Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

SICAIET
2024

Review Article
Collection: Artificial Intelligence and Emerging Technologies

Navigating Sarcasm in Multilingual Text: An In-Depth Exploration and Evaluation

Hamid Ghous[1], Mubasher H. Malik[2*], Javeria Altaf[3], Sundas Nayab[4], Iqra Sehrish[4], and Syed Ali Nawaz[5]

[1]Department of Computer Science, Institute of Southern Punjab, Multan, Pakistan.
[2]Vision, Linguistics & Machine Intelligence Research Lab, Multan, Pakistan.
[3]Department of Computer Science, Institute of Southern Punjab, Multan, Pakistan.
[4]Department of Information Technology, Institute of Southern Punjab, Multan, Pakistan.
[5]Department of Information Technology, Islamia University of Bahawalpur, Bahawalpur Punjab, Pakistan.
[*]Corresponding Author: Mubasher H. Malik. Email: mubasher@isp.edu.pk

Academic Editor: Salman Qadri **Published:** February 01, 2024

Abstract: Sarcasm means using words when you say something opposite from what you want to say, either to irritate someone, offend them, or just for fun. Detecting sarcasm in multiple languages is yet a challenging area of research. Identifying and understanding sarcasm in social media indicates people's thoughts about specific topics, news, and products. Many articles have been published on sarcasm detection using deep learning and machine learning methods. Moreover, very few systematic reviews have been conducted in this research area. This paper systematically reviews existing Artificial intelligence (A.I.) techniques in sarcastic text detection in different languages. The studies emphasized that in recent literature, machine learning and deep learning, especially recurrent neural networks (RNN), are the most commonly used techniques for sarcasm detection. Twitter is the most frequently used source and accuracy for performance measures. The articles covered in the literature survey also include sarcasm detection from social text, books, and code-mixed text, among other datasets. Finally, this paper briefly discusses the challenges in sarcasm detection and future research in this area.

Keywords: Sarcasm; Text Processing ‘News; Machine Learning; Deep Learning.

1. Introduction

Sarcasm detection in text has become an essential research topic in natural language processing NLP. Sarcasm is a figure of speech in which the intended meaning does not correspond to the text's literal sense. Interestingly, recognizing sarcasm is essential not only for such downstream tasks as sentiment analysis and dialogue systems, but sarcasm detection is complex because sarcasm very much relies on context, background information, and cultural details. This paper aims to study all the related articles to identify sarcasm in text. As much less work is done in multilingual, we learn all the associated articles, whether using mono-lingual, Bi-lingual, or code-mixed languages. We will explore what methods are used, the dataset gathered from which sources, what they aim to solve, how they test their results, which models perform best, and which models perform average to detect sarcasm in text. Throughout the review, we will study and identify which problems still occur in sarcasm detection, and we will also recommend suggestions for the following main task in this research area.

Comprehension of subtle linguistic forms is one of the significant challenges in the constantly changing NLP scenario. Verbal irony, especially sarcasm used for comical purposes or satirical comments, makes this even more complicated. Detecting and understanding sarcasm in the text is vital for improving accuracy. The topic of this research paper is sarcasm text identification, which involves a complex task (Aniruddha Ghosh, 2016).

There are several recent studies on sarcasm detection in social media texts because such texts do not contain the nonverbal cues that facilitate human understanding through the human voice and face. A variety of machine learning methods, such as support vector machines and convolutional neural networks (CNN), have been used to learn representative text patterns that identify sarcasm (Aniruddha Ghosh, 2016) (Tay†, 2018). As such, one of the significant limitations is that most studies focus on issues in English texts where sarcasm takes a different shape across cultures and languages. So, there is a need to tackle sarcasm in other regional languages. Specifically, (ADITYA JOSHI 2016) highlighted the importance of multilingual sarcasm detection that considers linguistic differences.

In this paper, A.I. techniques for multilingual sarcastic text detection are reviewed. We present some recent computational methods for the modeling of sarcasm. As much research is done in English, there is a need to detect sarcasm in more languages such as Arabic, Urdu, French, Czech, etc. These key factors concern contextual knowledge, showing incongruity between positive and negative sentiments, and K2L transfer (Aditya Joshi, 2015) (Hee, 2017). We also discuss the limitations of current approaches and give directions for further multilingual studies. Progress in multilingual sarcasm recognition will help sentiment analysis and other NLP applications in today's global society. The contribution of this study is a complete analysis of A.I. methods for recognizing sarcasm in text among different languages. Below, we outline the most critical approaches that have moved state-of-the-art sarcasm detection for languages such as English, Arabic, Czech, Dutch, and Italian (Francesco Barbieri, 2014) (Jihen Karoui, 2017). We emphasize critical issues in transferring computational methods across languages, including morphological variety and the necessity of using external knowledge sources. We also touch upon the limitations of current methods and provide suggestions for further multilingual work. There is a need for much more effort for multilingual sarcasm detection; it will assist with downstream activities like sentiment analysis, intent detection, and social media analytics for the ever-expanding international online world.

2. Literature Review

We intend to explore the field of multilingual sarcastic text identification, which reflects a dynamic and evolving field. Researchers are actively exploring novel techniques and integrating deep learning models, and there is a need to address challenges related to multilingual sarcasm from different regions. There is a need to review the literature in the following sections.

2.1. Multilingual Sarcastic Identification Using Social Media Reviews

The (lsa Scola, 2021) performed some pre-processing techniques, i.e., word embedding and Removing the bag of words on the dataset of 28,503 headlines. They used BiLSTM, BERT, SVM, and CNN methods, and accuracy achieved 87%, 91%, 79%, and 86%, respectively. (Deepali D.Londhe 2021) Based on a review for sentiment analysis. Language identification, classification, and normalization were performed on the Utsab. Barman et al. used the FIRE 2013 and FIRE 2014 datasets. They reviewed comparisons between different methods: Ling Pipe classifier and Textcat toolkit, Google Translate and Microsoft AP, Lexicon-based approach, BOG technique, and SVM. Results need more effort for multilingual sentiment analysis and mixed script sentiment analysis. (A. Palaniammal 2023) I performed pre-processing methods like classification, noise removal, and feature extraction on the Headline 2019 dataset. The paper used the ARO-MCEDNN Glove approach for word embedding methods and achieved 98.70% accuracy. (Ashwani Kumar 2023) I performed data cleaning, noise removal, normalization, handling missing values, and feature extraction on the dataset of information from tweets, discussions, and comments in English, Hindi, French, and German. They worked on different approaches: D.M. techniques, Text mining, NLP, A.I., and ML techniques (Vader et al. Machine)—the highest accuracy was achieved by two methods, 89% SVM and 83% Naïve byes.

Shekhar (Shekhar, 2021) performed tokenization and word embedding on the dataset of 9,752 words (code-mix text) in Hindi and English languages. They used methods like the Rule-framed approach, HLSTM learning model, Neural network and simulation model, and Voting technique—83% highest accuracy achieved. The (Poornachandran, 2022) proposed Word2Vec, FastText, XLM-R, and 1D-CNN and performed some pre-processing like Tokenization, Segmentation Noise removal, and transliteration on the dataset retrieved from FIRE 2020 & EACL 2021 in Malayalam & English languages. 77% accuracy achieved. (Matic Rašl, 2021) performed pre-processing techniques, i.e., Feature extraction and tokenization of the Twitter-based English dataset and Slovene dataset in English and Slovene languages. They proposed RoBERTa, Distil-Bert, and DistilBert-multilingual methods, and Results showed a 0.72% F1-score for English and 0.88% F1-score for Slovene. The Thara and (Thara, 2021) introduced a BERT, CamemBERT, DistilBERT, transformer models, and the dataset of 50K sentences extracted from YouTube (code-mixed) in English and Hindi. Basic pre-processing techniques, i.e., Feature extraction and tokenization, were performed on the dataset. Results showed 74% accuracy.

(Liana Ermakova, 2023) They performed Feature extraction on the datasets of the JOKER track at CLEF 2023, SemEval—2017, and SemEval-2021. They Proposed Ridge, NB, FastText, MLP, T5, and random forest methods and achieved 75% accuracy. The (Mohammad et al., 2023) introduced scarcity in low-resource languages. The dataset from SCIDN and MACI consists of 1.5M and 665k, respectively, in English, French, Spanish, and Italian. Data cleaning, selection, machine translation, and transformation were performed on the datasets. The proposed methodology used SVM, L.R., Random Forest, MLP, LSTM, XLMR, and mBERT methods for comparison and achieved 95% accuracy. (Amal Alqahtani, 2023) They proposed a review of sarcasm detection of the dataset between 2019 and 2022. It reviewed some pre-processing techniques like Feature extraction and tokenization as different methods reviewed like lexicons, traditional machine learning, deep learning, transformers Naive Bayes, KNN, RIPPER, C4.5 Decision Tree, BERT, LSTM, SVM, and the highest accuracy achieved 99.1%. (Aqsa Younas, 2020) introduced Deep learning approaches, mBERT, and XLM-R techniques on the MultiSenti Code-mixed dataset. They performed Categorization, Removal of a bag of words, and Labeling on the dataset. 0.65% F1 score obtained. (Elena Zotovaa 2021) Performed labeling, Feature extraction, And validation techniques on the - CIC Corpus - SemEval 2016 dataset and used SVM, XLNET, RoBERTa, mBERT, and XLM-RoBERTa methods and obtained a 75.10 F1 score. The (K Maity 2022) performed some pre-processing techniques, i.e., Removed missing values and irrelevant data and converted images into text data on the 25000 images or memes dataset. They used Proposed CLIP-CentralNet, ResNet, and mBERT. (Manjot Bedi 2021) based on a review for sentiment analysis. Alignment, Noise removal, and Feature extraction performed on the Utsab BMaSaC (code mixed) consist of 36,000 Hindi and 3,000 Eng. words for their analysis. They proposed Proposed MSH-COMICS, a neural architecture for sarcasm detection. Results achieved 83% for sarcasm and 87% for Humor classification.

(Ritesh Kumar, 2021) worked with SVM, BERT, ALBERT, and DistilBERT classifiers on the dataset HASOC (8000 posts from Twitter and Facebook) and TRAC-2 (aggressive language). They evaluated the loss that occurred to detect sarcasm. They performed Noise removal on the dataset. Accuracy achieved 80%. (ALAA RAHMA 2023) Performed pre-processing techniques like Data cleaning, Noise removal, Feature extraction, and Feature space on the review dataset between 2017 and 2022 on Arabic sarcasm detection. The paper used SVM, L.R., NB, LRCV, and deep learning models. (Sanzana et al., 2023) Tokenization and noise removal were performed on the dataset of 25,636 sarcastic comments from Facebook. They worked on different Transformer-based generative adversarial learning, SS-GAN, and GAN-BERT architecture—97.2% highest accuracy achieved. (Sayani Ghosal 2023) performed Noise removal, Data cleaning, Stop words, Empty/missing values removal, and Data analysis on the dataset of 10,000(hate speech) and 20,000(non-hate speech) from social media posts and news articles. They used methods like FSVMCIL, mBERT, the Morphological analysis method, the Hate similarity (H.S.) scheme, and the Word2Vec word embedding model. 85% accuracy achieved. The (arun kumar yadav 2023) worked on the Machine Learning and deep learning methods (CNN-BiLSTM) methodology on a consolidated dataset of 20600 instances. Noise removal, White space, and missing values were removed, and Data cleaning was performed on the dataset. It achieved 87% accuracy. The (A Ameur 2023) proposed self-training technique, AraBERT, Machine learning models(SVM and Logistic Regression) methods, performed some pre-processing like Noise removal, tokenization, and Fine tuning on the dataset retrieved from SemEval-2016 for hotel reviews, and 91% accuracy was achieved.

(Bharathi et al. 2023) performed pre-processing techniques, i.e., Normalization, Removing noise, Feature extraction, and Data analysis on the dataset of code-mixed (Tamil-English) 12,795 comments extracted from YouTube. They proposed BACD, FGACD, D.T. technique, TFIDF, and Bag of Words (BoW) with classical machine learning models, deep learning-based models, and N.B. classifier. Results showed 84% for code mixed. Moreover, this paper used pre-processing techniques, i.e., Segmentation, MFCC, and tokenization on the Mustard dataset, 690 videos, by using the different Audio segment models, Text models, and hybrid models (audio, text). The (SK Bharti, 2022) achieved 70.35% F1-score.

(CHRISTOPHER IFEANYI EKE 2021) performed data cleaning, noise removal, tokenization, and stemming on the datasets of published articles from 2008 to 2019 with content-based linguistic features.

They reviewed different Content-based feature extraction techniques, Bag of Words, word2vec, and n-gram methods. (Shivani Kumar, 2022) performed pre-processing techniques Classification, Noise removal, Vectorization, Cosine similarity and proposed Multi-modal Aware Fusion module, MCA2, Global Information Fusion (GIF), BARTmBART on the dataset of WITS dataset contains 2240 sarcastic text. 91% accuracy achieved. (Eke 2021) I contained the Malayalam–English code-mixed dataset and performed feature extraction and tokenization. They proposed XLM-R, Word2Vec, FastText, BiGRU, RNN, CNN, and LSTM. Results showed 0.76% F1-score. (Ramchandra Joshi 2021) [34] Performed Data collection, Aggregation, and Sampling techniques on the Constraint@AAAI 2021, which contains 8192 online posts collected from social media and used CNN, Multi-CNNBi-LSTM, CNNBi-LSTM, mBERT, IndicBERT methods and obtained 79.93%.

(Ibtissam Touahri, 2021) They reviewed the Tweets dataset of 5,030 Arabic political reviews. Some pre-processing techniques were performed, like data cleaning, noise removal, and feature extraction on the dataset. They used methods of LSTM architecture, SVM, Random Forest, RNN, and LSTM with a word2vec model and achieved 89.24% accuracy. (Christopher Ifeanyi Eke 2020) Performed different pre-processing techniques, i.e., Data acquisition, Feature extraction, Training, and Classification on the dataset of Internet Argument Corpus version 2 (IAC-v2) consisting of 3260 posts per class for sarcasm detection. They Proposed the BERT model, a deep learning model with Bi-LSTM, and obtained 89.24 accuracies. Table 1 showed the tabular review of Multilingual Sarcastic Identification using social media reviews.

Table 1. Multilingual Sarcastic Identification using Social Media Reviews – Tabular Review

Ref.	Dataset	Pre-processing	Language	Method	Results
(lsa Scola 2021)	consists of 28,503 headlines,	word embedding, Remove the bag of words,	English	BiLSTM, BERT, SVM, CNN,	87%, 91%, 79%, 86%
(Deepali D.Londhe 2021)	Utsab Barman et al. used the FIRE 2013 and FIRE 2014 datasets for their analysis.	Identification, Classification, Normalization,	Multiple languages	Ling Pipe classifier & Textcat toolkit, Google Translate and Microsoft A.P., SVM	89%
(A. Palaniammal 2023)	Headline 2019 dataset	Classification, Noise removal, Feature extraction,	English	ARO-MCEDNN, Glove approach for word embedding,	98.70%
(Ashwani Kumar 2023)	information from tweets, discussions, and comments	Data cleaning, Noise removal, Normalization, Handling missing values, Feature extraction	English, Hindi, French, German,	DM techniques, Text mining, NLP, A.I., ML techniques (Vader Algorithm)	89% 83%
(Shekhar 2021)	9,752words (code-mix text)	Tokenization word embedding	English Hindi	Rule-framed approach HLSTM learning model	83%
(Poornachandran 2022)	Dataset retrieved from FIRE 2020 & EACL 2021,	Tokenization, Segmentation Noise removal	Malayalam, English,	Word2Vec, FastText, XLM-R,	77%
(Matic Rašl 2021)	Twitter-based English dataset Slovene dataset	Tokenization Embedding Cross-validation	English, Slovene,	RoBERTa, Distil-Bert, Distil-Bert-multilingual	0.72 f1-score (English) 0.88 (Slovene)

(Thara 2021)	50K sentences extracted from YouTube.(code-mixed)	Feature extraction Tokenization	English Hindi	BERT, CamemBERT, DistilBERT,	74%
(Liana Ermakova 2023)	JOKER track at CLEF 2023 SemEval-2017 SemEval-2021	Feature extraction,	English, French	Ridge, NB, FastText	75%
(Mohammad Zia Ur Rehman a 2023)	SCIDN dataset contains 1.5M & MACI 665k comments	Data cleaning, Transliteration,	English, French, Spanish, Italian.	MLP, LSTM, XLMR, mBERT,	95% accu.
(Amal Alqahtani 2023)	Reviewed based (between 2019 and 2022)	Feature extraction, Tokenization	English	Lexiconstransformers, Naive Bayes, KNN, Decision Tree, BERT, STM, SVM.	99.1%
(Aqsa Younas 2020)	MultiSenti Code-mixed dataset	Categorization Removal of bag of words Labeling	English Roman Urdu	Deep learning approaches mBERT, XLM-R	0.65 F1 score
(Elena Zotovaa 2021)	CIC Corpus - SemEval 2016 dataset, Multilingual TW-10 corpus	Labeling Feature extraction Cross-validation	English	SVM, XLNET, RoBERTa, mBERT, XLM-RoBERTa	75.10 F1 score
(K Maity 2022)	25000 images or memes	Remove missing values and irrelevant data. Covert images into text data	English, Dutch, Hindi, Code-mixed	Proposed CLIP-CentralNet, ResNet, mBERT,	Improved 3.18% and 3.10% in terms of accuracy and F1 score,
(Manjot Bedi 2021)	MaSaC (code mixed) Consists of 36,000 Hindi and 3,000 English words.	Alignment, Noise removal, Feature extraction	Hindi, English	Proposed MSH-COMICS, a neural architecture for sarcasm	83% for sarcasm, 87% for Humor
(Ritesh Kumar 2021)	HASOC (8000 posts from Twitter and Facebook) & TRAC-2	Noise removal,	Hindi, Bangla, English	SVM, BERT, ALBERT and DistilBERT classifiers	80% accu
(ALAA RAHMA 2023)	review between 2017 and 2022 on Arabic sarcasm detection	Data cleaning, Noise removal, Feature extraction, Feature space,	Arabic	SVM, L.R., NB, LRCV, deep learning models	85% accu
(Sanzana Karim Lora 2023)	25,636 sarcastic comments on Facebook	Tokenization, Noise removal,	Bengali, English	Transformer-based generative adversarial learning,	97.2% accu

(Sayani Ghosal 2023)	10,000(hate speech) 20,000(non-hate speech) from social media posts and news articles	Noise removal, Data cleaning, Stop words, Empty/missing values removal, Data analysis,	Bengali,	FSVMCIL, mBERT, Morphological analysis method, Hate similarity (H.S.) scheme	85% accu
(Arun Kumar Yadav, 2023)	Consolidated dataset of 20600 instances.	Noise removal, missing values removed, Data cleaning,	Hindi, English,	Machine Learning, Deep Learning methods (CNN-BiLSTM)	87% accu
(A Ameur 2023)	SemEval-2016 for hotel reviews	Noise removal, Tokenization, Fine-tuning	Arabic	AraBERT, Machine learning models(SVM and Logistic Regression)	91%. accu
(Bharathi Raja Chakravarthi a 2023)	Code-mixed (Tamil-English) 12,795 comments extracted from YouTube.	Normalization, Removing noise, Feature extraction, Data analysis	Tamil, English	BACD, FGACD, D.T. technique, TFIDF, NB classifier	84% accu (code mixed)
(SK Bharti 2022)	Mustard dataset, 690 videos	Segmentation, MFCC, Tokenization,	English	Audio segment model, hybrid model(audio,text)	70.35% F1-score
(CHRISTOPHER IFEANYI EKE 2021)	published article from the period of 2008 to 2019 content-based linguistic features	Data cleaning, Noise removal, Tokenization, Stemming,	English & some other languages	-Content-based feature extraction techniques Bag of Words, word2vec,	90% accu
(Shivani Kumar, 2022)	WITS dataset contains 2240 sarcastic text	Classification, Noise removal, Vectorization, Cosine similarity	Hindi, Arabic, Italian	Multi-modal Aware Fusion module, MCA2, BARTmBART	91% accu
(Eke 2021)	Malayalam–English code-mixed dataset	Feature extraction, tokenization	Malayalam English	XLM-R, Word2Vec, FastText, BiGRU, RNN,	0.76% F1-score.
(Ramchandra Joshi 2021)	Constraint@AAAI 2021 contains 8192 online posts collected from social media.	Data collection, Aggregation, Sampling,	Hindi	Multi-CNN, Bi-LSTM, CNNBi-LSTM, mBERT, IndicBERT,	79.93% accu
(Ibtissam Touahri 2021)	5,030 Arabic political reviews.	Data cleaning, Noise removal, Feature extraction	Arabic,	LSTM architecture, SVM, Random Forest, RNN	89.24% accu
(Christopher Ifeanyi Eke 2020)	Internet Argument Corpus version 2 (IAC-v2) consists of 3260 posts per class for sarcasm	Data acquisition, Feature extraction, Training	English	Proposed BERT model, Deep learning model with Bi-LSTM,	89.24 accu.

detection.

2.2 Multilingual Sarcastic Identification Using Tweets Dataset

(Dalya Faraj 2021) Worked on different methods like Ensemble technique with AraBERT pre-trained model,Naive Bayes Multinomial, also performed Data cleaning ,Noise removal & normalization on the twitter's dataset 10,000 labeled Arabic tweets. Results produced 78% accuracy. The (Jens Lemmens 2020) (Jens Lemmens 2020) worked on the 5,000 tweets & 4,400 Reddit comments. They performed 10-fold cross-validation and then compared different models SVM,LSTM, CNN, MLP. Results showed 74% for Twitter & 67% for Reddit . (Lad 2023) Used 3466 & 3102 in English & Arabic tweets, respectively. It performed Tokenization, Verb extracting, Classification, Augmentation and then used different models RoBERTa for English binary classification, XLM-RoBERTa for Arabic binary classification, BERT for English multilabel classification. Highest accuracy achieved 80%. The (Bashar Talafha 2021) worked on Cross validation method, MSE method on the dataset of 1554 tweets. Basically they evaluated the loss occurred to detect sarcasm. They performed Tokenization and classification on the dataset of Arabic tweets. Evaluation loss occurred 0.011631458. The (Neha Garg, 2022) worked on the methodology consisting of naive Bayes, SVM, LDA, TF-IDF on the dataset of 18,000 tweets in English. Clustering, Lemmatization, Stemming, part of speech, Normalization, Stop word removal, and Feature extraction on the dataset. It achieved 91% accuracy. The (Matic Rašl, 2021) performed pre-processing techniques i-e Feature extraction, Tokenization Twitter-based English dataset & Slovene dataset in English & Slovene languages. They proposed RoBERTa, Distil-Bert, and DistilBert-multilingual methods, and Results showed a 0.72% F1-score for English and 0.88% F1-score for Slovene. Moreover, this paper Used some pre-processing techniques, i-e, normalization and noise removal on the dataset of 3,564 Tweets in Urdu by using the different SVM, MLP, ML & DL classifiers, LSTM approaches, the (MAAZ AMJAD1 2021) achieved 75% accuracy form SVM & 72% in MLP for sarcasm detection.

The (THARINDU RANASINGHE 2021) performed pre-processing techniques classification, identification & detection and proposed cross-lingual contextual embedding model, also used neural networks, transfer learning, supervised learning by classification techniques on the dataset of English (14,100), Arabic (8000), Bengali (4000), Danish (2,961), Greek (8,743), Hindi (8000), Spanish (66,00), Turkish (31,756). Results showed the robustness of cross-lingual contextual embedding, 0.8415 F1 (Bengali), 0.8532 F1 (Danish), 0.8701 F1 (Greek), 0.8568 F1 (Hindi), 0.7513 F1 (Spanish). (S. MUHAMMAD AHMED HASSAN SHAH 2023) performed different pre-processing techniques i-e Data analysis, Stop word removal, Stemming, Word embedding, N gram on the dataset of 27,000 Arabic tweets. They proposed a Modified Switch Transformer (MST) for detecting sarcasm, SWE, MARBERT, and MTL models and obtained a 90% higher recall rate and 81.66% Accuracy.

Plaza-Del-Arco, (FLOR MIRIAM PLAZA-DEL-ARCO 2021) Worked on different methods like Multi-task Learning (MTL) approach, Transformer-based model, also performed Data cleaning, Removing duplication & missing values Tokenization, Assigning equal weight on the twitter's dataset HatEval (6,600), MEX-A3T (10,475), interfaces (14,626), event (8,409) tweets. Results produced 78% macro-F1for HatEval & 86% for MEX-A3T, respectively. (Debby Alita, 2019) worked on the dataset of January 2018 to March 2018 tweets in Indonesian. They performed tokenizing, case-folding, stopword removal, classification, and feature extraction and then compared different model classification techniques, such as Naive Bayes Classifier, SVM, and Random Forest classifier. Results showed 61% for SVM and 62% for Random Forest method.

The (Rajshree Singh, 2022) introduced 5250 English-Hindi code-mixed tweets dataset. Noise removal, Data cleaning, n- gram was used for data pre-processing on the dataset and proposed an approach "the balancing layer" & other methods like Random forest, RBF-kernel, SVMs, DT. More than 90% accuracy was achieved. The (MY Khan, 2023) proposed a CR-based approach, Count vectorization, TF-IDF,Bag-of-words representation, and Unigram and n-gram word sequencing models on the dataset of 7,000 tweets in standard Urdu. Basic pre-processing techniques, i.e., Data cleaning and noise removal performed on the dataset. Results showed 85.5% accuracy.

(Shubhi Bansala 2024) They used pre-processing techniques, i-e, noise removal, removing null values, text conversion, feature extraction, feature refinement, and feature interaction on the 31,07,866 tweets,9,17,833 hashtags, & 4,78,120 users. They proposed TAGALOG, a graph-based neural network, and some other methods, mBERT, indicBERT, XLMR, and distilmBERT. Proposed methods achieved 82%. The

(AKSHI 2023) proposed a review on sarcasm detection of the dataset, a total of 1004 Hindi tweets. It performed pre-processing techniques like noise removal and feature extraction and used different methods like the Hybrid CNN-LSTM model, contexts - Emojis, ANN, and 97% accuracy was achieved. Table 2 showed the tabular review of Multilingual Sarcastic Identification using Tweets dataset.

Table 2. Multilingual Sarcastic Identification using Tweets Dataset – Tabular Review

Ref.	Dataset	Pre-processing	Language	Method	Results
(Dalya Faraj 2021)	10,000 labeled Arabic tweets	Data cleaning, Noise removal,	Arabic	Ensemble technique with AraBERT pre-trained model	78% accu
(Jens Lemmens 2020)	5,000 tweets & 4,400 Reddit comments	10-fold cross-validation.	English	SVM, LSTM, CNN, MLP,	74% accu (Twitter) 67% accu (Reddit)
(Lad 2023)	3466 English & 3102 Arabic tweets	Tokenization, Verb extracting, Classification, Augmentation,	English, Arabic,	Roberta XLM-Roberta BERT	80% Accu.
(Bashar Talafha 2021)	1554 tweets.	Tokenization, Classification,	Arabic	Cross-validation method, MSE,	0.011631458 Evaluation loss
(Neha Garg 2022)	18,000 tweets	Clustering Stemming, part of speech,	- English	naive Bayes, SVM, LDA, TF-IDF	91% accu
(Matic Rašl 2021)	Twitter-based English dataset Slovene dataset	Tokenization Embedding Cross-validation	English, Slovene,	Roberta, Distil-Bert, Distil-Bert-multilingual	0.72 f1-score (English) 0.88 (Slovene)
(MAAZ AMJAD1 2021)	3,564 Urdu Tweets	Normalization,	English, Urdu,	SVM, MLP, ML & DL classifiers, LSTM	75% accu (SVM), 72% accu (MLP)
(THARINDU RANASINGHE 2021)	English (14,100), Arabic (8000), Bengali (4000), Danish (2,961), Greek (8,743), Hindi (8000), Spanish (66,00), Turkish (31,756).	classification, identification, detection	Bengali, Danish, Greek, Hindi, Spanish	cross-lingual contextual embedding model, Neural networks, transfer learning, supervised learning by classification techniques	.8415 F1 (Bengali), 0.8532F1 (Danish), 0.8701F1 (Greek), 0.8568 F1 (Hindi), 0.7513 F1 (Spanish).
(S. MUHAMMAD AHMED HASSAN SHAH 2023)	27,000 Arabic tweets.	Data analysis, Stop word removal,	Arabic	Modified Switch Transformer (MST) for detecting sarcasm SWE, MARBERT, MTL	recall rate90% R.F. Accuracy 81.66%

(FLOR MIRIAM PLA-ZA-DEL-ARCO 2021)	HatEval (6,600), MEX-A3T (10,475), interTASS (14,626), emoEvent (8,409) tweets	Data cleaning, Removing duplication & missing values Tokenization,	Spanish	Multi-task Learning (MTL) approach, Transformer-based model	78%
(Debby Alita 2019)	January 2018 to March 2018 tweets in the Indonesian language	tokenizing, case-folding classification, feature extraction	Indonesian language	Classification Naive Bayes Classifier, SVM, Random Forest classifier	61% (SVM), 62% (Random Forest)
(Rajshree Singh 2022)	5250 English-Hindi code-mixed tweets.	Noise removal, Data cleaning, n- gram,	Hinglish (Hindi, English)	proposed approach "the balancing layer," Random forest, SVMs,	more than 90%
(MY Khan 2023)	Seven thousand tweets in standard Urdu.	Data cleaning, Noise removal,	Urdu	proposed CR-based approach, Count vectorization	85.5%
(Shubhi Bansala 2024)	31,07,866 tweets, 9,17,833 hashtags, & 4,78,120 users.	noise removal, conversion of text, feature extraction,	Bangla, Hindi, Kannada, Gujrati, Tamil,	TAGALOG proposed graph-based neural network mBERT, indicBERT, XLMR, distilmBERT	82% (hit rate)
(AKSHI 2023)	A total of 1004 Hindi tweets	Noise removal, Feature extraction,	Hindi	Hybrid CNN-LSTM model, contexts - Emojis, ANN	97%

2.3 Multilingual Sarcastic Identification Using Multiple Reviews on Movies

(Bharathi et al. 2023) introduced the dataset of DravidianCodeMix (consists of offensive language data from Tamil, Malayalam, and Kannada movie trailers on YouTube). They used Feature extraction and NLTK package for data pre-processing on the dataset and then used Ensemble of models, GA-based approaches, Fusion of MPNet, and CNN model. As a result, a 98% F1 score was achieved. (ABDUL GHAFOOR 2021) Used some pre-processing techniques, i-e, POS Tagging and validation on the IMDB English movie review dataset methods like Google translator API, Machine learning, and deep learning models for translation in English, Urdu, Hindi, and German, achieved 90%, 87%, 85%,90% accuracy, respectively.

They used pre-processing techniques, i.e., Segmentation, MFCC, and Tokenization on the Mustard dataset, and 690 videos using the Audio segment model, Text model, and hybrid model (audio, text). The (SK Bharti, 2022) achieved 70.35% F1-score. (Akshi Kumar, 2021) the proposed XGBoost model, random forest, and SVM were introduced on the 690 dialogues from four famous television shows' datasets. They performed data cleaning and noise removal on the dataset, and 93% accuracy was obtained. Table 3 showed the tabular review of Multilingual Sarcastic Identifications using Movie reviews.

Table 3. Multilingual Sarcastic Identification using Multiple Reviews on Movies – Tabular Review

Ref.	Dataset	Pre-processing	Language	Method	Results
(Bharathi Raja Chakravarthi 2023)	DravidianCode-Mix(consists of offensive language data from Tamil, Malayalam, and	Feature extraction, NLTK package was used for data pre-processing.	Tamil, Malayalam, Kannada,	Ensemble of models, GA-based approaches, Fusion of MPNet	98% F1-score

	Kannada movie trailers on YouTube.			and CNN model,	
(ABDUL GHAFOOR 2021)	IMDB English movie review dataset	POS Tagging, Cross-validation	English, Urdu, Hindi,	Google translator API, Machine learning and deep learning models,	90% 87% 85% 90%
(SK Bharti 2022)	Mustard dataset, 690 videos	Segmentation, MFCC, Tokenization,	English	Audio segment model, hybrid model(audio,text)	70.35% F1-score
(Akshi Kumar 2021)	Six hundred ninety dialogues from four famous television shows.	Data cleaning, Noise removal,	English	Proposed XGBoost model, Random forest, SVM,	93%

3. Discussion

The section focuses on future directions of sarcasm detection from the analysis of this study. The majority of recent research focuses on English, ignoring other languages. Future directions include:

1) Considering more languages: The core of the current sarcastic detection is monolingual or bi-lingual, ignoring other languages as they only pay attention to English. In this direction, one realistic avenue is to work more in multilingualism by considering multiple languages.
2) Tweet correctness techniques: Twitter was frequently used data source to train sarcasm detection model evaluation in most articles studied. On the other hand, tweets are probably full of misspelled words that might interfere with the model's performance. An alternative way of evolution for them is using an automatic word-correction method in the early stages of the development process to a set of sarcastic recognition systems.
3) Exploring other social network sources: commonly used datasets in the reviewed articles were only Twitter and Reddit. Although both of them share the ability to provide quality data, the other social networks that should be explored further in this domain include Facebook and Instagram.
4) Building multi-modal sarcasm detection models: Recent studies on sarcasm detection mainly include limited text-based data. So, there is a need to work more on analyzing multi-modal models.

Use of emojis: This is why an emoji has to be inserted on as many messages used in some posts of these social networker accounts for a sarcastic text from anywhere else around the world they are. This implies that many other such research should also be conducted on the newly emerging ideas that may lead to further improvement in classification models.

4. Conclusions

This paper provides a systematic review of A.I. techniques for sarcasm detection. Identifying gaps in current literature on the identification of sarcasm significantly contributes to Natural Language Processing (NLP) specifically targeted at the key areas. The review focuses on the research gap in this research area. Very little work is done in multiple languages for sarcasm detection. Moreover, this would lead to enhanced accuracy of sarcasm detection models. In addition, the study will focus on additional elements that go beyond linguistic factors like behavioral, contextual, and visual cues in an effort to address multilingual sarcasm detection. Machine learning would be superseded by deep leaning as a better performance in feature extraction is observed, increasing efficiency and accuracy for sarcasm identification. In addition, this review recognizes the need to identify sarcasm in social media comments. Lastly, the paper emphasizes that a multilingual approach is vital in sarcasm identification. It recognizes that people speak in their mother tongue more truthfully. Hence, it is necessary to extend research beyond English datasets. Future research should focus on developing approaches for multilingual sarcasm detection.

There is a need to find methods that work well with different languages. Cross-lingual contextual embedding has shown robustness in detecting sarcasm in various languages. More datasets need to be created for languages other than English.

References

1. A Ameur, S. H. (2023). Domain adaptation approach for Arabic sarcasm detection in hotel reviews based on hybrid learning. Elsevier, 11.
2. A. Palaniammal, P. A. (2023). researchgate.net, 13.
3. ABDUL GHAFOOR, A. S. (2021). The impact of translating resource-rich datasets to low-resource languages through multi-lingual text processing. IEEE, 13.
4. ADITYA JOSHI, P. B. (2016). Automatic Sarcasm Detection: A Survey. Joshi, A., Bhattacharyya, P., & Carman, M. J. (2017). Automatic sarcasm detection: A survey. ACM Computing Surveys (CSUR), 50(5), 1-22., 17.
5. Aditya Joshi, V. S. (2015). Harnessing Context Incongruity for Sarcasm Detection. 1, 6.
6. Agarwal, M. S. (2020). Tomato leaf disease detection using convolution neural network. Procedia Computer Science. 293–301.
7. Akshi Kumar, S. D. (2021). Explainable artificial intelligence for sarcasm detection in dialogues. hindawi.com.
8. AKSHI, A. K. (2023). Hybrid deep learning model for sarcasm detection in Indian indigenous language using word-emoji embeddings. Association for Computing Machinery (ACM), p. 22.
9. ALAA RAHMA, S. S. (2023). A Comprehensive Survey on Arabic Sarcasm Detection: Approaches, Challenges and Future Trends. IEEE, 20.
10. Amal Alqahtani, L. A. (2023). Text-based Sarcasm Detection on Social Networks: A Systematic Review. (IJACSA) International Journal of Advanced Computer Science and Applications,, 17.
11. Amjad, K. &. ((2021)). Critical review on multi-crops leaves disease detection using artificial intelligence methods. Int J Sci Eng Res, 12, 2.
12. Aniruddha Ghosh, T. V. (2016). Fracking Sarcasm using Neural Network.
13. Aqsa Younas, R. N. (2020). Sentiment Analysis of Code-Mixed Roman Urdu-English Social Media Text using Deep Learning Approaches. IEEE, p. 6.
14. Arun kumar yadav, A. K. (2023). Hate Speech Recognition in Multilingual Text: Hinglish Documents. National Institute of Technology Hamirpur, p. 19.
15. Ashvani Kumar, S. K. (2023). RESEARCH ON THE TECHNIQUE OF SENTIMENT ANALYSIS WITH SARCASM DETECTION IN MULTILINGUAL DATA. irjmets.com, 8.
16. Bashar Talafha, M. E.-A.-K. (2021). Sarcasm detection and quantification in Arabic tweets. ieeexplore.ieee.org, 8.
17. Batmavady, S. &. (2019). Detection of cotton leaf diseases using image processing. Int J Rec Technol Eng (IJRTE), 8.
18. Bharathi Raja Chakravarthi, R. P. (2023). Detecting abusive comments at a fine-grained level in a low-resource language. www.elsevier.com/locate/nlp, 17.
19. Bharathi Raja Chakravarthi, M. B. (2023). Offensive language identification in Dravidian languages using MPNet and CNN. Elsevier, p. 18.
20. Christopher Ifeanyi Eke, A. A. (2020). Sarcasm identifcation in textual data: systematic review, research challenges and open directions. ResearchGate, p. 45.
21. CHRISTOPHER IFEANYI EKE, A. A. (2021). Context-based feature technique for sarcasm identification in benchmark datasets using deep learning and BERT model. IEEE, 18.
22. D.Kokane, M. C. (2016). A Review: To Detect and Identify Cotton leaf disease based on pattern recognition technique. International Journal Of Engineering And Computer Science, 4,, 110–118.
23. Dalya Faraj, M. A. (2021). Sarcasmdet at sarcasm detection task 2021 in arabic using arabert pretrained model. aclanthology.org, 6.
24. De Luna, R. G. ((pp. 1414-1419).). Automated image capturing system for deep learning-based tomato plant leaf disease detection and recognition. In TENCON 2018-2018 IEEE Region 10 Conference,IEEE, (2018, October). .
25. Debby Alita, S. P. (2019). Analysis of emoticon and sarcasm effect on sentiment analysis of Indonesian language on Twitter. download.garuda.kemdikbud.go.id, 10.

26. Deepali D.Londhe, A. K. (2021). Challenges in multilingual and mixed script sentiment analysis. IEEE, 6.
27. Despotović, I. G. (2015). MRI segmentation of the human brain: challenges, methods, and applications. Computational and mathematical methods in medicine, 2015.
28. Dhikhi, T. T. (2019). Measuring the size of an object using computer vision. International Journal of Innovative Technology and Exploring Engineering, 8(4), 424-426.
29. Durmuş, H. G. (2017). Disease detection on the leaves of the tomato plants by using deep learning. 6th International conference on agro-geoinformatics,IEEE, (pp. 1-5).
30. Eke, C. I. (2021). Multi-Feature Fusion Framework for Automatic Sarcasm Identification in Twitter Data. search.proquest.com, 24.
31. Elena Zotovaa, R. A. (2021). Semi-automatic Generation of Multilingual Datasets for Stance Detection in Twitter. Elsevier, 29.
32. FLOR MIRIAM PLAZA-DEL-ARCO, M. D.-G.-L.-V. (2021). A multi-task learning approach to hate speech detection leveraging sentiment analysis. IEEE, 12.
33. Francesco Barbieri, H. S. (2014). Modelling Sarcasm in Twitter, a Novel Approach. Association for Computational Linguistics, 50.
34. Ganatra, N. &. (2018). A survey on diseases detection and classification of agriculture products using image processing and machine learning. International Journal of Computer Applications, 1-13.
35. Gharghory, S. M. (2020). Performance Analysis of Efficient Pre-trained Networks based on Transfer Learning for Tomato Leaf Diseases Classification. International Journal of Advanced Computer Science and Applications., 11(8).
36. Hee, C. V. (2017). Can machines sense irony?: exploring automatic irony detection on social media. 1Examples by Kumon-Nakamura et al. (1995), 178.
37. Hsu, T.-S. W.-C. (2015). An improvement stereo vision images processing for object distance measurement. International Journal of Automation and Smart Technology, 5(2), 85-90.
38. Ibtissam Touahri, A. M. (2021). Enhancement of a multi-dialectal sentiment analysis system by the detection of the implied sarcastic features. Elsevier, 13.
39. Jens Lemmens, B. B. (2020). Sarcasm detection using an ensemble approach. aclanthology.org, 6.
40. Jihen Karoui, F. B.-G. (2017). Exploring the Impact of Pragmatic Phenomena on Irony Detection in Tweets: A Multilingual Corpus Study. http//www.spinoza.it, 11.
41. K Maity, P. J. (2022). A Multitask Framework for Sentiment, Emotion and Sarcasm aware Cyberbullying Detection from Multi-modal Code-Mixed Memes. SIGIR ’22, July 11–15, 2022, Madrid, Spain, 11.
42. Lad, R. (2023). Sarcasm Detection in English and Arabic Tweets Using Transformer Models. digitalcommons.dartmouth.edu, 59.
43. Lee, J. M. (2021). Real distance measurement using object detection of artificial intelligence. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(6), 557-563.
44. Li, D. Q. (2013). A new moving object detection approach with adaptive double thresholds. 2013 International Conference on Computational and Information Sciences, 102-105.
45. Liana Ermakova, A.-G. B. (2023). The JOKER Corpus: English-French Parallel Data for Multilingual Wordplay Recognition. SIGIR ’23, July 23–27, 2023, Taipei, Taiwan, 11.
46. Llorca, C. Y. (. (2018, March)). Image-based pest and disease recognition of tomato plants using a convolutional neural network. Proceedings of international conference technological challenges for a better world.
47. lsa Scola, I. S. (2021). Detección de Sarcasmo con BERT. Fundación Dialnet, 12.
48. MAAZ AMJAD1, N. A. (2021). Threatening language detection and target identification in Urdu tweets. ieeexplore.ieee.org, 12.
49. Manjot Bedi, S. K. (2021). Multi-modal Sarcasm Detection and Humor Classification in Code-mixed Conversations. ieeexplore.ieee.org, 13.

50. Matic Rašl, M. Ž. (2021). Transformer-based Sarcasm Detection in English and Slovene Language. researchgate.net, 5.
51. Mohammad Zia Ur Rehman a, S. M. (2023). User-aware multilingual abusive content detection in social media . Elsevier, 12.
52. MY Khan, T. A. (2023). Cognitive relationship-based approach for Urdu sarcasm and sentiment classification. IEEE, 30.
53. Neha Garg, K. S. (2022). Text pre-processing of multilingual for sentiment analysis based on social network data. academia.edu, 9.
54. nnabel, L. S. (2019, April)). Machine learning for plant leaf disease detection and classification–a review. International conference on communication and signal processing (ICCSP),IEEE, (pp. 0538-0542).
55. Patki, S. S. (2016). A review: Cotton leaf disease detection. . IOSR Journal of VLSI and Signal Processing, 6(3), 78-81.
56. Poornachandran, S. T. (2022). Social media text analytics of Malayalam–English code-mixed using deep learning. Journal of big Data, 2022•Springer, 25.
57. Prashar, K. T. ((2019, April)). CNN based on overlapping pooling method and multi-layered learning with SVM & KNN for American cotton leaf disease recognition. International Conference on Automation, Computational and Technology.
58. Pu, L. T.-C. (2016). Novel object-size measurement using the digital camera. 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 543-548.
59. Rajshree Singh, D. R. (2022). A Novel Balancing Technique with TF-IDF Matrix for Short Text Classification to Detect Sarcasm. kalaharijournals.com, 6.
60. Ramchandra Joshi, R. K. (2021). Evaluation of deep learning models for hostility detection in Hindi text. ieeexplore.ieee.org, 5.
61. Revathi, P. &. (2012). Homogenous segmentation based edge detection techniques for proficient identification of the cotton leaf spot diseases. International Journal of Computer Applications, 47(2), 875-888.
62. Ritesh Kumar, B. L. (2021). Aggressive and Offensive Language Identification in Hindi, Bangla, and English: A Comparative Study. Springer Nature Singapore Pte Ltd 2021, 20.
63. S. MUHAMMAD AHMED HASSAN SHAH, S. F. (2023). Arabic Sentiment Analysis and Sarcasm Detection using Probabilistic Projections Based Variational Switch Transformer. IEEE, 17.
64. Sanzana Karim Lora, I. J. (2023). A transformer-based generative adversarial learning to detect sarcasm from Bengali text with correct classification of confusing text. Heliyon, 18.
65. Sayani Ghosal, A. J. (2023). Inculcating Context for Emoji Powered Bengali Hate Speech Detection using Extended Fuzzy SVM and Text Embedding Models. ACM Trans. Asian Low-Resour. Lang. Inf. Process, 24.
66. Shah, N. &.-4. ((2019, February)). Detection of disease in cotton leaf using artificial neural network. Amity International Conference on Artificial Intelligence (AICAI) IEEE, (pp. 473-476).
67. Shekhar, S. (2021). Neural Network Pun Material Identifiction Framework based on Artificial Intelligence Learning. IOP Publishing, 10.
68. Shivani Kumar, A. K. (2022). When did you become so smart, oh wise one?! Sarcasm Explanation in Multi-modal Multi-party Dialogues. arxiv.org, 13.
69. Shubhi Bansala, K. G. (2024). Multilingual personalized hashtag recommendation for low-resource Indic languages using graph-based deep neural network. ResearchGate, 36.
70. SK Bharti, R. G. (2022). Multi-modal sarcasm detection: a deep learning approach. hindawi.com, 25.
71. Tay†, Y. (2018). Reasoning with Sarcasm by Reading In-between. Tay, Y., Tuan, L. A., Hui, S. C., & Su, J. (2018). Reasoning with sarcasm by reading in-between. arXiv preprint arXiv:1805.02856., 11.
72. Thara, S. A. (2021). Transformer based language identification for malayalam-english code-mixed text. . IEEE Access,, 14.
73. THARINDU RANASINGHE, M. Z. (2021). Multilingual offensive language identification for low-resource languages. arxiv.org, 13.
74. Tripathy, S. ((2021, November).). Detection of cotton leaf disease using image processing techniques. In Journal of Physics:. IOP Publishing., Conference Series (Vol. 2062, No. 1, p. 012009).

75. Tsironis, V. B. ((2020).). Tomatod: evaluation of object detection algorithms on a new real-world tomato dataset. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 43, 1077-1084.
76. Udawant, P. &. (2019). Diseased portion classification & recognition of cotton plants using convolution neural networks. International Journal of Engineering and Advanced Technology, 8(6), 3492-3496.
77. Zilvan, V. R. (2022). Convolutional variational autoencoder-based feature learning for automatic tea clone recognition. Journal of King Saud University-Computer and Information Sciences, , 34(6), 3332-3342.