowl Based Syst 284:111173

Wadhawan A, Aggarwal A (2021) Towards emotion recognition in Hindi-English code-mixed data: a transformer based approach. arXiv preprint. arXiv:2102.09943

Wang D, Li J, Xu K, Wu Y (2017) Sentiment community detection: exploring sentiments and relationships in social networks. Electron Commer Res 17:103–132

Wang Y, Reiterer H, Epstein L, Gordy C, Raynor H, Beddome M, Kilanowski C, Paluch R, Bech P, Olsen L et al (2019) Assistive conversational agent for health coaching: a validation study. Methods Inf Med 58(01):009–023

Wang Y, Huang G, Li M, Li Y, Zhang X, Li H (2023) Automatically constructing a fine-grained sentiment lexicon for sentiment analysis. Cogn Comput 15(1):254–271

Widyassari AP, Rustad S, Shidik GF, Noersasongko E, Syukur A, Affandy A et al (2022) Review of automatic text summarization techniques & methods. J King Saud Univ Comput Inf Sci 34(4):1029–1046

Wikarsa L, Thahir SN (2015) A text mining application of emotion classifications of Twitter's users using naive bayes method. In: 2015 1st International conference on wireless and telematics (ICWT). IEEE, pp 1–6

Wu Y, Guo Y (2020) Dual adversarial co-learning for multi-domain text classification. In: Proceedings of the AAAI conference on artificial intelligence, vol 34, pp 6438–6445

Xiang Z, Du Q, Ma Y, Fan W (2017) A comparative analysis of major online review platforms: implications for social media analytics in hospitality and tourism. Tour Manage 58:51–65

Xu T, Zhou Y, Wang Z, Peng Y (2018) Learning emotions EEG-based recognition and brain activity: a survey study on BCI for intelligent tutoring system. Procedia Comput Sci 130:376–382

Xu J, Hu Z, Zou J, Bi A (2019a) Intelligent emotion detection method based on deep learning in medical and health data, vol 8. IEEE, pp 3802–3811

Xu H, Liu B, Shu L, Yu PS (2019b) Bert post-training for review reading comprehension and aspect-based sentiment analysis. arXiv preprint. arXiv:1904.02232

Xu G, Li W, Liu J (2020) A social emotion classification approach using multi-model fusion. Futur Gener Comput Syst 102:347–356

Xue Z, Ko T-Y, Yuchen N, Wu M-KD, Hsieh C-C (2018) ISA: Intuit smart agent, a neural-based agent-assist chatbot. In: 2018 IEEE international conference on data mining workshops (ICDMW). IEEE, pp 1423–1428

Yamin MNM, Aziz KA, Siang TG, Aziz NAA (2023) A systematic literature review on emotion recognition system in Malaysia. In: Proceedings of the international conference on artificial life and robotics, Sapporo, Japan, pp 24–28

Yan W, Zhou L, Qian Z, Xiao L, Zhu H (2021) Sentiment analysis of student texts using the CNN-BIGRU-AT model. Sci Program 2021:1–9

Yang Z, Dai Z, Yang Y, Carbonell J, Salakhutdinov RR, Le QV (2019) XLNet: generalized autoregressive pretraining for language understanding. In: Advances in neural information processing systems, vol 32

Yang L, Lin H et al (2011) Text based emotion transformation analysis. Comput Eng Sci 33(9):123

Yasaswini K, Puranik K, Hande A, Priyadharshini R, Thavareesan S, Chakravarthi BR (2021) IIITT@ DravidianLangTech-EACL2021: transfer learning for offensive language detection in dravidian languages. In: Proceedings of the first workshop on speech and language technologies for Dravidian languages, pp 187–194

Ye Z, Zuo T, Chen W, Li Y, Lu Z (2023) Textual emotion recognition method based on ALBERT-BiLSTM model and SVM-NB classification. Soft Comput 27(8):5063–5075

Yenter A, Verma A (2017) Deep CNN-LSTM with combined kernels from multiple branches for IMDB review sentiment analysis. IEEE, pp 540–546

Yin W (2020) Meta-learning for few-shot natural language processing: a survey. arXiv preprint. arXiv:2007.09604

Yu Y, Kim Y-J (2020) Attention-LSTM-attention model for speech emotion recognition and analysis of IEMOCAP database. Electronics 9(5):713

Zhang L, Wang S, Liu B (2018a) Deep learning for sentiment analysis: a survey. Wiley Interdiscip Rev Data Min Knowl Disc 8(4):1253

Zhang Y, Song D, Li X, Zhang P (2018b) Unsupervised sentiment analysis of twitter posts using density matrix representation. In: Advances in information retrieval: 40th European conference on IR research, ECIR 2018, Grenoble, France, 26–29 March 2018, proceedings, vol 40. Springer, Cham, pp 316–329

Zhang WE, Sheng QZ, Alhazmi A, Li C (2020) Adversarial attacks on deep-learning models in natural language processing: a survey. ACM Trans Intell Syst Technol (TIST) 11(3):1–41

Zhang J, Yan K, Mo Y (2021) Multi-task learning for sentiment analysis with hard-sharing and task recognition mechanisms. Information 12(5):207

Zhao X, Wong C-W (2023) Automated measures of sentiment via transformer-and lexicon-based sentiment analysis (TLSA). J Comput Soc Sci. https://doi.org/10.36227/techrxiv.21781109.v1

Zucco C, Liang H, Di Fatta G, Cannataro M (2018) Explainable sentiment analysis with applications in medicine. In: 2018 IEEE international conference on bioinformatics and biomedicine (BIBM). IEEE, pp 1740–1747