INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

Detecting Multi-label emotions from code-mixed Facebook Status Updates

Swarnangini Sinha[1]*, **Kanak Saxena**[2], **Nisheeth Joshi**[3]

1 Research Scholar, Department of Computer Science, Banasthali Vidyapith, Rajasthan, 304022, India
2 Department of Computer Application, Samrat Ashok Technological Institute, Vidisha, 464001, Madhya Pradesh, India
3 Department of Computer Science, Banasthali Vidyapith, Rajasthan, 304022, India

Abstract

Objectives: With the growth of Social Media and the increasing use of English-Hindi (Hinglish) in linguistically diverse countries such as India, it is becoming increasingly important to analyse Hinglish-language content on Social Media platforms like Facebook. Prior sentiment and emotion analyses have only focused on single-label classification, ignoring the possibility of coexisting emotions within one instance. By analysing code-mixed Facebook status updates, the study aims to investigate multiple emotions. **Method:** 15,995 English-Hindi mixed Facebook status updates are annotated with emotions like joy, sadness, anger, fear, trust, disgust, surprise, anticipation, and love. Different pre-processing techniques are used to normalize the noisy data to produce more accurate results. We apply five different multi-level classification algorithms with word-level and character n-gram approaches to test the best classification results. **Findings:** The results of the experiment indicate that a status update can evoke multiple emotions rather than just one. Precision, recall, F1 score, and accuracy using both Micro and Macro averaging are used to evaluate the performance of different classifiers. As compared to other classification algorithms, the Classifier Chains algorithm with its 2-6-gram approach has the highest accuracy of 86% with a precision of 0.98. As compared to other classifiers, the Classifier Chains algorithm offered better results due to its ability to consider the correlations between class labels. **Applications:** The article focuses on the multi-label emotion classification task, which examines whether a Facebook status update shows none, one, or more of the nine emotions as outlined by Plutchik's wheel of emotions. Considering the emotion of a text can support decision-making processes in various ways.

Keywords: Social Media; Emotion Analysis; Code-Mixed; MultiLabel 29 Classification; Emojis; Lexicons

1 Introduction

The initiative of the Indian government towards digitalization and the spread of the Internet in recent years has resulted in the massive growth in the digital population of

India which is predicted to cross the figure of 650 million Internet users by the year 2023[1]. Nowadays the social media is the buzzword amongst youngsters. According to a survey conducted by Pew Research Centre[2], about 95% of youngsters have smartphones and 45% of them are found to be constantly online. The availability and access to smartphones and Internet connectivity at a cheaper cost has resulted in an estimation of more than 3 billion people will be using social media in near future across the world. The widespread use of social media has resulted in the generation of an enormous amount of social media text. The text written by people of different strata belonging to varied cultural backgrounds has given birth to a very informal way of textual communication with diverse linguistic distinctions. In a country like India where people belonging to multilingual societies, it is observed that informal communication is mainly a mix of transliterated Hindi i.e. (Roman Hindi) with English or English with their regional languages called code-mixing. In India, most of the social media texts are written using code-mixing more precisely using the mix of English and Hindi languages.

More than 2.6 billion people[3] use Facebook each month, making it the most popular social network in the world until recently. Young people from low-income families still find it appealing. Statistics published by Statista[4] indicate that India has more than 290 million Facebook users followed by the United States with 180 million. Several things can be done on Facebook, including creating profiles, uploading photos and videos, sending messages, and keeping in touch with friends, family, and colleagues. People use Facebook to communicate with one another in an informal, and fast manner. According to Facebook surveys, approximately 83% of women and 75% of men are Facebook users among Internet users. Facebook users are also estimated to have 155 friends on average. According to Facebook usage statistics by age groups, 62% of people aged 65+ and 72% between 50-64 are active Facebook users. A whopping 88% of internet users aged 18-29 use Facebook, along with 84% of users aged 30-49. Facebook is used by 82% of college graduates. Having reviewed these statistics, we identified Facebook as the platform that represents a wide variety of emotions expressed in Facebook status updates.

People express their views and opinions in the form of emotions. Facebook has revolutionized the way people communicate with each other. Facebook channelizes the opinions, facts, emotions expressed by people on different topics in the form of text messages. Emotion analysis focuses on detecting the polarity (positive or negative) and the emotional state such as joy, anger, sadness, etc[5]. The earlier work on sentiment analysis and emotion analysis has mainly classified the text into either positive, negative, or neutral sentiments[6], or identified single emotion from the text data. On the other hand, in a multi-label emotion classification, a piece of text is found to be associated with all the emotions, a subset of emotions or no emotions[7–9]. Hence, in this article, we focus on the multi-label emotion classification task, which aims to identify the presence of all, some, or no emotions. In this study, we are using eight primary emotions (joy, sadness, anger, fear, trust, disgust, surprise, and anticipation) and one secondary emotion (love) as described in Plutchik's wheel of emotions[10].[Figure 1]
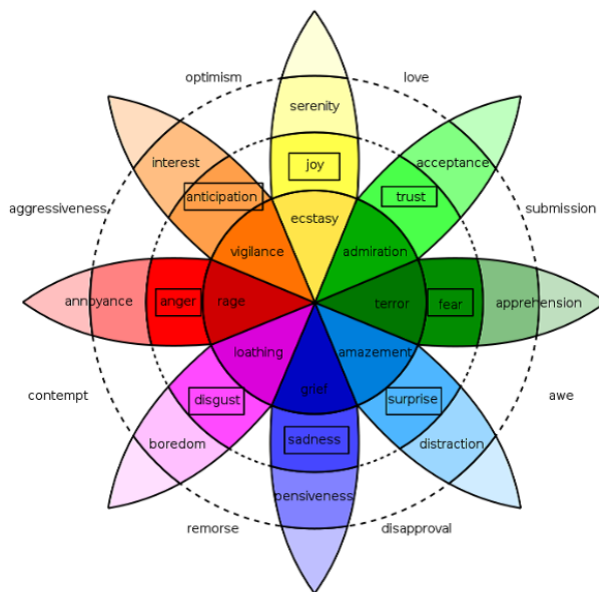

Fig 1. Plutchik's Wheel of /Emotions

There have been several early approaches for analysis of code-mixed data, mostly focusing on pre-processing [11–14], language identification [15–18], lexicon building [19], sentiment classification [20–22] and subjectivity analysis [23]. Traditional algorithms are built to work on single-label classification problems. Hence, in a lot of prior approaches, the multilabel problem is divided into multiple single label problems so that they can be applied with the existing single label algorithms. The emphasis of such classification is on finding the single label classifier's predictions individually which are then transformed into multi-label predictions. The most natural approaches called one to rest and binary relevance are used in multi-label problems to transform them into multiple individual binary classification problems, one problem for each label. Then, an independent binary classifier is trained to predict the relevance of one of the labels [24,25]. These are the simplest classification approaches to implement multi-label output. However, both these approaches of classification do not consider the possible correlations between class labels. This leads to chances of over fitting the label combinations because it does not take into consideration the association between the previously observed combinations of labels.

In this article, the other classification models like Classifier chains, Label power set, and a k-Nearest Neighbor (KNN) adaptive algorithm using a lazy learning approach named multi-Label k-Nearest Neighbor (ML-kNN) are presented. In all these algorithms the correlation between the underlying labels is considered. The classification on Facebook status updates with the given features shows that Classifier chains achieved better performance as compared to other experimented models.

The contribution of this work is summarised as under.

- Multi-label classification paradigm in Emotion Analysis domain.
- A real dataset composed of Facebook status updates labelled with nine emotions.
- Creation of Hindi-English emotion lexicons of 67,000 words labelled with nine emotions.
- Creation of 3000 emoji lexicons labelled with nine emotions.

To recognize the coexistence of different emotions, the proposed architecture uses a dataset of 15,995 code mixed Facebook status updates. Different algorithms are used to detect the coexistence of different emotions. Based on the results of the experiments, we found that our system is better than the state-of-the-art systems, surpassing them. In the rest of the article, the following information is presented. Described in Section 2 is relevant work on multi-label problem transformation and emotional analysis of Facebook statuses.

The rest of the article is arranged as follows. In Section 2, we overview the related work on multi-label problem transformation methods and emotion analysis of Facebook status updates. In Section 3, it is described how the exploratory analysis was conducted. Section 4 provides a detailed explanation of the methodology. In Section 5, we report the experimental results and evaluation of algorithms followed by the conclusion.

2 Materials and Methods

2.1 Exploring the Dataset

English-Hindi code mixed 32,693 Facebook status updates of 80 people are collected. The status updates collected are in the json format which consists of all the information such as timestamps, posts, titles, tags, and attachments. We have used the Python programming language for implementing the system. The status updates collected are stored together in the CSV File. Extensive semi-automated processing is carried out to remove all the noisy status updates. In the annotation phase, we further removed all those status updates which were not expressing any emotions. Finally, we have got the corpus containing 15,996 Facebook status updates with 99066 words. Among the 15,996 Facebook status updates 321 status updates were in Devanagari Script, 146,45 status updates in Roman Script, and 1002 in Mixed Script. The dataset annotation was carried out for segregating the status updates depending upon the language used in it. For each status update, a label was assigned to its source language. In the first phase, the database was divided using two kinds of labels namely, 'eng' and 'hin'.

S1: "**happy** birthday dear wishing you **a magical** birthday".

'eng' label was assigned to status updates which are present in English language vocabulary, such as "magical", and "happy" used in S1.

S2: "किस्मत न जैसा चाहा वैस ढल गए हम बहुत सभल क चल फिर भा **फिसल गए** हम किसा न **विश्वास ताडा** ता किसा न दिल आर लागा का लागा की बदल गए हम"

'hin' label was assigned to status updates which are present in the Devanagari Hindi vocabulary such as "विश्वास ताडा", and "फिसल गए" used in S2.

S3: "happy birthday dubey ji **kamyab** our **khush** rahiye"

The label 'mix' was assigned to status updates that are present in Roman (Transliterated) Hindi such as "khush", and "kamyab" used in S3.

Emotion annotation of the status updates was done with nine primary emotions, namely, Joy, Trust, Anticipation, Sadness, Anger, Fear, Disgust, Surprise, and one secondary emotion i.e., Love separately for all three kinds of labels. The initial Exploratory Data Analysis for the total number of posts in each emotion category and posts with multiple emotions reveals the following facts about the corpus as depicted in [Figures 2 and 3].
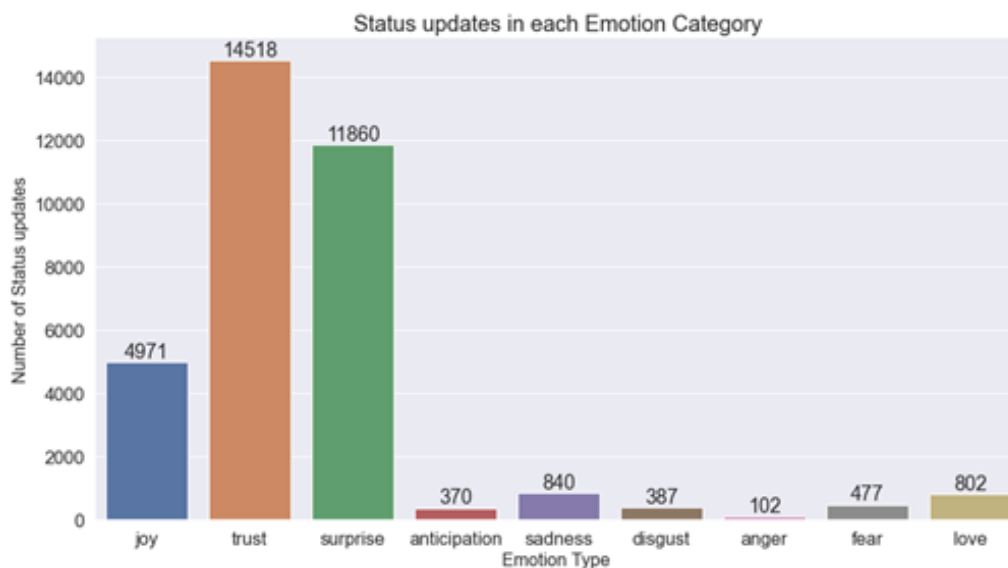

Fig 2. Count of status updates under each label

Fig 3. Count of status updates with multiple labels

2.2 System Architecture

After annotating the corpus, we try to detect emotions in all three kinds of noisy status updates. Thus, a data preparation task was employed to clean the data followed by feature identification and extraction and finally, the classification of emotions as joy, trust, anticipation, surprise, disgust, sadness, anger, fear, and love. The steps are described in sequential order as under.

2.2.1 Pre-processing of the noisy Status Updates

Noisy status updates: The unstructured status updates containing special characters, URLs, blank lines, multiple spaces, repeated characters, punctuations, and numbers were all cleaned up in a pre-processing step.

Conversion of Emoticon and Emoji to Words: In social media, people prefer to converse or express their feelings in the form of emoticons and emojis along with text data. In emotion analysis, they give some valuable information and so removing them might lead to losing important information related to underlying emotions. Hence, they are converted to their text equivalents.

Stop words: Stop words corpus obtained from NLTK was used to eliminate most unproductive words which provide little information about individual status updates. While doing so words like not, never that are depicting negation are preserved.

Slangs and Abbreviations: Most of the time status updates are written using popular slags and abbreviations. As they are used to convey emotions they are converted to their respective full forms.

Spelling correction: Typos are common in text data available on social media and thus spelling mistakes are corrected before the analysis.

Translation: We used Google translate app to translate all transliterated Hindi (Roman Hindi) scripts to Devanagari Hindi script. Some words were not translated by this app correctly. We have created a list of 67,000 transliterated words with their corresponding Hindi and English meanings. The Devanagari Hindi script is then translated to English using the google translate app.

2.2.2 Feature Engineering

The process of creating features by extracting information from the data was employed using various techniques as described below. These feature vectors are then used to train our machine learning models.

Word N-Grams: Bag-of-words is the simplest method of extracting features from the text. However, it treats each word as an independent word and counts its frequency. We use the word n-gram where n ranges from 2 to 6. However, it was improved by using Bag-of-words with n-grams because they captured the context around each word.

Character N-Grams: Character N-Grams are language independent and more reliable even though they increase the dimensionality of the problem. They also give information about their content and context. We use character n-gram where n ranges from 2 to 6.

TF-IDF Vectorization: TF-IDF weight is used as a statistical measure to identify how important the word is in that document. The value of a word increases proportionally to count, but it is inversely proportional to the frequency of the word in the corpus.

NRC lexicons: NRC Lexicon [26] containing 14,182 English unigrams was used for this study. Emotion is associated with each word of the lexicon if it has an association score of 1, otherwise, an association score of 0 is assigned. With weights associated with each lexicon, it was possible to preserve all the emotions represented by a single lexicon. If the training and testing sets belong to the same domain, the emotion lexicon improves classification accuracy significantly. Additionally, NRC lexicons have been augmented with 3000 emoji-based words to help represent emotions.

Emoji: Social media posts are largely dominated by emojis. Emotions are expressed nowadays using emojis rather than text messages. Emojis were assigned labels for nine different emotions to improve the classification of their emotion and sentiment.

2.2.3 Classification Models

The pre-processed corpus is used to develop feature vectors with the traditional Bag-of-Words (Bow)approach with word-n-gram and character n-gram techniques. The TF-IDF vectorizer is used to create an index for each word in the vocabulary. Before applying the baseline machine learning model, we balanced the dataset so that the train and test datasets have equally distributed class labels. We split the data into 70% for the training dataset and 30% for the test dataset. The class labels are mutually dependent in single-label classification whereas in multi-label classification as multiple class labels are not mutually exclusive special machine learning algorithms are required for predicting the output.

One-Vs-Rest: The algorithm uses the heuristic method to divide the multi-label problem into individual one label per class binary classification problems. And for an unseen instance, the predictions are made based on the class with maximum confidence. The algorithm also assumes that the classes are mutually exclusive without considering any underlying dependencies between them.

Binary Relevance: In this case, an ensemble of single-label binary classifiers is trained, one for each class. Each classifier forecasts if an instance belongs to a particular class or not. Thus, the multi-level output is the union of all classes that were predicted. However, the classifier does not take into consideration the correlations between class labels[27].

Classifier Chains: A chain of binary classifiers is constructed to decompose the multi-label classification problem into many binary problems. The labels of unseen instances are predicted sequentially by using the output of all preceding classifiers feature input for the subsequent classifiers[28].

Label Powerset: The algorithm considers the correlation between different class labels. It uses a training dataset to transform the multi-label problem into one binary classifier for every label combination[29].

ML-kNN : The algorithm first identifies the k-nearest neighbors for each instance of the training set. The maximum a posteriori principle is used to determine the label sets for the unseen instances which are based on the label sets of neighboring instances[30].

3 Results and Discussion

We trained different classification algorithms on the dataset of Facebook status updates containing 15995 status updates with nine labels and evaluated their performance. The classifiers used the BoW approach augmented with Word N-gram and Character N-gram feature extraction techniques with TF-IDF vectorization. Based on the data and the extent of pre-processing, the results are quite satisfactory. We trained a binary Relevance classifier for each of the nine labels as our baseline classifier. We report the performance of all the models using four evaluation metrics like accuracy Rate, Precision Rate, Recall Rate, and F1 Score using both Micro and Macro averaging. By using Micro-averaging, the average of all True Positives, True Negatives, False Positives, and False Negatives is calculated for each class. When Macro-averaging is performed, different sets of precision and recall are calculated to achieve the system's average accuracy and recall. In multi-label classification, instead of evaluating accuracy based solely on precision and recall metrics, predictions are made based on the labels falling into multiple classes rather than inaccurately guessing no labels at all.

Positive class predictions are measured in precision against all positive class predictions. Alternatively, it can be defined as the number of true positives divided by the number of false positives.

The precision is the ratio

$$\text{TP}/(\text{TP}+\text{FP}) \tag{1}$$

The Recall represents how successfully the fraction of the relevant instances are classified by the classifier. The Recall is calculated by dividing true positives by the sum of true positives and the number of false negatives.

The recall is the ratio

$$TP\,/\,(TP\,+\,FN) \tag{2}$$

Precision and recall determine the F1 Score, which is weighted average of both. Multilabel classification is determined by averaging the F1 scores of individual classes.

The F1 score is:

$$F1\,=\,2*(precision*recall)\,/\,(precision\,+\,recall) \tag{3}$$

The accuracy metric measures the ratio of correct predictions over the total number of instances evaluated.

$$Accuracy =\,(TP+TN)/(TP+TN+FP+FN) \tag{4}$$

The results of the experiments conducted using different classifiers are compared as shown in [Table 1]. Overall, we obtain good accuracy values for every model. We observed that char n-gram TF-IDF vectors surpass the word n-grams. We chose the Classifier Chains model above the remaining models because it has the highest test set accuracy, which is near to the training set accuracy.

Table 1. Performance Evaluation of Classification Models

Model	Features	Averaging	Precision	Recall	F1-Score	Accuracy (Training Set)	Accuracy (Test Set)
One Vs Rest	Word ngram(2,6)	micro avg	0.96	0.87	0.91	0.82	0.8
	Character ngram(2,6)	micro avg	0.98	0.92	0.95	0.88	0.85
	Word ngram(2,6)	macro avg	0.76	0.34	0.38	0.82	0.8
	Character-ngram(2,6)	macro avg	0.96	0.55	0.63	0.88	0.85
Binary Relevance	Word ngram(2,6)	micro avg	0.46	0.92	0.61	0.36	0.28
	Character ngram(2,6)	micro avg	0.75	0.96	0.84	0.72	0.7
	Word ngram(2,6)	macro avg	0.35	0.76	0.38	0.36	0.28
	Character ngram(2,6)	macro avg	0.42	0.74	0.49	0.72	0.7
Classifier Chains	Word ngram(2,6)	micro avg	0.95	0.87	0.91	0.82	0.79
	Character-ngram(2,6)	micro avg	0.98	0.93	0.95	0.88	0.86
	Word-ngram(2,6)	macro avg	0.76	0.34	0.37	0.82	0.79
	Character-ngram(2,6)	macro avg	0.96	0.55	0.64	0.88	0.86
Label Powerset	Word-ngram(2,6)	micro avg	0.91	0.87	0.89	0.82	0.79
	Character-ngram(2,6)	micro avg	0.98	0.9	0.93	0.86	0.83
	Word-ngram(2,6)	macro avg	0.75	0.33	0.35	0.82	0.79
	Character-ngram(2,6)	macro avg	0.97	0.48	0.56	0.86	0.83
ML-kNN	Word-ngram(2,6)	micro avg	0.93	0.88	0.91	0.78	0.74
	Character-ngram(2,6)	micro avg	0.95	0.91	0.93	0.83	0.82
	Word-ngram(2,6)	macro avg	0.63	0.38	0.42	0.78	0.78
	Character-ngram(2,6)	macro avg	0.81	0.6	0.67	0.83	0.82

4 Conclusion

A corpus of English-Hindi code-mixed Facebook Status updates has been annotated with nine emotions like joy, trust, surprise, anticipation, sadness, disgust, anger, fear, and love, along with language tags indicating English, Hindi, and Mixed. Additionally, we compare different classification algorithms for classifying the emotion of status updates. The use of slang, abbreviations, and emoticons improves accuracy, as evidenced by experiments. The Char N-Gram feature vector has performed better than the Word N-Gram feature vector. Based on code-mixed social media messages in Hindi and English, this article describes the initial efforts to predict different emotions coexisting in a status update. For future work, word embeddings may be used as the feature vectors to classify emotional states using deep learning techniques.

References

1) Internet usage in India - statistics & facts. 2020. Available from: https://www.statista.com/topics/2157/internet-usage-in-india.
2) Teens, Social Media and Technology 2018. 2018. Available from: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018.
3) Leading countries based on Facebook audience size as of January 2021 (in millions). 2021. Available from: https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users.Dateaccessed:05/06/2021.
4) Facebook by the Numbers: Stats, Demographics & Fun Facts. . Available from: https://www.omnicoreagency.com/facebook-statistics.
5) Mohammad SM. Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text. *Emotion Measurement*. 2016;p. 201–237. Available from: https://doi.org/10.1016/B978-0-08-100508-8.00009-6.
6) Koustav R, Shruti R, Rafiya B, Kalika B, Monojit C, Niloy G. Understanding Language Preference for Expression of Opinion and Sentiment: What do Hindi English Speakers do on Twitter. *Proceedings of EMNLP*. 2016;p. 1131–1141. Available from: https://www.aclweb.org/anthology/D16-1121.pdf.
7) Clercq OD, Balahur A, Sedoc J, Barriere V, Tafreshi S, Buechel S. Shared Task: Predicting Empathy and Emotion in Reaction to News Stories. *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis Proceedings of the 11th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. 2021;p. 92–104. Available from: https://aclanthology.org/2021.wassa-1.10.pdf.
8) Wadhawan A, Aggarwal A. Towards Emotion Recognition in Hindi-English Code-Mixed Data: A Transformer Based Approach. *arXiv preprint* . 2021. Available from: https://arxiv.org/pdf/2102.09943v2.pdf.
9) Deepanshu V, Aditya B, Vinay S, Sarfaraz AS, Manish S. Corpus Creation and Emotion Prediction for Hindi-English Code-Mixed Social Media Text. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop. Akhtar Syed Sarfaraz. 2018;p. 128–135. Available from: https://doi.org/10.18653/v1/N18-4018.

10) Robert P, Kellerman H. A general Psychoevolutionary Theory. In: Theories of Emotion. Academic Press. 1980;p. 3–33. Available from: https://doi.org/10.1016/B978-0-12-558701-3.50007-7.
11) Shashank S, Srinivas PYKL, Chandra BR. Sentiment analysis of code - Mix script. *International Conference on Computing and Network Communications*. 2015;p. 530–534. doi:10.1109/CoCoNet.2015.7411238.
12) Shashank S, Srinivas PYKL, Chandra BR. Text normalization of code mix and sentiment analysis. *International Conference on Advances in Computing, Communications, and Informatics (ICACCI)*. 2015;p. 1468–1473. doi:10.1109/ICACCI.2015.7275819.
13) Jamatia A, Gamback B, Das A. Part-of-Speech Tagging for Code-Mixed English-Hindi Twitter and Facebook Chat Messages. *Proceedings of Recent Advances in Natural Language Processing*. 2015;p. 239–248. Available from: https://1library.net/document/q01g3lxz-speech-tagging-mixed-english-hindi-twitter-facebook-messages.html.
14) Raghavi KC, Chinnakotla MK, Shrivastava M. "Answer ka type kya he?": Learning to Classify Questions in Code-Mixed Language. *Proceedings of the 24th International Conference on World Wide Web*. 2015;p. 853–858. Available from: https://doi.org/10.1145/2740908.2743006.
15) Utsab B, Amitava D, Wagner J, Jennifer F. Code Mixing: A Challenge for Language Identification in the Language of Social Media. *Proceedings of 1st Workshop on Computational Approaches to Code Switching*. 2014;p. 13–23. Available from: https://www.aclweb.org/anthology/W14-3902.pdf.
16) Utsab B, Joachim W, Grzegorz C, Jennifer F. DCU-UVT: Word-Level Language Classification with Code-Mixed Data. *Proceedings of The First Workshop on Computational Approaches to Code Switching*. 2014;p. 127–132. Available from: https://aclanthology.org/W14-3915.pdf.
17) Amitava D, Björn G. Identifying Languages at the Word Level in Code-Mixed Indian social media Text. *Proceedings of 11th Int Conf Nat Lang Process Goa*. 2014;p. 169–178. Available from: https://www.aclweb.org/anthology/W14-5152.pdf.
18) Yogarshi V, Spandana G, Jatin S, Kalika B, Monojit C. POS Tagging of English-Hindi Code-Mixed Social Media Content. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2014;p. 974–979. Available from: https://www.aclweb.org/anthology/D14-1105.pdf.
19) Souvick G, Satanu G, Dipankar D. Part-of-speech Tagging of Code-Mixed Social Media Text. *Proceedings of 2nd Workshop on Computational Approaches to Code Switching*. 2016;p. 90–97. Available from: https://www.aclweb.org/anthology/W16-5811.pdf.
20) Dinkar S, Savitha M, Debraj R, Devansh S, Kashyap D. Sentiment analysis of mixed language employing Hindi-English code switching. *International Conference on Machine Learning and Cybernetics (ICMLC)*. 2015;p. 271–276. doi:10.1109/ICMLC.2015.7340934.
21) Rupal B, Yashvardhan S, Shubham S. Sentiment Analysis for Mixed Script Indic Sentences. *International Conference on Advances in Computing*. 2016;p. 524–529. doi:10.1109/ICACCI.2016.7732099.
22) Abdul-Mageed M, Mona D, Sandra K. SAMAR: Subjectivity and sentiment analysis for Arabic social media. *Computer Speech and Language*. 2014;28(1):20–37. doi:10.1016/j.csl.2013.03.001.
23) Das A, Bandyopadhyay S. Subjectivity Detection in English, and Bengali: A CRF-based Approach. In: Proceedings of 7th International Conference on Natural Language Processing Macmillan Publishers. 2009. Available from: http://amitavadas.com/Pub/ICON_Subjectivity_2009.pdf.
24) Grigorios T, Ioannis K, Ioannis V. Multi-label Classification: An Overview. *International Journal of Data Warehousing and Mining* . 2009;3:1–13. doi:10.4018/jdwm.2007070101.
25) Read J. Scalable Multi-label Classification. Hamilton, New Zealand. 2010. Available from: https://researchcommons.waikato.ac.nz/bitstream/handle/10289/4645/thesis.pdf.
26) Saif M, Peter T, D. Crowdsourcing a Word-Emotion Association Lexicon. *Computational Intelligence*. 2013;29(3):436–465.
27) Jesse R, Bernhard P, Geoff H, Eibe F. Classifier Chains for Multi-label Classification. *Machine Learning*. 2011;85(333). Available from: https://doi.org/10.1007/s10994-011-5256-5.
28) Dominik H, Robin S, Weiwei C, Eyke H. Multilabel classification for exploiting cross-resistance information in HIV-1 drug resistance prediction. *Bioinformatics*. 2013;29(16):1946–1952. Available from: https://doi.org/10.1093/bioinformatics/btt331.
29) Newton S, Alvares CE, Carolina MM, Diana LH. A Comparison of Multi-label Feature Selection Methods using the Problem Transformation Approach. *Electronic Notes in Theoretical Computer Science*. 2013;292(292):135–150. doi:10.1016/j.entcs.2013.02.010.
30) Min-Ling Z, Zhi-Hua Z. *ML-KNN: A lazy learning approach to multi-label learning Pattern Recognition*. 2007;40(7):2038–2048. doi:10.1016/j.patcog.2006.12.019.