Enhancing Hinglish Language Student Sentiment Analysis In Educational Domain Through Cross-Validated Deep Learning Models

Deepti Singh Kshatriya[1] and Dr. Snehlata Barde[2]
[1]Research Scholar, Department of MSIT, MATS University, Raipur (CG), India
[2]Professor, Department of MSIT, MATS University, Raipur (CG), India

Abstract

Sentiment Analysis (SA) is a class of data mining algorithms that extracts latent representations of emotion from textual corpora. Student feedback is essential for assessing teachers' effectiveness and evaluating the quality of their instruction. Understanding and analyzing student feedback is vital for educational institutions to enhance campus experiences. However, when dealing with student feedback in Hinglish, a unique blend of Hindi and English, conventional approaches fall short. This article presents a comprehensive multi-step approach to sentiment analysis for Hinglish student feedback. We begin by collecting diverse datasets from postgraduate students across 17 different colleges and classify with machine learning and deep learning-based analysis model to improve the learning experience. In this study, CountVectorizer (CV) and Term Frequency-Inverse Document Frequency (TF-IDF) models are employed for feature extraction and classification models Logistic Regression, Keras Sequential Deep Learning, and GloVe word embedding with CNN Algorithms are used to analyze our model. The results show that GloVe word embedding with the CNN model is the best model with the highest accuracy, precision and f1-score.

Keywords:
Sentiment Analysis, Machine Learning, CNN, Logistic Regression

DOI: 10.48047/ecb/2022.11.11.77

1. INTRODUCTION

Sentiment analysis, also referred to as opinion mining, is a method of natural language processing that involves the automatic detection, extraction, and quantification of subjective data from text data using computational algorithms. In a text, such as a tweet, review, or article, it is used to ascertain the emotional content, attitude, or opinion expressed. Sentiment analysis can be carried out at various levels, such as the document, sentence, or aspect levels, and it has a wide range of uses, including product monitoring, students' emotion identification, market analysis, client feedback evaluation, and political campaign evaluation.

Students' sentiments and opinions are a rich source of data that can be used to enhance institutions and policies as well as analyze how students behave in relation to a course, a topic, or instructors[1]. Opinion mining and sentiment analysis may both appear to be identical, but there is a small distinction between the two: the other refers to extracting and analyzing people's opinions for a specific entity while the former relates to discovering sentiment phrases and words expressing emotions.

Student feedback serves as a valuable resource for educational institutions striving to improve their campuses. However, the analysis of student feedback in Hinglish—a language that seamlessly blends elements from both Hindi and English—presents unique challenges. In this article, we present a comprehensive multi-step approach to sentiment analysis for Hinglish student feedback. Our methodology encompasses data collection, lexicon creation, feature extraction, sentiment analysis, and real-world sentiment prediction, offering actionable insights for colleges and universities .Using natural language processing (NLP), sentiment analysis from word-based, sentence-level, or document-level corpora is performed. Unfortunately, manual processing of feelings is problematic due to the high amount of document[2]. Data processing must therefore be automated.

The classification of data, such as feedback reviews and automatic suggestions based on past user experience, could be resolved via opinion mining. The effectiveness of teaching and the performance of instructors must be evaluated through the utilization of student feedback[3]. When compared to conventional rating systems, free-form text comments generated from open-ended inquiries are rarely thoroughly examined. The current study used document level sentiment analysis, machine learning, and deep learning to automatically determine the emotion polarity of text-based feedback in order to address this issue.

In many application fields, particularly in business and social networks, sentiment analysis has been used for a variety of purposes. Product and service reviews are some well-known applications of sentiment analysis in business[4], capital markets [5],management of client relationships [6], and business strategies and studies [5], among others.

2. RELATED WORK

Sentiment analysis (SA) is a variant of natural language processing (NLP) that uses text analysis as well as associated technologies to categories subjective content into groups of opinions, emotions, or any other type of category. The problem of vector representation of text is crucial to sentiment analysis since it affects the efficacy and accuracy of the created SA models[7].

At its core, Hinglish embodies the cultural amalgamation that defines India. It seamlessly blends elements of Hindi, one of India's most widely spoken languages, with English, the global lingua franca. The result is a language that effortlessly combines the rich, melodic cadences of Hindi with the succinct, universally recognized vocabulary of English[8].

Vivek [9] Has used bi-directional sequence models for Mathur et al.[10] Hinglish Offensive Tweet (HOT) and he found that For effective word representations of Hinglish text, embedding learned from the Hinglish data set may be insufficient and require additional training.

Feedback from students is gathered in a variety of ways. Clickers, mobile phones, and social networking sites like Facebook and Twitter are among them. Prabha et al., used different Deep Learning model, LSTM ,Multi-head attention,LSTM with attention layer and Fusion (Multi-head attention+Embedding+LSTM) model for Vietnamese students feedback .They Concluded that when compared to the other three versions, the Fusion model is the best[11].

Lazrig and Humpherys [12] proposed Nine machine learning algorithms for five Experimental Datasets (student feedback, dataset without neutral, Movies ,Airlines and Airlines without Neutral). They observed that 98% accuracy was obtained with naive Bayes when neutral sentiment was excluded.

Deep Learning is a multi-layered neural network subclass. A neural network can extract high-level features because of these many layers. Words need to be converted into vectors before establishing a Deep-learning model for natural language processing. This method of grouping words with similar significance into close positions within the same frequency of a vector space is known as "word embedding"[13]. The "Word2Vec" model is one of the ones used for word embedding. Puspa and Seneewong Na Ayutthaya [14] Compare several conventional deep learning models (CNN,LSTM and Bidirectional LSTM) for sentiment analysis of Thai Children stories. Finally they concluded that The CNN model performed the best for sentiment classification in this dataset.

Ren et al.,[3] proposed aspect-level sentiment evaluation with deep learning and dictionary-based techniques to calculate student evaluations of teaching.They used Attention layer, Bi-LSTM layer, Average pooling layer and Word embedding layer for calculate emotion orientation automatically. Finally they concluded that aspect-level sentiment analysis determined by the topic dictionary, with word embedding, average pooling, Bi-LSTM, attention, and output layers is the best result for calculating student evaluations of teaching.

3. PROPOSED METHODOLOGY

The suggested system's implementation schematic is depicted in Figure 1. The modules for this proposed effort include data collection, preprocessing, feature extraction, and sentiment approaches.

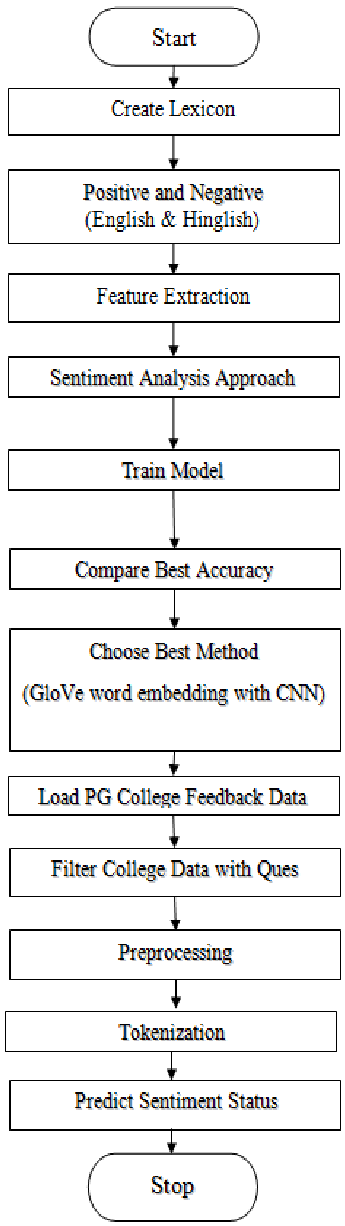


Fig 1. Implementation of Hinglish Feedback Sentiment Analysis

3.1 DATA COLLECTION

3.1.1 Real-Word Hinglish Student Feedback:

There are several sources for gathering data and divided them into three distinct categories depending on their features. These groups include Surveys/Questionnaires, Social media and blogs and Platforms for research and education (Coursera, ResearchGate, edX, Kaggle, LinkedIn etc). In figure 2 show, nearly 3000 post-graduate students from different colleges participated by filling out questionnaires. In order to ensure the quality of their Campus and enhance the learning experience for everyone involved, it is crucial that we collect feedback in Hinglish language student at the postgraduate level. Institutions may improve the learning experience and make students and teachers feel valued and respected by developing a thorough feedback procedure.

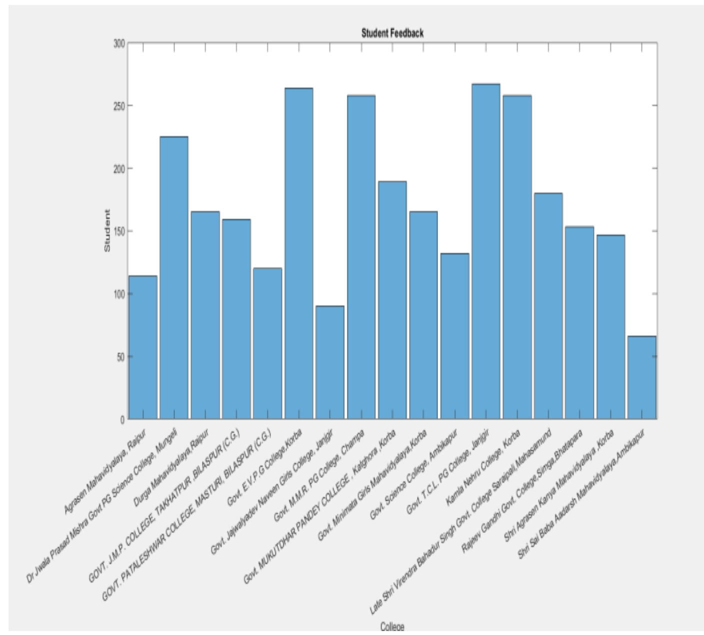


Fig 2. Total number of College Student Feedback

3.1.2 Comprehensive Lexicon Generation

Hinglish is a linguistically complex language, drawing inspiration from both Hindi and English. It incorporates slang, colloquialisms, and non-standard spellings, resulting in a wide range of word variations. The absence of standardized grammar rules and the dynamic nature of the language pose significant challenges for lexicon creation. We leverage Kaggle, a reputable data source, to download Hindi and English lexicons. These lexicons serve as the foundation for our Hinglish lexicon.

3.2 FEATURE EXTRACTION

Natural Language Processing (NLP) approaches enhance text analysis by offering some basic tools for automated sentiment analysis. The statistical and neural methods subcategories of NLP techniques are the most prevalent[15]. In this research, CountVectorizer and TF-IDF Vectorizer were used to analyze the dataset and determine the sentiment of feedback. The CountVectorizer counts the number of times each word occurs in a document.

Furthermore, the TF-IDF Vectorizer, which was used in the experiment, is a way of determining data relevant to words.

3.2.1 *CountVectorizer:*

The countervectorizer technique employs a bag of words that only processes data from the quantity of words and ignores text structure[16]. Every phrase in every textual piece is converted into a vector. By transforming the string representation into a numeric array, this technique [17].The amount of distinct words in a document serves as the vector's input, and each word will be given an index.

3.2.2 *Tf-Idf:*

The Term Frequency-Inverse Document Frequency (TF-IDF) Vectorizer technique is a form of statistical analysis that assesses the significance of a phrase in information retrieval of information. Two components of the frequency can be separated: first, the term frequency (TF). TF stands for the number of times a word occurs in a document, and Inverse document frequency (IDF) stands for the inversed value of the number of times the word occurs in all documents[18]. Take this phrase from a dataset of student feedback as an example. "The campus is excellent." In this case, "excellent" will be more significant than "The" or "campus." The term "campus" may appear in as much feedback as possible in the documents. As a result, word documents will be used frequently. The corresponding number when TF-IDF is calculated will be small. The only word that will be accepted as input is one with a higher TF-IDF number.

3.3 SENTIMENT ANALYSIS APPROACHES

3.3.1 *Logistic Regression with Lasso:*

Logistic Regression is a method of classification that solves the binary classification issue. In models that include a double situation, the outcome is typically described as 0 or 1. The function considered in logistic regression is

$$p(x) = \frac{exp^{\beta_o+\beta_1 x}}{1+exp^{\beta_o+\beta_1 x}} \quad (1)$$

Here in the Eq.(1) value of p(x) lies between 0 and 1 and x fluctuates between $-\infty$ to ∞. To utilize the above function for sentiment analysis .However, in most cases, the x values that make sense are constrained, as seen in the illustration in Fig. 3.we must first extract appropriate information from the input content, which can include word frequencies or the existence or absence of specific words or phrases. The logistic regression function is then computed using these features to determine the likelihood of positive sentiment. We can categorize the sentiment as positive if the probability is higher than a specific threshold.

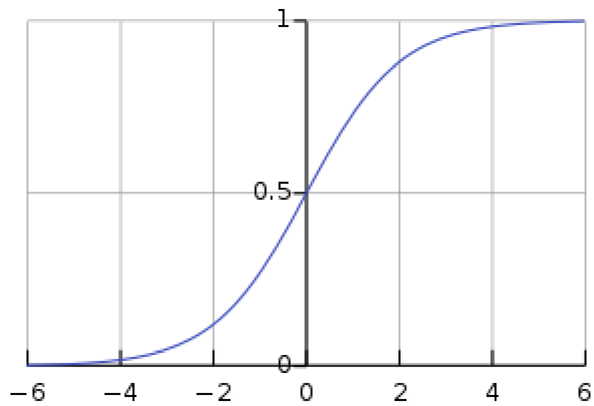


Fig.3. Logistic Function

We settled on LASSO regression, a development of Logistic Regression. When overfitting happens, which can be minimized through regularization, LASSO regression is advantageous to use. The goal of regularization is to minimize the loss function, so the lower the coefficient, the better. For LASSO regression, the coefficients may eventually be reduced to zero in order to enhance model efficiency and lower variance.

3.3.2 *Convolutional Neural Network:*

The CNN features powerful pooling layers and provides a standard architecture for translating variable-length words and phrases of fxed length distributed vectors[19].Fig.4. represents an example of a CNN model. The model is a clarified convolutional network based on CNN for phrase classification. However, we just use one channel to keep things simple. First, we must convert the phrase into a matrix, where each row represents a word vector[20]. The word vectors have a dimensionality of d, while the sentence matrix has a dimensionality of s×d when a sentence has a length of s.

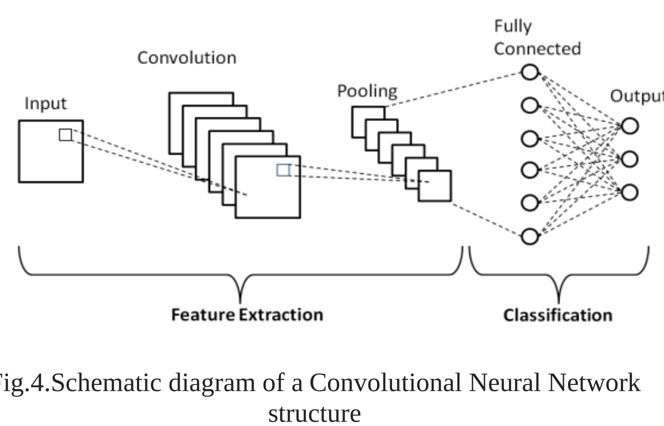


Fig.4.Schematic diagram of a Convolutional Neural Network structure

4. RESULT AND DISCUSSION

In the first experiment, the Count Vectorizer approach is used to extract features, and logistic regression with the l1 penalty and the C argument set to 0.4 is used to classify the data. The goal for getting these scores is to examine how well the model generalizes to new, previously unseen data (i.e., the test data) after being trained on the training data. Overfitting may be present if the training score is significantly greater than the test score. The accuracy of training dataset 90.31% and test accuracy is 88.23%

In second experiment text classification pipeline using the TF-IDF vectorization technique and a logistic regression model in scikit-learn.The Logistic Regression class from scikit-learn is then used in the code to create a logistic regression model . The value of the class_weight argument is 'balanced', which modifies the class weights to take into account unbalanced data. Lasso regularization is applied to the model by setting the penalty argument to "l1," and the C argument is set to "0.4," which regulates the regularization's strength. The cross_validate scikit-learn tool is then used to analyze the pipeline's performance using k-fold cross-validation. The 'accuracy' scoring argument measures the percentage of successfully identified cases. For each fold of the cross-validation, the function delivers the training and testing accuracy scores, which are reported to the console. The accuracy of training dataset 86.77% and test accuracy is 85.33%

In third experiment Simple Keras Sequential Deep learning model with binary feature is used, each feature map is subjected to a pooling function to produce a fixed-length vector, which is then extracted from using 1-max pooling. We'll utilize the Sequential API in this experiment because it lets you build models layer by layer. We will then include the layers that are input, hidden, and output because layers are the cornerstone of deep learning models. Between them, we'll utilize Dropout and dense on each layer to avoid overfitting. The accuracy of training dataset 94% and test accuracy is 88.38%.

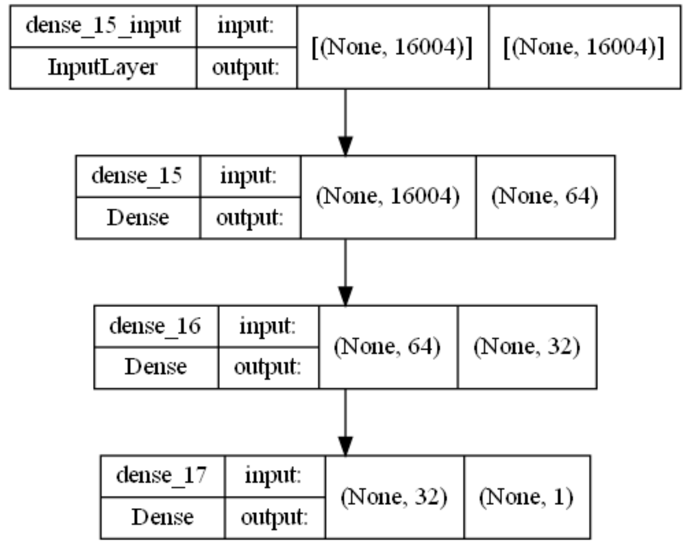
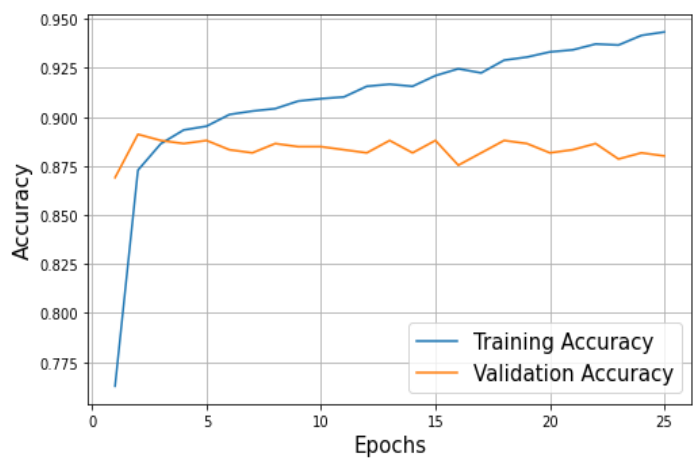


Fig.5. Dense layer architecture for text classification using



CNN

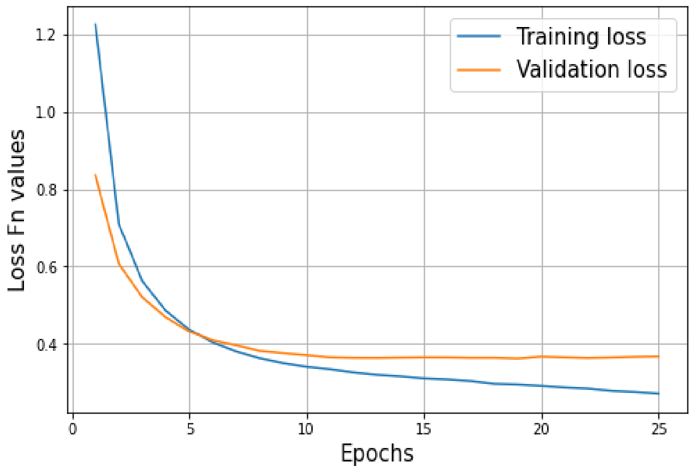
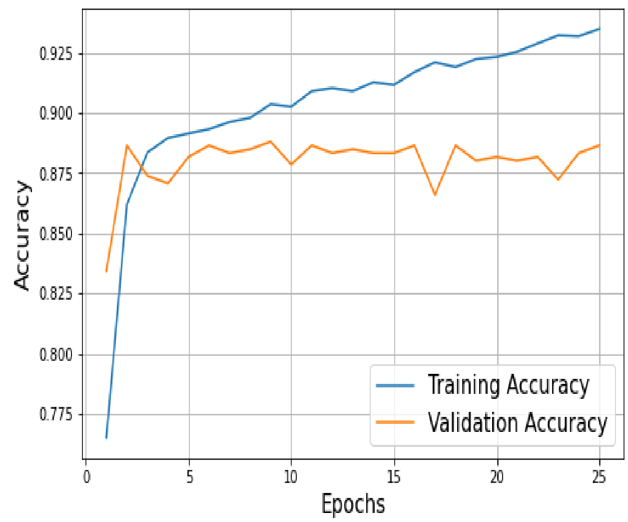


Fig.6.Training and validation accuracy and loss values for sentiment analysis task using Deep learning with sequential API

Figure 6 demonstrates that the training accuracy is greater than the validation accuracy. In this figure shown, both the training loss and the validation loss are decreasing and stabilize at a particular point. This denotes an ideal fit or a model that is neither overfitting nor underfitting.

In fourth experiment we employ an embedding is a type of learnt text representation in which words with similar meanings are represented similarly. GloVe, which stands for Global Vectors for Word Representation, is an acronym. It's an unsupervised learning system from Stanford that creates word embeddings from the global phrase co-occurrence matrix of a corpus[19]. The main goal of the GloVe embedding is to determine the relationship between the words using statistics. The models' accuracy is increased using the Embedding approach. a Convolutional Neural Network (CNN), which has excelled in solving challenges involving document classification. With a kernel size of 3 and a rectified linear (or "relu") activation function, a conservative CNN configuration is utilized. This configuration consists of 64 filters , which are used in parallel. The output of the convolutional layer is then cut in half by a pooling layer that follows. The 'features' retrieved by the CNN are then represented by flattening the 2D output from the CNN component of the model into a single lengthy 2D vector. Standard Multilayer Perceptron layers are used as the model's back end to interpret the CNN characteristics. An output value between 0 and 1 representing the review's positive and negative sentiment is produced by the output layer using a sigmoid activation algorithm. The network is then fitted using the training data. Because the task we are learning is a binary classification problem, we utilize a binary cross entropy loss function. We employ the effective Adam stochastic gradient descent implementation and monitor accuracy as well as loss during training. 25 epochs through the training set of data, are used to train the model. On the training dataset, we can get 96.06% accuracy. The model's accuracy on the test dataset at the end of the run is 91.5%.



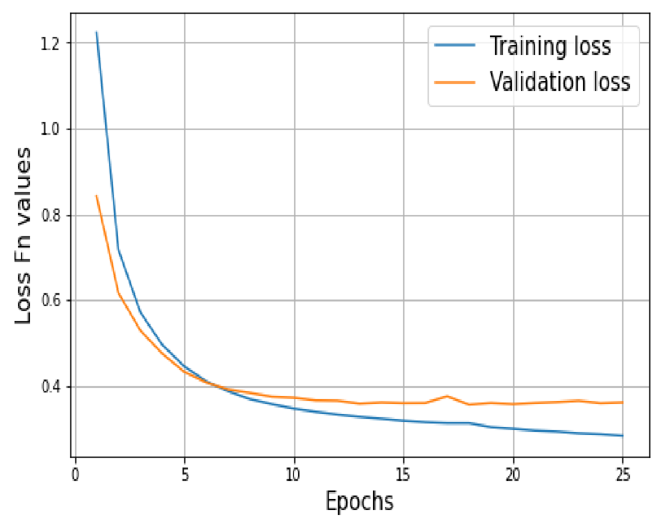


Fig.7.Training and validation accuracy and loss values for sentiment analysis task using GloVe word embedding+ CNN

According to Figure 7, the training accuracy is greater than the validation accuracy. This represents the best model. And the training and validation losses both reduce and stabilize at a certain point. This denotes an ideal fit or a model that is neither overfitting nor underfitting.

Table.1. Predict the performance of classification models on a test set using Accuracy measures.

Model	Accuracy		
	Negative	Neutral	Positive
CV+LR	88.32	91.89	90.72
TF-IDF+LR	87.49	87.52	85.03
Keras Sequential DL	94.39	93.09	94.52
GloVe+CNN	95.23	96.82	96.13

Table.2. Predict the performance of classification models on a test set using Precision measures.

Model	Precision		
	Negative	Neutral	Positive
CV+LR	87.38	91.39	93.75
TF-IDF+LR	85.69	89.34	84.93
Keras Sequential DL	92.09	93.59	93.25
GloVe+CNN	98.63	99.03	98.39

Table.3. Predict the performance of classification models on a test set using Recall measures.

Model	Recall		
	Negative	Neutral	Positive
CV+LR	65.55	54.67	68.56
TF-IDF+LR	70.83	69.34	75.03
Keras Sequential DL	90.19	86.82	79.25
GloVe+CNN	93.99	91.33	91.39

Table.4. Predict the performance of classification models on a test set using F1-score measures.

Model	F1-Score		
	Negative	Neutral	Positive
CV+LR	74.90	68.41	79.20
TF-IDF+LR	77.55	78.07	79.67
Keras Sequential DL	91.13	90.07	85.68
GloVe+CNN	96.25	95.02	94.76

The accuracy and precision are calculated for negative, neutral, and positive sentiments. These results are shown in Table 1, 2, 3 and 4, it is clear that our suggested GloVe with CNN model performs better in the evaluation measures. We Choose GloVe with CNN Model for predict sentiment analysis of PG students feedback.

4.1 Student Feedback Data Loading

Transitioning from model development to real-world application, we load Hinglish language PG student feedback specific to a particular domain from 17 different colleges. These feedbacks encapsulate the genuine experiences and opinions of students within that domain.

4.2 PREPROCESSING & Tokenization

Data preprocessing and Tokenization is an important stage in sentiment analysis since it can enhance the level of accuracy of the interpretation by reducing noise and making sure the data is in the right format for analysis.

4.2.1 Removing Punctuation:

During this stage, all punctuation from the text will be eliminated (‘!”#$%&'()*+,-./:;?@[\]^_`{|}~’).

4.2.2 *Removing Stopwords and Lower Casing:*

Stopwords are frequently occurring words that are removed from texts since they add nothing to the analysis. Lower casing is one of the most frequent preprocessing steps involves changing the text's case, ideally to lower case.

4.2.3 *Tokenization:*

At this point, the text has been broken into smaller sections. To fix our issue, we use word tokenization.

4.2.4 *Stemming/Lemmatization:*

In order to analyse the inflected or derived words as a single entity, it is necessary to reduce them to their base or root form.

4.3 Predict Sentiments: In figure 3 depicts the sentiment analysis results for students from diverse colleges, all expressing their sentiments in Hinglish. The figure categorizes these sentiments into positive, negative, and neutral classes

This visual representation's central component is a rich tapestry of student input. These are not just words, they are the voices of people navigating the complex world of education, each with their own unique viewpoint and story to tell. We learn important things about the emotional range of these children by this figure, which is illustrated by their response.

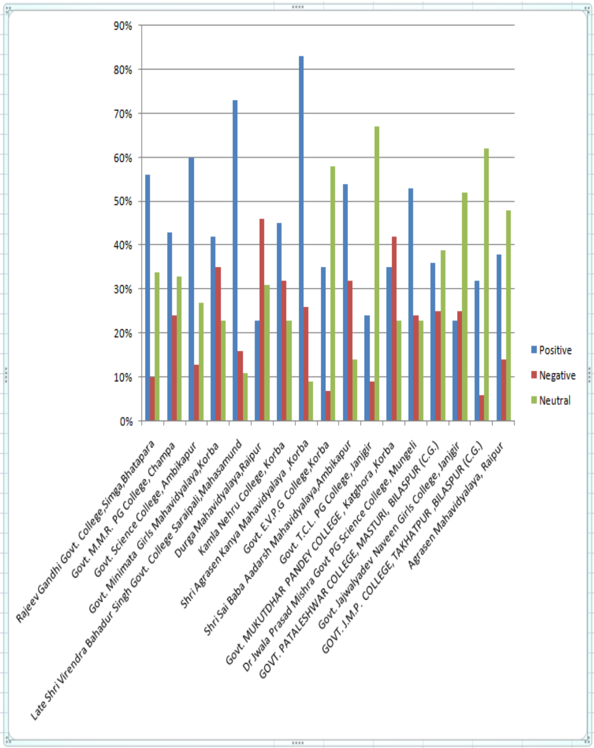


Fig 8. College Student Feedback Analysis

4.3.1 Positive Sentiments:

The positive emotion category, shown in Figure 8, highlights the elements of students' educational journeys that they value and celebrate. It captures the prevailing moments, the gratitude for committed teachers, the sense of achievement after conquering obstacles, and the pleasure experienced from fulfilling relationships with other members of the college community. These encouraging words serve as touching tales that serve as a reminder of the transformational impact of education.

4.3.2 Negative Sentiments:

The negative emotion category in Figure 8 reveals, in contrast, the difficulties and disappointments that students experience. It discusses the demands of academic life, the low points, and the occasions when expectations do not match reality. Even though they are open and even difficult to address, these unfavorable opinions are crucial signs of potential areas where the educational system needs to be improve

4.3.3 Negative Sentiments:

The neutral emotions, which fall in the middle of this feeling range and are neither too pleasant nor blatantly negative, are present. They stand for the middle ground, where students can convey factual data, observations, or personal experiences without evoking strong feelings. These impartial feelings give a well-rounded view of numerous facets of college life and add significant context to the overall sentiment landscape.

5. CONCLUSION

This article offers a comprehensive approach to sentiment analysis for Hinglish student feedback, addressing the linguistic intricacies of the language and providing actionable insights for educational institutions. Feedback from students is a precious tool for instructors because it provides ideas and guidance for enhancing a course. This work extends the use of student satisfaction feedback beyond constructing models utilizing deep learning and machine learning by extracting the useful data that is embedded in students' free text comments. In this study, we experimented with Count Vectorizer, TF-IDF and GloVe word embedding, sentiment features individually and in combinations to classify the sentiment of post-graduate students from different colleges using Logistic Regression And CNN, According to the accuracy(Cross Validation) outcomes, it appears that the fourth experiment model, which used CNN and the GloVe embedding, performed better than the all four experiment models, which used a straightforward deep learning model using binary features and Logistic Regression model .

Reference

[1] Z. Kastrati, F. Dalipi, A. S. Imran, K. P. Nuci, and M. A. Wani, “Sentiment analysis of students’ feedback with nlp and deep learning: A systematic mapping study,” *Appl. Sci.*, vol. 11, no. 9, 2021, doi: 10.3390/app11093986.

[2] D. S. Kshatriya and S. Barde, “Comparative Analysis of Machine Learning Approaches for Sentiment Analysis of Students ’ Online Learning Feedback During COVID -19,” vol. 11, no. 3, p. 2022, 2022, doi: DOI:10.26634/jit.11.3.19093.

[3] P. Ren, L. Yang, and F. Luo, “Automatic scoring of student feedback for teaching evaluation based on aspect-level sentiment analysis,” *Educ. Inf. Technol.*, vol. 28, no. July, pp. 797–814, 2023, doi: https://doi.org/10.1007/s10639-022-11151-z Automatic.

[4] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, “Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning,” *IEEE Access*, vol. 8, pp. 23522–23530, 2020, doi: 10.1109/ACCESS.2020.2969854.

[5] A. E. O. Carosia, G. P. Coelho, and A. E. A. Silva, “Analyzing the Brazilian Financial Market through Portuguese Sentiment Analysis in Social Media,”

Appl. Artif. Intell., vol. 34, no. 1, pp. 1–19, 2020, doi: 10.1080/08839514.2019.1673037.

[6] N. Capuano, L. Greco, P. Ritrovato, and M. Vento, “Sentiment analysis for customer relationship management: an incremental learning approach,” *Appl. Intell.*, vol. 51, no. 6, pp. 3339–3352, 2021, doi: 10.1007/s10489-020-01984-x.

[7] L. Rao, “Sentiment Analysis of English Text with Multilevel Features,” *Sci. Program.*, vol. 2022, 2022, doi: 10.1155/2022/7605125.

[8] D. Vijay, A. Bohra, V. Singh, S. S. Akhtar, and M. Shrivastava, “Corpus creation and emotion prediction for hindi-english code-mixed social media text,” *NAACL HLT 2018 - 2018 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Student Res. Work.*, vol. 2018-Janua, pp. 128–135, 2018, doi: 10.18653/v1/n18-4018.

[9] V. K. Gupta, “‘Hinglish’ Language -- Modeling a Messy Code-Mixed Language,” no. NeurIPS, 2019, [Online]. Available: http://arxiv.org/abs/1912.13109

[10] P. Mathur, R. Sawhney, M. Ayyar, and R. R. Shah, “Did you offend me? Classification of Offensive Tweets in Hinglish Language,” *2nd Work. Abus. Lang. Online - Proc. Work. co-located with EMNLP 2018*, pp. 138–148, 2018, doi: 10.18653/v1/w18-5118.

[11] K. S. D. Prabha, “Sentiment analysis of student feedback using multi - head attention fusion model of word and context embedding for LSTM,” *J. Ambient Intell. Humaniz. Comput.*, no. 4, 2020, doi: 10.1007/s12652-020-01791-9.

[12] I. Lazrig and S. L. Humpherys, “Using Machine Learning Sentiment Analysis to Evaluate Learning Impact,” *Inf. Syst. Educ. J.*, vol. 20, no. 1, p. 20, 2022, [Online]. Available: https://isedj.org/;https://iscap.info

[13] Y. Goldberg and O. Levy, “arXiv : 1402 . 3722v1 [cs . CL] 15 Feb 2014 word2vec Explained : Deriving Mikolov et al .’ s,” no. 2, pp. 1–5, 2014.

[14] K. Pasupa and T. Seneewong Na Ayutthaya, “Thai sentiment analysis with deep learning techniques: A comparative study based on word embedding, POS-tag, and sentic features,” *Sustain. Cities Soc.*, vol. 50, no. May, p. 101615, 2019, doi: 10.1016/j.scs.2019.101615.

[15] L. Tan, O. K. Tan, C. C. Sze, and W. W. Bin Goh, “Emotional Variance Analysis: A new sentiment analysis feature set for Artificial Intelligence and Machine Learning applications,” *PLoS One*, vol. 18, no. 1 January, pp. 1–22, 2023, doi: 10.1371/journal.pone.0274299.

[16] S. Chowdhury and M. P. Schoen, “Research Paper Classification using Supervised Machine Learning Techniques,” *2020 Intermt. Eng. Technol. Comput. IETC 2020*, 2020, doi: 10.1109/IETC47856.2020.9249211.

[17] A. Prasetya, N. Susetyo, F. Putri, and P. Widiharso, “Letter detection : an empirical comparative study of different ML classifier and feature extraction,” vol. 5, no. 1, pp. 1–7, 2023.

[18] S. Valentina and W. Songpan, “Features Extraction Based on Probability Weighting for Fake News Classification on Social Media,” *Curr. Appl. Sci. Technol.*, vol. 23, no. 2, 2023, doi: 10.55003/cast.2022.02.23.014.

[19] K. Shanmugavadivel, V. E. Sathishkumar, S. Raja, T. B. Lingaiah, S. Neelakandan, and M. Subramanian, “Deep learning based sentiment analysis and offensive language identification on multilingual code-mixed data,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–13, 2022, doi: 10.1038/s41598-022-26092-3.

[20] H. Peng, Z. Zhang, and H. Liu, “A Sentiment Analysis Method for Teaching Evaluation Texts Using Attention Mechanism Combined with CNN-BLSTM Model,” vol. 2022, 2022.

ABOUT THE AUTHORS

Deepti Singh Kshatriya is a Ph.D. Research Scholar(CS), Department of Computer Science from MATS University, Raipur, Chhattisgarh, India. She is working as an Assistant Professor, Department of Computer Science & IT, Kamla Nehru Mahavidyalaya, Korba, Chhattisgarh, India accreditated by NAAC and affiliated to Atal Bihari Vajpayee Vishwavidyalaya, Bilaspur, Chhattisgarh, India . She have 12yrs of Teaching Experience and managing projects in academic level and giving sessions on different subjects like Embedded System, Data Mining, Machine Learning, Operating System, Software Engineering, Oracle, Visual Basic, etc. and handling Projects.

Dr. Snehlata Barde is a Professor of Computer Science at MATS University, Raipur, Chhattisgarh, India where she has been a faculty member since 2018. Her key role was of researcher, academician and motivator. She holds Doctorate in Computer Science. Her research interest includes Digital image processing ,forensic science, pattern recognition ,Machine Learning etc. She has published several articles in peer-reviewed journals and also presented her research at several national and international conferences and other role to get involved to motivate students for their future career guidance.