Engineered Science
DOI: https://dx.doi.org/10.30919/es1650

Comparative Analysis of Emotion Detection from Social Media Text using Machine Learning and Deep Learning

Vinayak Malavade,[1,2,*] Virat Giri,[1] Shruti Patil,[3,*] Deepali Jadhav,[4] Sheetal Kusal[3] and Jyoti Deshmukh[5]

Abstract

Emotions plays a crucial role in daily life. Detection of emotions from social media text is paying attention in sentiment analysis, social monitoring. Enhancement of internet technology increased use of social media. Nowadays social media text is used for wide varieties of application as it contains implicit expressions, feelings, thoughts, moods. Sentiment analysis extracts sentiments, aspect, emotion, opinion and view from the text, videos, audio. Extraction of accurate sentiments or emotions from users and to analyze such data is a challenging task. Various approaches are developed to extract emotion from text which includes classification algorithms, machine learning, deep learning techniques, rule-based approach, hybrid approach etc. Extraction of implicit emotion from user text is still a challenging area. This study aims to give insights into analysis of different machine learning algorithms such as Support Vector Machine, Random Forest Classifier, Decision Tree, Extreme Gradient Boosting, Naive Bayes Classifier and deep learning techniques such as Convolutional Neural Network, Long-Short Term Memory, Hybrid models for detection of emotion from social media text based on two different online available datasets, Performance of models is measured with performance metrics accuracy and losses.

Keywords: Social media text; Emotion detection; Machine learning; Deep learning.
Received: 29 November 2024; Revised: 08 March 2025; Accepted: 12 May 2025
Article type: Research article.

1. Introduction

The use of internet technology and Social Media (SM) is increasing enormously. All over the world (7.8 billion individuals) approximately 50.64% of the earth population make use of SM.[1] Increased use of SM platforms has generated huge amounts of data. Such a huge amount of data shows crucial role in different real time applications. Sentiment Analysis (SA) is the application in which a lot of study is made but still it is associated with varieties of issues. It is associated with extraction of sentiments from data. It extracts opinion, emotion, aspect, view, feelings from the text, images, videos or audios. SA classifies the words in to positive, negative and neutral.[2] SA can be categorized in to four main categories *i.e.* Fine-grained sentiment, Aspect-based, Intent analysis and Emotion Detection, Emotion Detection (ED) is a sub-type of SA.[3] Different approaches or algorithms are developed to perform SA. Recent studies in SA focus to improve emotion-based SA of dialect speech using XGBoost and RNN using acoustic features. Gate Control Unit (GCU) was developed for acoustic features.[4] Optimized Deep Learning (ODL) Model (FTDLM) proposed to detect facial expressions for peoples automatically with neurological disorders in facial muscle to identify facial expressions for stroke, Parkinson's, Alzheimer's, and Bell's palsy.[5]

ED is a type of SA associated with extraction of the user's implicit feelings, thoughts, mood, subjective information etc. Social media becomes a platform in which people implicitly share their views, feelings, moods etc. Understanding emotions expressed in SM text is useful for various applications, including SA, customer feedback interpretation, and mental health monitoring, and enhancing human-computer interactions. It is Natural Language Processing (NLP) task which is used to extract contextual information. Extraction of such implicit emotions from text is one of the challenging tasks. Traditional approaches to ED have relied heavily on manual annotations and rule-based systems, which are often limited in scalability and accuracy. With advancements in Machine Learning (ML), particularly in NLP, it has become possible to develop sophisticated models capable of automatically detecting emotions with high precision. ED is an NLP technique that detects the emotion in text and gives emotion type.[2] In the last decade evolution of

[1]*Department of Computer Science and Engineering, Sanjay Ghodawat University, Kolhapur, Maharashtra, 416012, India*
[2]*Department of Computer Engineering (Regional Language), Pimpri Chinchwad College of Engineering, Nigdi, Pune, Maharashtra, 411044, India*
[3]*Department of Computer Engineering, Sanjay Ghodawat Institute, Atigre, Kolhapur, Maharashtra, 416012, India*

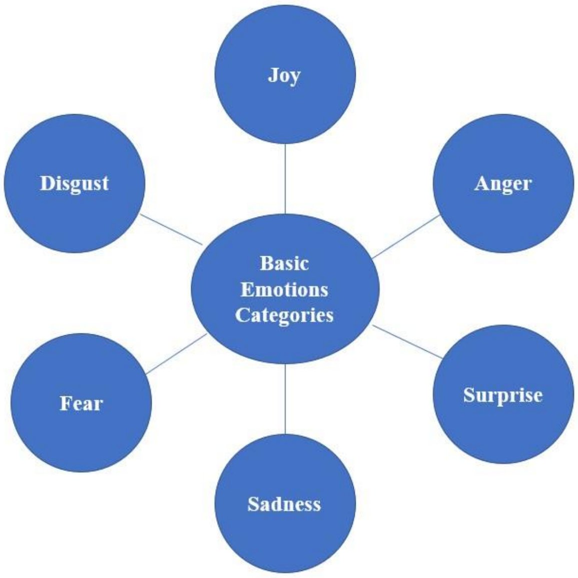


Fig. 1: Ekamn's basic emotion categories.

DL techniques gives significant improvements in results for different domains. ML is a subfield of artificial intelligence involving traditional algorithms like Support Vector Machines (SVM), Logistic Regression (LR), Decision Tree (DT) classifier, Random Forests (RF), Stochastic Gradient Descent (SGD), Gradient Boost (GB), Extreme GB (XGB), K- Nearest Neighbors (KNN) , Multinomial Naive Bayes (MNB) and Naive Bayes (NB) classifier, etc. which learns patterns from data and gives predictions without explicit program. ML often requires manual feature extractions. DL subfield of AI and a special branch of ML involving architectures such as Artificial Neural Network (ANN) Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM) which mimics the structure of a neural network (human brain) also called ANN to extract features. DL models are more complex and need a large dataset. ML algorithms are used to explore simple tasks with smaller datasets whereas DL techniques used to process large applications or unstructured data like image audio, video etc.

From the last two decades, the evolving era of AI and its techniques giving promising results in various applications. ML and DL are the booming areas of research. ML provided valuable insights in various applications such as healthcare for early disease detection such as predicting late onset neonatal sepsis using ensemble methods, mode shows F1 score 0.7829, accuracy 0.8553 and AUROC 0.9266.[6] Solutions for geotechnical engineering challenges, such as predicting the seismic stability of excavations with the help of RF models optimized with algorithms like whale, dragonfly, and sparrow search optimization.[7] also, contribute in automation of the complex process of control system design *e.g.* synthesize the control functions for stabilizing object movement along a predefined path, synthesis of motion etc.[8] Prediction of Engineering Demand Parameter (EDP) *e.g.* vulnerability Information of huge tubular buildings or structures.[9] Hybrid DL models CNN-GRU helps in Infrastructure durability and building processes such as predicting ultra-high-performance concrete (UHPC) flowability, operations etc.[10]

This research paper presents an analysis of various ML algorithms and DL techniques used for ED by evaluating the performance of different algorithms, and also aims to find the best effective algorithm for accurately capturing the emotions hidden in SM content. Study can contribute to the researchers in the domain of detection of emotions, providing insights for efficient SA based systems development in real-world applications. Different studies defined different numbers of emotions. Paul Ekman projected the emotion model baseline. The study proposes six basic emotions as surprise, sadness, fear, anger, disgust and happiness. These six emotions are considered as a basic emotion types.[11] Two more emotions added such as trust and anticipation into proposed eight primary emotions including basic emotion.[12] Fig. 1 shows Paul Ekman's basic categories of emotions which includes Joy, Sadness, Anger, Fear, Surprise and Disgust.

Existing work on ED includes various approaches such as lexicon based, dictionary based, ML algorithms, DL techniques, rule-based, hybrid approach of DL etc. Review on text ED is presented in which papers from various publications like Science Direct, IEEE, Scopus etc. from 2005 to 2021 are considered, and reviews the applications in research. It gives overview on models of emotion, methods for feature extraction, various datasets, and its challenges along with future scope.[13] Comparative analysis of various ED techniques from SM data is focused which includes ML approaches with different hyper- parameters tuning and DL approaches using the pre-trained model, hybrid Ensemble Techniques (ET).[3] Work on ML models by making use of term frequenc-inverse document frequency (TF-IDF), count vectors are focused, also DL model were used. DL models proved better results than ML models in performance with an accuracy of 92%.[14] Problem of twitter sentiment analysis with the help of DL techniques were focused. Recurrent Neural Network (RNN) along with CNN models are constructed and model architectures were trained on 600,000 tweets.[15] ML algorithms and Lexicon-based approach were used for ED, also developed web application.[16] Supervised ML algorithms are including NB, SVM algorithms. Study on ML techniques to large language models (LLMs) is performed in which results shows fine-tuned Mistral 7B model outperforms with

[4]*Department of Artificial Intelligence and Machine Learning, Symbiosis Institute of Technology, Symbiosis International Deemed University, Pune, Maharashtra, 412115, India*
[5]*Department of Information Technology, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India*
[6]*Department of Artificial Intelligence and Data Science, Marathwada Mitra Mandal's Institute of Technology, Lohagaon, Pune, Maharashtra, 411047, India*
[*]*E-mail:* vnmalavade@gmail.com (V. Malavade), shruti.patil@sitpune.edu.in (S. Patil)

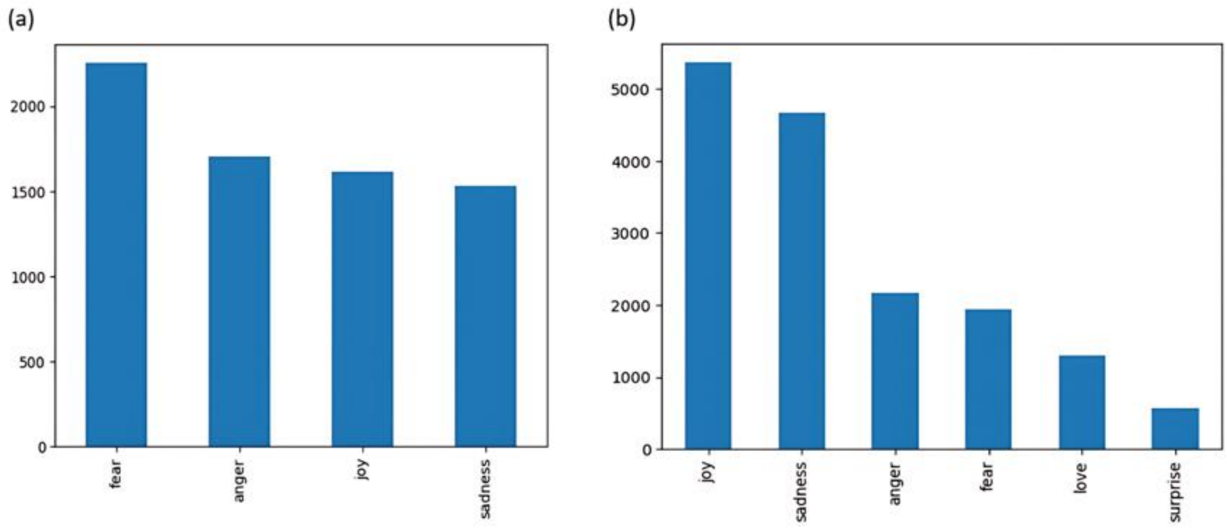


Fig. 2: Datasets: (a)Dataset_1; (b) Dataset_2.

an of 76% of accuracy. SVM attained an accuracy of 64%.[17] Strengths and weaknesses of each algorithm and how various ML algorithms are used for classifying emotions in text are discussed.[18] A per-trained model Bidirectional Encoder Representations from Transformers (BERT) shows significant results in the domain. BERT is used to capture contextual information also retrained models are used to improve performance of ED models.[19] A survey on various emotion identification approaches presented along with its features, their advantages and limitations. The survey is made implicit and explicit Emotion Recognition (ER) in text in which they highlight the impacts of NLP concepts such as tagging, parsing etc. on different approaches. Results showed that DL approaches give enhanced results than the other approaches. Use of handcrafted features and word embedding in ML or DL approaches represents feature scope.[20] Text based Emotion identification and ED model were proposed using NLP and NN, where NN represents promising results.[21] ED from text it advances, challenges, and opportunities discussed in which authors discuss on recent proposals with respect to their contributions, results, strengths, and weaknesses.[22] Semantic analysis of text with the help of DL using Big data for ED proposed.[23] Sentiment analysis is lacking in some applications thus requires ED. Study on various levels of sentiment analysis, emotion models and text ED presented.[24] New application of Roberta model for to finding emotion change over time of different citizens of world is developed.[25] Two million tweets were collected and analyzed using Transfer Learning and Robustly Optimized BERT Pre-training Approach (RoBERTa). A multi-class emotion classifier system was formed by using RoBERTa and Twitter dataset. RNN models, LSTM, BiLSTM, and GRU performances are discussed and evaluated using the ISEAR dataset for ED model. The result shows GRU achieved the highest score achieved 60.26% accuracy, BiLSTM and LSTM accuracies are 59.3% and 57.65% respectively.[26] DL models for detection of emotions *e.g.* Happy, Sad and Angry from the text dialogues using a combined approach of semantics and sentiment-based representations is proposed.[27] Detection of additional emotional classes such as Surprise, Fear, and Disgust with training on context of the dialogue is still limited in this research. Investigation on new approach for improving code-switched ED from text *i.e.* encoder architecture for monolingual and bilingual features proposed in which two

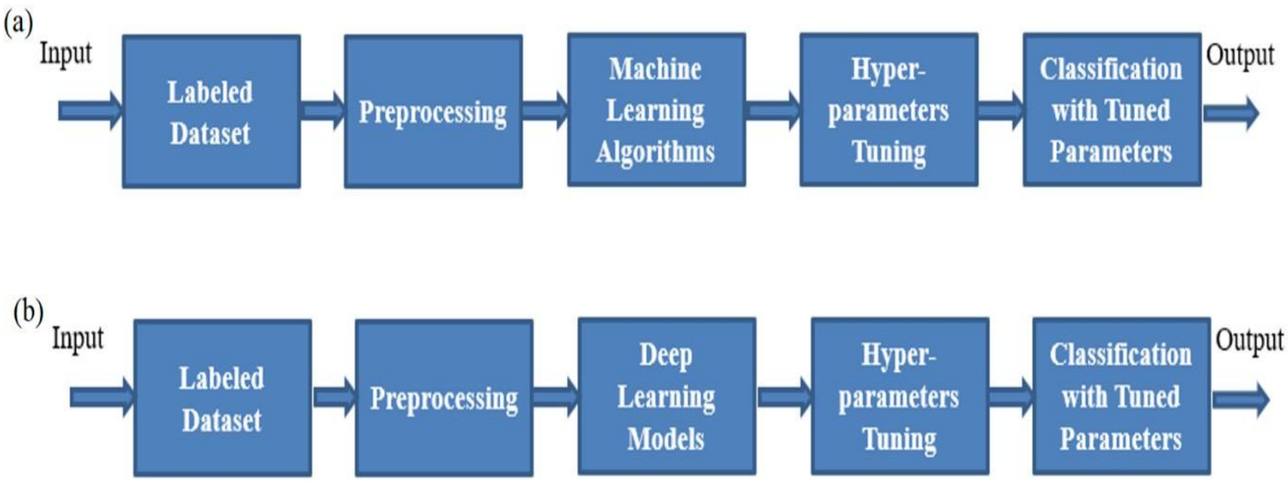


Fig. 3: Emotion Detection Steps: (a) ML; (b) DL.

languages were used with parallel translation.[28] Dataset of 12000 Hindi-English code-mixed annotated text were created. To get feature vectors a pre-trained bilingual model is used and classification is done using DNN. Hybrid model CNN-BiLSTM gives better accuracy of 83.21% for detection of emotions.[29] Combined sparse and dense representations are used along with ET approach using pre-trained, dense word embedding. Models developed and experiment on the datasets of ED from different domains.[30] Classification using approaches of Bi-LSTM, DNN, CNN, self-attention is proposed and applied on different datasets, also pre-trained models were compared for encoding of words.[31] Ontology-based models for ED from text proposed.[32] After considerable advancements, still several challenges remain in domain from

Table 1: Existing techniques and its performance for ED.

Reference	Dataset	Techniques / Algorithms	Performances (Accuracy in %)	Challenges / Limitation / Future scope
[2]	Kaggle, 20000 sentences	DL, LR, SVM, NB	NB: 73.08%, SGD: 89.79%, LR: 88.10%, DL: 89.76%	Limited Size of dataset
[3]	Kaggle, 20000 sentences	MNB, SVM, KNN, DT, LR, RF, Bi-LSTM and BERT	BI-LSTM and BERTL: 94%, LR: 84%, Hybrid models: 82%	Exploration of other application domains
[14]	Kaggle, 20000 sentences	DT,RF, SVM, LR, DL	ANN: 92%.	Imbalance dataset
[16]	Kaggle, 20000 sentences	Lexicon-based approach, NB, SVM, NN	Hybrid model NN, CNN and RN: 90%,	Multiple emotions, sarcasm, irony idioms, metaphors, homonyms in text
[17]	ISEAR (Dataset 1)	BERT, Falcon 7B, and Mistral 7B (M7B)	M7B: 76%, SVM: 64%.	Exploration of one-shot, multi-shot learning techniques
[18]	EmoEval2022	LR, SVM, GB, RF	LR: 62.23%	Limited size of dataset
[21]	Comments from SM 27000 words	NLP and NN, NRCLex	NRCLex : 64.0%, NLP: 83.0%, NN: 99.0%.	Other DL techniques to improve performance
[26]	ISEAR (Dataset 1)	LSTM, Bi-LSTM, GRU	GRU 60.26%, Bi-LSTM 59.3%, LSTM 57.65%.	Pre-processing can be enhanced, performance can be tested with different datasets with respect to quantity and quality
[34]	ISEAR, WASSA, Emotion-stimulus	Natural language processing (NLP) and DL, RD, SVM, proposed LED_LSMD_MDLT model	RF 81, SVM 83, LED_LSMD_MDLT 85 for ISEAR RF 82, SVM 84, LED_LSMD_MDLT 86 for WASSA RF 85, SVM 89, LED_LSMD_MDLT 91 for Emotion-Stimulus	Not Mentioned
[35]	SemEval 2019 Task 3	Ensemble DL	RNN-based models: 67.03%, Ensemble model 77.07%.	Complexity of the training process, overhead, lack of auto-tuning techniques, Imbalance dataset
[36]	SemEval 2019 Task3 and ISEAR	BERT-CNN	BERT-CNN 94.7% semeval2019 task3 dataset and 75.8% for ISEAR dataset.	Performances can be improved using NN and other pretrained transformer models, Large dataset with multiple languages can be used
[39]	ISEAR, WASSA, Emotion-Stimulus	SVM, RF, NB, DT, GRU, Bi-GRU, CNN, CNN+Bi-GRU+SVM	SVM 78.97, Bi-GRU 79.46%, CNN F1-score 80.76. Hybrid model 80.11%.	ET can be used to improve the results in regional languages. Real-time model can be developed

Table 2: Dataset_1 Emotions Data Distribution.

Emotions	Size
Fear	2252
Anger	1701
Joy	1616
Sadness	1533

SM text including multilingual text, performance metrics, noisy data etc. Alternative to supervised approach unsupervised approach is used for emotions classes at the sentence level.[33] Metaheuristic DL approach were developed for ED from live social media data based on linguistics.[34] English conversations emotion classification with DL and transformer-based model are studied and analyses performance, also ensemble learning which uses technique of majority voting for performance improvement using SemEval 2019 Task 3 public dataset where ensemble model shows improved performance.[35] Use combined BERT and CNN model. Combined BERT and CNN model represents better results over baseline performance using the SemEval 2019 task3 and ISEAR datasets, achieves an 94.7% and 75.8% accuracy for respective datasets.[36] CNN with Sequence and word embedding model is proposed for ED. Attention mechanism is used to focus on words using CNN. Results shows good precision and accuracy for detection of emotion from text.[37] Various AI techniques such as ML, DL, pre-trained models are compared for ED in text-based dialogue and performance of each model evaluated based on balanced and unbalanced conversational data.[38]

Table 3: Dataset_2 Emotions Data Distribution.

Emotions	Size
Fear	2161
Anger	2434
Joy	6057
Sadness	5247
Love	1463
Surprise	638

Existing study of ED with respect to datasets used, models focused and performance of each model showcased in Table 1 along with its future scope. As per literature various DL models along transformer technique and attention mechanism proven improved results for ED. Freely available online datasets were used for analysis of models.

Various models and architectures are proposed which have their own pros and cons. Major issues are identified are highlighted as a research gaps are mentioned below

1) Robustness of ML and DL models (inconsistency of models with different datasets) highlights need of model improvement

2) Relevant Feature extraction is crucial in model performance but still remain challenging task

3) Identification of multiple emotions in text, sarcasm, irony idioms, metaphors, homonyms in text.

4) Availability of a large annotated balanced dataset is major challenge.

5) Lack of auto tuning techniques for efficient model.

To address some issues, we have proposed various models of ML and DL with different architectures and compared performance along with accuracy and loss.

Table 4: Sample annotated data for ED.

Text	Class
Okay, this is hilarious	Joy
I feel pretty pathetic most of the time	Sadness

2. Methodology

2.1 Datasets

Study presents comparative analysis of ED from SM text using ML and DL techniques. The process starts with data collection. We have used two different datasets freely available online which includes 7102 samples for Dataset_1 with emotions fear, anger, joy, sadness and 18000 samples for Dataset_2 with emotions joy, sadness, anger, fear, love, surprise distribution of emotions for respective datasets is as shown in Fig. 2a and Fig. 2b Both datasets are available on online platforms and are used for analysis. The number of classes as well as its emotion count are represented in the same graphs for both datasets. To confirm the robustness of these models we have run different algorithms on two different datasets. The comparative analysis gives performance of each algorithm with respective tuned hyper-parameters on both datasets. All models are applied on the same datasets. Model performances are compared and visualized to identify better algorithms for ED from SM text.

Table 2 and Table 3 represents a number of samples of respective emotion categories used for ED models for two

Table 5: DL Hyper-parameters.

Tuned Hyper-parameters	
DL Models	Hyper-parameters
Bi-LSTM	lr = [0.0001, 0.0005] lu = [64, 128] dr = [0.3, 0.5]
LSTM	lr = [0.0001, 0.0005] lu = [64, 128] dr = [0.3, 0.5]
CNN	lr = [0.0001, 0.0005] nf = [32, 64] ks = [3, 5] dr = [0.3, 0.5]
CNN LSTM	lr = [0.0001, 0.0005] nf = [32, 64] ks = [3, 5] lu = [64, 128] dr = [0.3, 0.5]
LSTM CNN	lr = [0.0001, 0.0005] lu = [64, 128] nf =[32,64] ks = [3, 5] dr = [0.3, 0.5]
LSTM CNN LSTM	lr = [0.0001, 0.0005] lu = [64, 128] nf = [32, 64] ks = [3, 5] dr = [0.3, 0.5]
CNN LSTM CNN	lr = [0.0001, 0.0005] lu = [64, 128] nf = [32, 64] ks = [3, 5] dr = [0.3, 0.5]

Table 6: ML Result Analysis (accuracy) with respect to Dataset_1 with and without hyper parameter tuning.

	Dataset_1	
ML Algorithms	Without Hyper-parameters in %	With Hyper-parameters in %
Decision Tree	80.15	81.35
RF	82.96	83.25
XG Boost	85.15	84.79
SVM	83.88	83.53
NB	74.38	80.78

different datasets. Dataset_1 with four categories and Dataset_2 with six categories of emotion. Whereas Table 4 represents samples of annotated data used for ED which shows text and respective emotion category.

2.2 Proposed framework

Fig. 3a and Fig. 3b outline the steps used to perform ED from SM text using ML and DL respectively. Labeled datasets are used to perform comparative analysis of ED from SM text collected from online SM platforms. Input text passed for prepossessing steps to remove unwanted noise from data *e.g.* stops words removal, to remove emoticons etc. Prepossessing steps include lowering the text, stop-words removal and tokenization. Once the whole text is converted into lower case it is passed to remove punctuation's, stop words, emoticons etc. A dictionary of contraction words is used to convert contraction words to actual English words. Finally processed text is passed for lemmatization. Lemmatization is used to get meaningful words from prepossessed text. Prepossessed data is converted to numerical form using word-to-vec for ML and word embedding for DL. Vectored data is divided into training 80%and testing with 20%. Training labeled vectored data is applied to different ML and DL models. Models are trained using training data and tuned with different hyper-parameters for respective algorithms and techniques. For Machine Grid search cross validation is used to get tuned hyper parameters. Finally tuned parameters are passed to the ML model. Performance of models is measured with training and testing accuracy. For DL different combinations of hyper-parameters are used to hyper-tune DL models. Finally, hyper tuned parameters are passed for further classification of respective ML and DL models.

2.3 Hyperparameter tuning for ML and DL models

Various hyper parameters are used including criterion gini index, maximum depth as 'None' samples leaf as 1, samples split as 5, random splitter for the DT. Maximum depth as 'None', samples leaf as 1, samples split as 2, estimators as 200 are used in RFC. 'colsample_bytree' as 1.0, learning rate as 0.2, maximum depth as 7, minimum child weight as 1, number of estimators as 200, subsample as 1.0 used in XGBoost algorithm. Regularization parameter (C) as 1, gamma as 0.001 and linear kernel is used in SVM. Alpha value as 1.0, 'fit_prior' as 'False' are the hyper parameters used in NB classifier for Dataset 1. For Dataset 2 various hyper parameters are used which includes criterion as gini index, max depth is 'None', samples leaf as 4, min samples split as 10 for DT. For RF Classifier maximum depth is 'None', samples leaf is 2, samples split is 5, number of estimators as 200 used as hyper parameters. 'colsample_bytree' as 1.0, learning rate as 0.2, maximum depth is 7, minimum child weight is 1, number of estimators are 150, and sub sample 0.8 are the hyper parameters used in XGBoost Algorithm. SVC uses regularization parameter (C) as 10, gamma as 0.1, and Radial function kernel is used for SVM algorithm. Alpha value as 1.0 and 'fit_prior is False' are the parameters used in NB. Table 4 gives a brief about hyper-parameters used to hyper tuned various DL models. Each model executed with various combination of given hyper-parameters to get the better results. Table 5 gives a summary of various hyper-parameters used for ED using DL models. Performance of each model is analyzed with combinations of each hyper-parameter with another. Best parameters are identified with respect to performance metric accuracy and respective loss. learning_rates (lr), lstm_units_options (lu), dropout_rates (dr), num_filters_options (nf), kernel_sizes (ks) are the common parameters used in these models to get enhanced performance of DL models.

3. Results and discussion

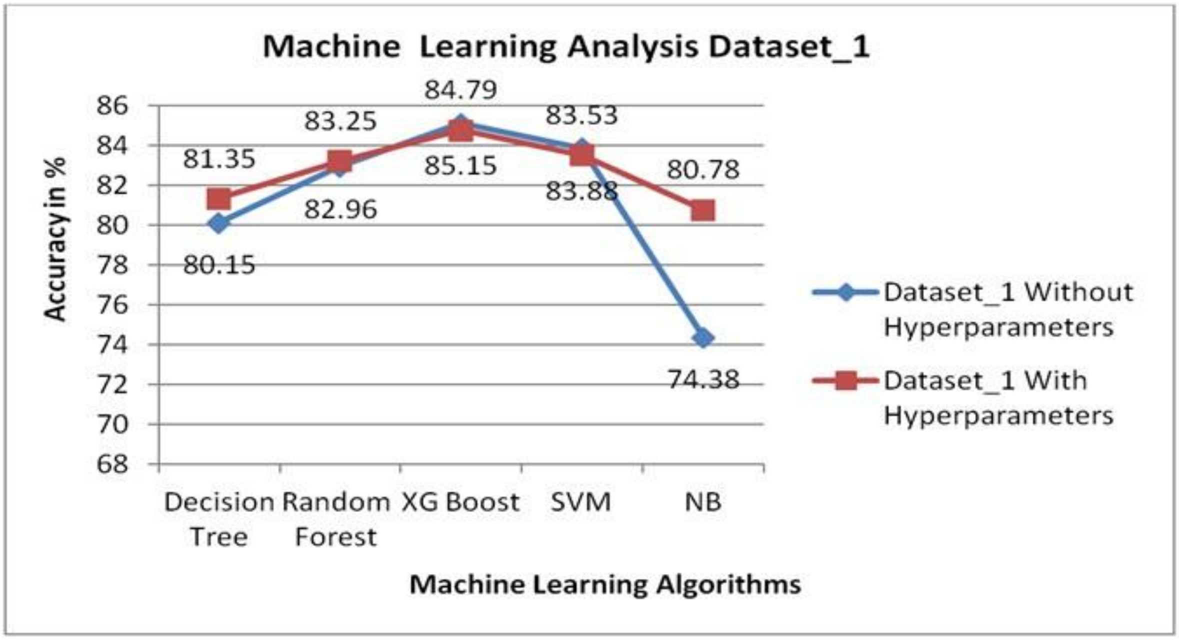
Here, we outline the findings from our analysis of different ML and DL models for ED from SM text. Various model performances are evaluated with accuracy and loss of the trained models with and without hyper parameters. Table 6 numerically analyze performance of various ML algorithms on Dataset_1 and Table 7 analyze performance of various ML algorithms such as DT, RF, XGB, SVM and NB on Dataset_2 with and without hyper-parameters. Hyperparameters used to evaluate performances represented in Table 5. XGBoost model perform well for Dataset_1 as well as Dataset_2 with and without hyper-parameters.

Performance metric accuracy of different ML models used for analysis on Dataset_1 with and without hyper parameters shown in Fig. 4a and Fig. 4b represents accuracy of different ML models used for analysis on Dataset_2 with and without hyper parameters. Hyper parameters tuning represents better results as per simple ML models. In both datasets Extreme Gradient Boosting (XGBoost) algorithm shows better results 84.79% in Dataset_1 and 87.68% in Dataset 2 followed by RF classifier and SVM algorithm.

Table 7: ML Result Analysis (accuracy) with respect to Dataset_2 with and without hyper parameter tuning.

Dataset_2		
Column Header Goes Here	Without Hyperparameters in %	With Hyperparameters in %
Decision Tree	84.28	84.56
RF	86.75	86.84
XG Boost	87.40	87.68
SVM	81.15	86.31
NB	68.78	80

(a)



(b)

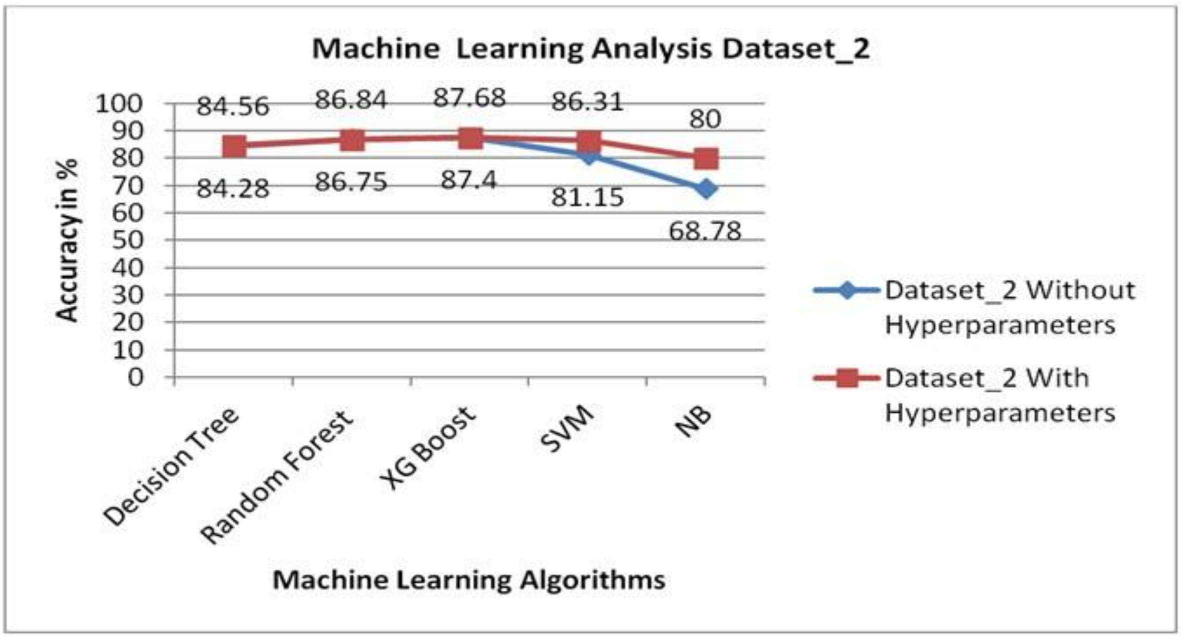


Fig. 4: ML Algorithms Performance Analysis: (a) Dataset_1; (b) Dataset_2.

Fig. 6 represents graphical representation of number of epochs verses accuracy and number of epochs verses loss for different DL models used for analysis on Dataset_2. In both datasets models used to evaluate performances includes Bi-LSTM, LSTM, CNN, CNN-LSTM, LSTM-CNN, CNN- LSTM-CNN and LSTM-CNN-LSTM.

Table 8 represents numerical analysis of different DL algorithms on two different datasets with best hyper-parameters tuning with help of Dataset_1 and Dataset_2 for each different model of DL. Hybrid models of CNN and LSTM proves better results over traditional DL models.

Fig. 7 represents a graphical view of performance of various DL models with hyper tuned parameters for respective datasets. Each model is evaluated with performance metric accuracy and loss of model. DL models represent data driven results as Dataset_2 represents better accuracy as compared to Dataset_1.

As per graphical results presented in Fig. 5 and Fig. 6, we can conclude that some of the models are overfitted because of limited and unbalanced dataset. Each model is executed with all various combinations of mentioned hyperparameters presented in Table 4 Out of all possible runs, the best runs are considered as the final result. Hybrid DL models, CNN-LSTM and CNN-LSTM-CNN model achieved the highest accuracy across both datasets. Hybrid model CNN-LSTM performing best on Dataset_2 gives accuracy 92.40%. The Bi-LSTM and CNN models also performed well, especially in Dataset_1, where Bi-LSTM reached accuracy of 87.04%. Proposed models compared with existing baseline models. Proposed models compared with existing baseline models. Table 9 represents comparison of analysis of accuracy with respect to various ML and DL techniques. Various datasets are

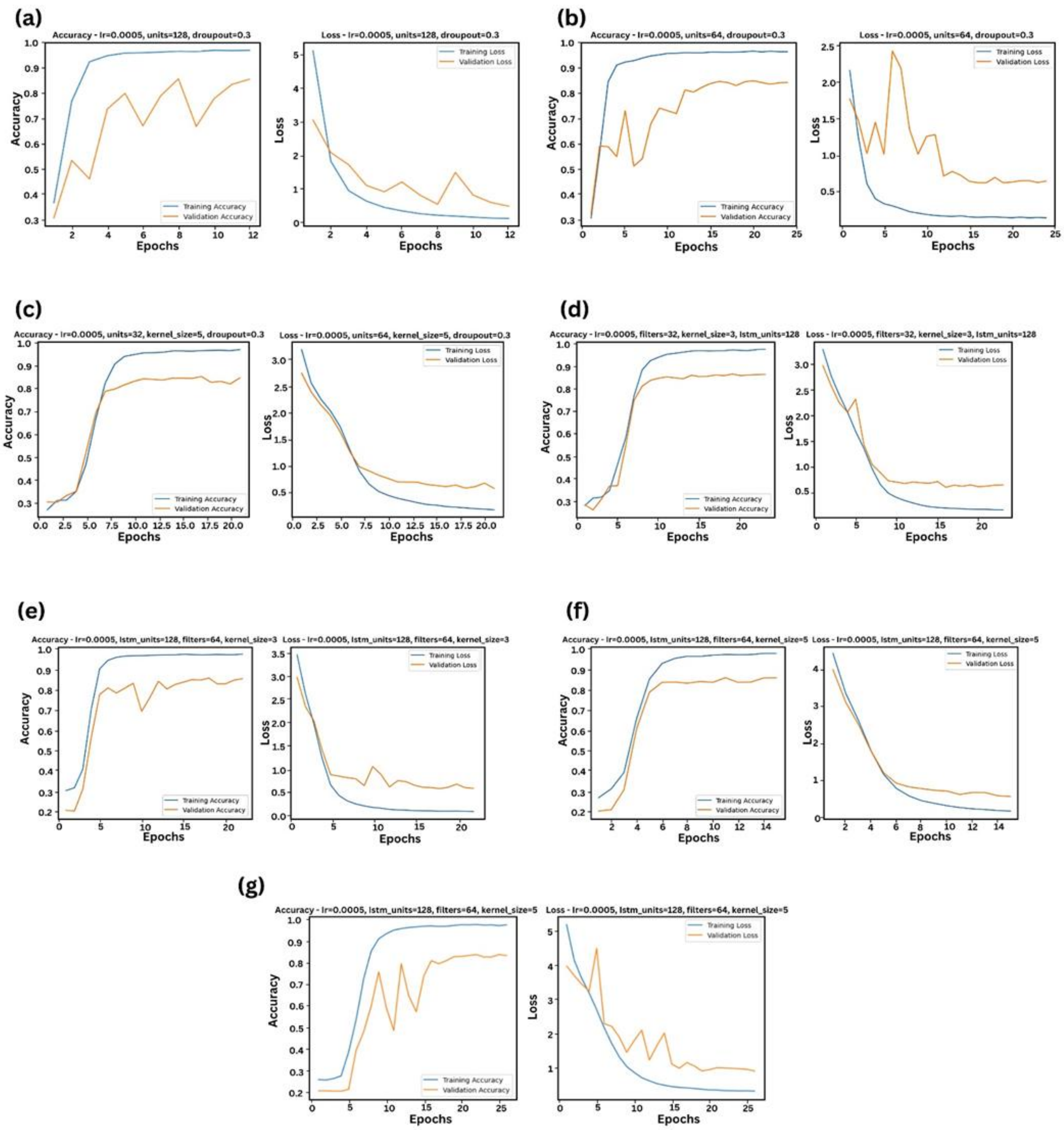


Fig. 5: DL Techniques Performance Analysis Dataset_1: (a) Bi-LSTM; (b) LSTM; (c) CNN; (d) CNN LSTM; (e) LSTM CNN; (f) CNN LSTM CNN; (g) LSTM CNN LSTM.

Table 8: DL Result Analysis (accuracy) with respect to Dataset_1 and Dataset_2 with tuned hyper-parameter.

Models / Dataset	Dataset_1 Accuracy in %	Dataset_2 Accuracy in %
Bi-LSTM	87.04	86.50
LSTM	83.52	86.60
CNN	85.07	91.90
CNN LSTM	85.49	92.40
LSTM CNN	85.49	90.30
LSTM CNN LSTM	83.52	90.40
CNN LSTM CNN	87.32	91.50

Table 9: Comparison of state-of-art methods with our results.

References	Dataset	Highest Performance in Accuracy (%)	
		ML	DL
[2]	(Dataset 2) Kaggle, 20000 sentences	SGD: 89.79%,	DL: 89.76%
[3]	(Dataset 2) Kaggle, 20000 sentences	LR: 84%	BI-LSTM and BERT: 94%,
[14]	(Dataset 2) Kaggle, 20000 sentences	NA	ANN: 92%.
[16]	(Dataset 2) Kaggle, 20000 sentences	NA	Hybrid model NN, CNN and RN: 90%,
[17]	ISEAR (Dataset 1)	SVM: 64%.	M7B: 76%
[26]	ISEAR (Dataset 1)	NA	GRU 60.26%,
[34]	ISEAR, WASSA, Emotion-stimulus	Proposed LED_LSMD_MDLT 85% for ISEAR, 86% for WASSA, 91% for Emotion-Stimulus	NA
[36]	SemEval 2019 Task3 and ISEAR	NA	BERT-CNN 94.7% for semeval2019 task3 dataset, 75.8% for ISEAR dataset.
[39]	ISEAR, WASSA, Emotion-Stimulus dataset	SVM 78.97	Bi-GRU 79.46, Hybrid model 80.11%.
Proposed Models	(Dataset 2) Kaggle, 20000 sentences	XGB: 87.68	CNN LSTM: 92.40
Proposed Models	ISEAR (Dataset 1)	XGB: 85.15	CNN LSTM CNN: 87.32

applied on ML and DL models. In state-of-arts ML models, SGD proven highest accuracy of 89.79% for online Kaggle dataset of 20000 samples whereas BERT-CNN DL model proven highest accuracy of 94.7% for semeval2019 task3 dataset also BI-LSTM and BERT gives better accuracy of 94% for online Kaggle dataset of 20000 samples. In this work, we have proposed models which are executed and compared with various hyperparameters. In ML we have used various combinations of hyperparameters. Each parameter is executed and evaluated with combinations of other parameters. In DL we have proposed various architectures of DL techniques along with hybrid models of CNN and LSTM. Each model is executed with different epochs, layers, units, filters etc. Best fit model showcased in a given study. Hyperparameter tuning shows slightly improved accuracy for ML algorithms whereas NB and SVM showing better accuracy gain. XGB gives better accuracy for both datasets which shows robustness of model. In DL hybrid models highlights the benefits of both convolutional and recurrent layers. Sequential model (Bi-LSTM) performed well for ISEAR dataset whereas CNN based models performed well on dataset 2. Pretrained model shows highest accuracy of 94% whereas proposed simple hybrid model of LSTM and CNN showcase 92% without increasing complexity of the model. XGB algorithm gives improved performance over LR and SVM model. Size of datasets affects performance of models. Still there are chances to improve robustness of models by training models on large datasets.

4. Challenges

There are several challenges to detect emotion from SM text. Some of the listed bellow

- As SM is a casual way of communication, many users use their regional language. Also it is associated with code mixed (CM) or Code Switched (CS) text thus multi- language text and ED from such CM or CD text is one of the major challenging tasks.
- Extraction of contextual information is still lagging due to various issues like, sarcasm, idioms etc.
- Extraction of linguistic features plays a major role in extraction of context of text. Due to different linguistic structures, idioms, phrases, etc. Extraction of linguistic context is a challenging task.
- Availability of annotated datasets is measure concern in the DL.

SM is informal way of text representation where users post text with shortcuts, non-dictionary words, emojis, incorrect

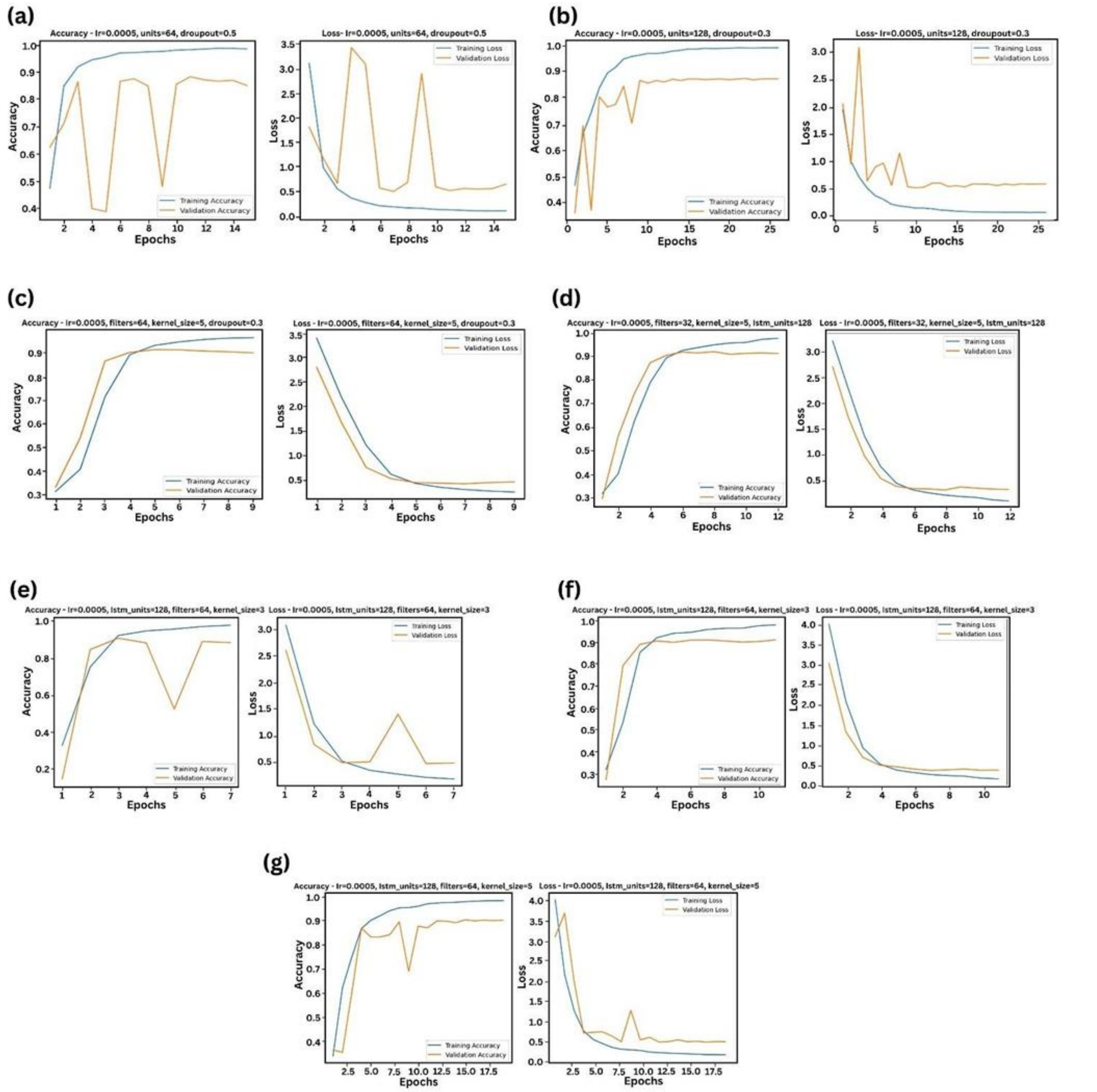


Fig. 6: DL Techniques Performance Analysis Dataset_2: (a) Bi-LSTM; (b) LSTM; (c) CNN; (d) CNN LSTM; (e) LSTM CNN; (f) CNN LSTM CNN; (g) LSTM CNN LSTM.

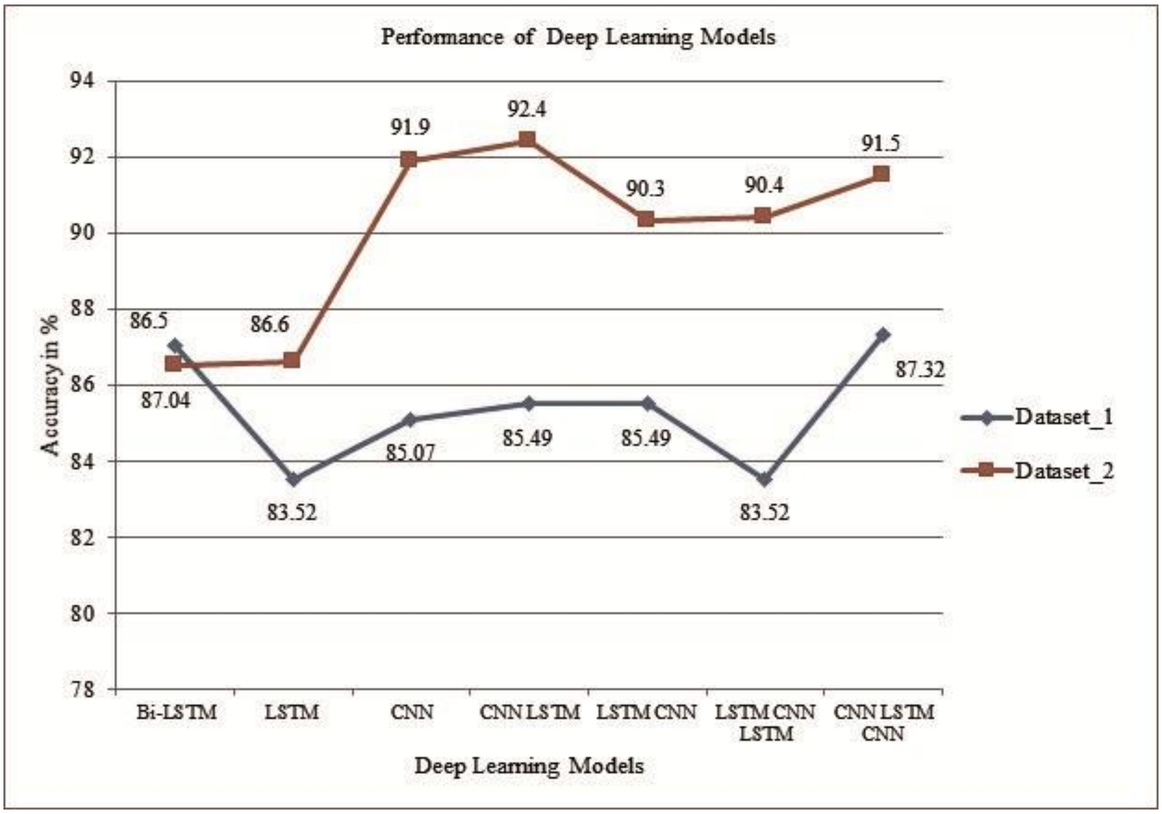


Fig. 7: Performance of DL Models on various datasets.

grammar, short forms etc. due to such issues extraction of implicit emotions form such huge amount of SM text is challenging task.

5. Conclusion

Implicit ED from social media text gives insights in different domains like medical, e-commerce, defense etc. This research focuses on comparative numerical analysis of different ML algorithms used for implicit ED from social media text. Online datasets used for analysis consist of labeled text with respect to various categories of emotions such as fear, anger, joy sadness etc. Result analysis shows that Extreme Gradient Boosting (XGBoost) algorithm shows better results with accuracy of 84.79% and 87.68% in two different datasets followed by RF Classifier and SVM algorithm. DL results indicate that hybrid CNN and LSTM models outperforms well, especially CNN-LSTM models. DL approach highlights the combined benefit of convolutional and recurrent layers for extracting both spatial and temporal features. Extraction of implicit information from data available on online platforms (social media) is a difficult task due to various challenges. In current study development of a reliable ED model is still lagging. Performance of the model can be improved by minimizing overfitting with the help of balanced datasets along with different performance metrics, hyper-parameter tuning, hybrid approaches, ET, and advanced DL techniques like attention mechanism or transformer techniques. Also, Real-time ED from mixed language text is the emerging area of research in the field of SA.

Acknowledgements

We would like to acknowledge the online platforms such as Kaggle, GitHub for availability of online datasets.

Conflict of Interest

There are no conflicts to declare.

Supporting Information

Not applicable.

CRediT Statement

Vinayak Malavade: contributed to conceptualization, implementation, formal analysis, and writing the original draft. **Virat Giri:** guided, supported and supervised the work, and assisted in reviewing and editing. **Shruti Patil:** supported the research for data curation, analysis, and validation. **Deepali Jadhav:** supported for validation and updating draft of the paper. **Sheetal Kusal:** supported for validation, visualization and assisted in writing the draft. **Jyoti Deshmukh:** provided guidance in formal analysis and assisted in reviewing and editing the manuscript.

References

[1] S. Kusal, S. Patil, K. Kotecha, R. Aluvalu, V. Varadarajan, AI based emotion detection for textual big data: techniques and contribution, *Big Data and Cognitive Computing*, 2021, **5**, 43, doi: 10.3390/bdcc5030043.

[2] H. Pujara, P. Babel, Emotable - emotion detection based social media application using machine learning and deep learning, *2022 IEEE Conference on Interdisciplinary Approaches in*

Technology and Management for Social Innovation (IATMSI), December 21-23, 2022, Gwalior, India. IEEE, 2022, 1-6, doi: 10.1109/IATMSI56455.2022.10119332.

[3] D. Shukla, S. K. Dwivedi, A comparative study of text-based emotion detection techniques for emotion recognition on social media data, *2023 IEEE 7th Conference on Information and Communication Technology (CICT)*, December 15-17, 2023, Jabalpur, India. IEEE, 2023, 1-6, doi: 10.1109/CICT59886.2023.10455381.

[4], S. Thimmaiah, R. Jayaram, A new gate control unit-recurrent neural network structure for audio-based sentiment analysis, *Engineered Science*, 2024, **30**, doi: 10.30919/es1180

[5], D. G. Takale, A. V. Dhumane, P. N. Mahalle, T. Jadhav, P. P. Gawali, A. Buchade, Optimize deep learning model for intensive care of neurological disorders patients based on facial expression, *Engineered Science*, 2024, **32**, doi: 10.30919/es1364

[6], R. D. Jathanna, D. Acharya, L. E. Lewis, K. Makkithaya, Early detection of late onset neonatal sepsis using machine learning algorithms, *Engineered Science*, 2023, **26**, doi: 10.30919/es976

[7], D. R. Kumar, W. Wipulanusat, J. Sunkpho, S. Keawsawasvong, W. Jitchaijaroen, P. Samui, Machine learning approaches for the prediction of the seismic stability of unsupported rectangular excavation, *Engineered Science*, 2024, **28**, doi: 10.30919/es1083

[8], A. Diveev, E. Sofronova, N. Konyrbaev, S. Ibadulla, Stabilisation system synthesis for motion along the trajectory by evolutionary machine learning control, *Engineered Science*, 2024, **29**, doi: 10.30919/es1130

[9], M. Zain, S. Keawsawasvong, C. Thongchom, I. Sereewatthanawut, M. Usman, L. Prasittisopin, Establishing efficacy of machine learning techniques for vulnerability information of tubular buildings, *Engineered Science*, 2023, **27**, doi: 10.30919/es1008

[10], A. M. Al-Hinawi, R. A. Alelaimat, E. Alhenawi, M. I. AlBiajawi, Hybrid deep learning approach for accurate prediction of flowability in ultra-high-performance concrete, *Engineered Science*, 2024, **30**, doi: 10.30919/es1182

[11] P. Ekman, Basic Emotions, *Handbook of Cognition and Emotion*, 1999, 45–60, doi: 10.1002/0470013494.CH3.

[12] R. Jan, A. A. Khan, Emotion mining using semantic similarity, *International Journal of Synthetic Emotions*, 2018, **9**, 1-22, doi: 10.4018/ijse.2018070101.

[13] S. Kusal, S. Patil, J. Choudrie, K. Kotecha, D. Vora, I. Pappas, A systematic review of applications of natural language processing and future challenges with special emphasis in text-based emotion detection, *Artificial Intelligence Review*, 2023, **56**, 15129-15215, doi: 10.1007/s10462-023-10509-0.

[14] K. Rai, P. Kumar, Comparative analysis of machine learning and deep learning techniques in text based emotion detection, *2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)*, April 28-30, 2023, Greater Noida, India. IEEE, 2023, 260-263, doi: 10.1109/CISES58720.2023.10183612.

[15] Y. A. Jasim, M. G. Saeed, M. B. Raewf, Analyzing social media sentiment: twitter as a case study, *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 2023, **11**, 427-450, doi: 10.14201/adcaij.28394.

[16] K. Machová, M. Szabóova, J. Paralič, J. Mičko, Detection of emotion by text analysis using machine learning, *Frontiers in Psychology*, 2023, **14**, 1190326, doi: 10.3389/fpsyg.2023.1190326.

[17] S. H. N. Esfahani, M. Adda, Classical machine learning and large models for text-based emotion recognition, *Procedia Computer Science*, 2024, **241**, 77-84, doi: 10.1016/j.procs.2024.08.013.

[18] T. Madhu Midhan, P. Selvaraj, M. Harshavardan Kumar Raju., M. Bhanu Prakash Reddy., T. Bhaskar, Classification of mental health and emotion of human from text using machine learning approaches, *6th International Conference on Information Systems and Computer Networks (ISCON),* March 3-4, 2023, Mathura, India. IEEE, 2023, 1-7, doi: 10.1109/ISCON57294.2023.10111973.

[19] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: pre-training of deep bidirectional transformers for language understanding, North American Chapter of the Association for Computational Linguistics, 2021

[20] N. Alswaidan, M. El Bachir Menai, A survey of state-of-the-art approaches for emotion recognition in text, *Knowledge and Information Systems*, 2020, **62**, 2937-2987, doi: 10.1007/s10115-020-01449-0.

[21] S. Arun Kumar S. and A. Geetha, Emotion Detection from Text using Natural Language Processing and Neural Networks, *International Journal of Intelligent Systems and Applications in Engineering*, 2024, **12**, 609–615, https://www.ijisae.org/index.php/IJISAE/article/view/4707

[22] F. A. Acheampong, W. Chen, H. Nunoo-Mensah, Text-based emotion detection: advances, challenges, and opportunities, *Engineering Reports*, 2020, **2**, e12189, doi: 10.1002/eng2.12189.

[23] J. Guo, Deep learning approach to text analysis for human emotion detection from big data, *Journal of Intelligent Systems*, 2022, **31**, 113-126, doi: 10.1515/jisys-2022-0001.

[24] P. Nandwani, R. Verma, A review on sentiment analysis and emotion detection from text, *Social Network Analysis and Mining*, 2021, **11**, 81, doi: 10.1007/s13278-021-00776-6.

[25] J. Choudrie, S. Patil, K. Kotecha, N. Matta, I. Pappas, Applying and understanding an advanced, novel deep learning approach: a covid 19, text based, emotions analysis study, *Information Systems Frontiers*, 2021, **23**, 1431-1465, doi: 10.1007/s10796-021-10152-6.

[26] D. Yohanes, J. S. Putra, K. Filbert, K. M. Suryaningrum, H. A. Saputri, Emotion detection in textual data using deep learning, *Procedia Computer Science*, 2023, **227**, 464-473, doi: 10.1016/j.procs.2023.10.547.

[27] A. Chatterjee, U. Gupta, M. K. Chinnakotla, R. Srikanth, M. Galley, P. Agrawal, Understanding emotions in text using deep learning and big data, *Computers in Human Behavior*, 2019, **93**, 309-317, doi: 10.1016/j.chb.2018.12.029.

[28] X. Zhu, Y. Lou, H. Deng, D. Ji, Leveraging bilingual-view parallel translation for code-switched emotion detection with

adversarial dual-channel encoder, *Knowledge-Based Systems*, 2022, **235**, 107436, doi: 10.1016/j.knosys.2021.107436.

[29] T. T. Sasidhar, B. Premjith, P. Soman K, Emotion detection in hinglish(Hindi+English) code-mixed social media text, *Procedia Computer Science*, 2020, **171**, 1346-1352, doi: 10.1016/j.procs.2020.04.144.

[30] J. Herzig, M. Shmueli-Scheuer, D. Konopnicki, Emotion Detection from Text via Ensemble Classification Using Word Embeddings, *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval,* Amsterdam, The Netherlands. ACM, 2017, 269–272, doi: 10.1145/3121050.3121093.

[31] M. Polignano, P. Basile, M. de Gemmis, G. Semeraro, A Comparison of Word-Embeddings in Emotion Detection from Text using BiLSTM, CNN and Self-Attention, *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization,* Larnaca Cyprus, ACM, 2019, 63–68, doi: 10.1145/3314183.3324983.

[32] M. Haggag, S. Fathy, N. Elhaggar, Ontology-based textual emotion detection, *International Journal of Advanced Computer Science and Applications*, 2015, **6**(9), doi: 10.14569/ijacsa.2015.060932

[33] A. Agrawal, A. An, Unsupervised emotion detection from text using semantic and syntactic relations, *IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology,* December 4-7, 2012, Macau, China. IEEE, 2012, 346-353, doi: 10.1109/WI-IAT.2012.170.

[34] S. Mubeen, D. N. Kulkarni, M. R. Tanpoco, D. R. D. Kumar, M. Lakshmu Naidu, T. Dhope, Linguistic based emotion detection from live social media data classification using metaheuristic deep learning techniques, *International Journal of Communication Networks and Information Security* (*IJCNIS*), 2022, **14**, 176-186, doi: 10.17762/ijcnis.v14i3.5604.

[35] A. Thiab, L. Alawneh, M. AL-Smadi, Contextual emotion detection using ensemble deep learning, *Computer Speech & Language*, 2024, **86**, 101604, doi: 10.1016/j.csl.2023.101604.

[36] A. R. Abas, I. Elhenawy, M. Zidan, M. Othman, BERT-CNN: a deep learning model for detecting emotions from text, *Computers, Materials & Continua*, 2022, **71**, 2943-2961, doi: 10.32604/cmc.2022.021671.

[37] K. Shrivastava, S. Kumar, D. K. Jain, An effective approach for emotion detection in multimedia text data using sequence based convolutional neural network, *Multimedia Tools and Applications*, 2019, **78**, 29607-29639, doi: 10.1007/s11042-019-07813-9.

[38] S. D. Kusal, S. G. Patil, J. Choudrie, K. V. Kotecha, Understanding the performance of AI algorithms in text-based emotion detection for conversational agents, *ACM Transactions on Asian and Low-Resource Language Information Processing*, 2024, **23**, 1-26, doi: 10.1145/3643133.

[39] S. K. Bharti, S. Varadhaganapathy, R. K. Gupta, P. K. Shukla, M. Bouye, S. K. Hingaa, A. Mahmoud, Text-based emotion recognition using deep learning approach, *Computational Intelligence and Neuroscience*, 2022, **2022**, 2645381, doi: 10.1155/2022/2645381.