CHUNKER BASED SENTIMENT ANALYSIS AND TENSE CLASSIFICATION FOR NEPALI TEXT

Archit Yajni, Ms. Sabu Lama Tamang

Department of Mathematics, Sikkim Manipal Institute of Technology, Rangpo, Sikkim

ABSTRACT

The article represents the Sentiment Analysis (SA) and Tense Classification using Skip gram model for the word to vector encoding on Nepali language. The experiment on SA for positive-negative classification is carried out in two ways. In the first experiment the vector representation of each sentence is generated by using Skip-gram model followed by the Multi-Layer Perceptron (MLP) classification and it is observed that the F1 score of 0.6486 is achieved for positive-negative classification with overall accuracy of 68%. Whereas in the second experiment the verb chunks are extracted using Nepali parser and carried out the similar experiment on the verb chunks. F1 scores of 0.6779 is observed for positive -negative classification with overall accuracy of 85%. Hence, Chunker based sentiment analysis is proven to be better than sentiment analysis using sentences.

This paper also proposes using a skip-gram model to identify the tenses of Nepali sentences and verbs. In the third experiment, the vector representation of each sentence is generated by using Skip-gram model followed by the Multi-Layer Perceptron (MLP)classification and it is observed that verb chunks had very low overall accuracy of 53%. In the fourth experiment, conducted for Tense Classification using Sentences resulted in improved efficiency with overall accuracy of 89%. Past tenses were identified and classified more accurately than other tenses. Hence, sentence based tense classification is proven to be better than verb Chunker based sentiment analysis.

KEYWORDS

Skip –gram model, MLP classification, Parser, Verb chunks, Tense classification, Sentiment Analysis.

1. INTRODUCTION

Sentiment Analysis is a well-known text classification technique in computer science that examines people's ideas and perspectives and categorises them according to their nature into various classifications [1]. However, the terms sentiment analysis and opinion mining are sometimes used synonymously because opinion mining is a tool that contextualises polarity ratings in terms of topics, facets, and objectives. Sentiment detection is the process of categorising a sentence or text into classifications that are neutral, positive, or negative (sometimes) depending on the polarity of the sentences [2]. There are 45 million native speakers of the Nepali language throughout the world, notably in Nepal, Bhutan, Myanmar, and various regions of India, including Sikkim, West Bengal (Darjeeling district), Uttaranchal, and Assam [3]. Despite the popularity, the Nepali language continues to remain understudied. There are yet more languages, like Nepali, which can be taken into consideration for the sentiment analysis assignment. Until now, numerous languages, including English, Chinese, Persian, and Arabic, have been used for the task. Due to the lack of a well-annotated corpus, SA in the Nepali language is not as straightforward as it might first appear. This research work is extended to tense classification and in future can be applied to another type of review, like social–media comments,

election reviews, movie reviews, book reviews, etc. Also, the proposed approach can also be applied to other native Indian languages.

The accurate identification and classification of verb tenses is crucial in natural language processing and computational linguistics for various applications like machine translation, sentiment analysis, information retrieval, and question answering systems. Traditionally, rulebased systems struggled to handle the complexities of natural language. However, neural networks have revolutionized natural language processing by offering robust, data-driven approaches. Traditionally, tense classification has relied on rule-based systems and linguistic heuristics, which often struggle to handle the complexities and ambiguities of natural language. However, in recent years, the advent of deep learning techniques, particularly neural networks, has revolutionized the field of natural language processing by offering more robust and data-driven approaches to tackle complex language understanding tasks. In particular, the application of neural networks to tense classification has shown great promise in achieving state-of-the-art performance and overcoming the limitations of rule-based methods.

1.1. Contribution

This research aims to improve sentiment analysis by overcoming challenges in handling Nepali text data. With a vast amount of user-generated content in Nepali, Sentiment Analysis can help us in generating valuable information. This corpus was collected from Kaggle and sentiment analysis. The approach can be extended to various domains as in tense classification, social media comments, election reviews, movie reviews, and book reviews, and can be applied to other native Indian languages in the future.

This research paper also explores tense classification using skip gram model, focusing on core concepts, methodologies, and advancements. It provides an understanding of challenges and solutions proposed by skip gram-based models. The paper contributes to the ongoing discourse on artificial intelligence's role in linguistics and natural language understanding by reviewing the current state of the art and presenting empirical results.

1.2. Skip- Gram Model

TheSkip-gram model is a method for vector representation of words from a huge amount of unstructured text data, which was introduced by Mikolov et al. [4]. This model tries to maximize word classification which is based on another word in the same sentence.

Furthermore, this method predicts the context words using the main word and predict words within a certain range before and after the current word. Suppose a text is composed by a sequence of words w_o, w_1, w_2, w_3…... w_N. Then for word w, context of $w$is given by its left and right neighbourhood. Then to each word w a vector representation v is assigned, and the probability that w_0 is in the context of $w_i$is defined as the SoftMax of their vector product.

$$p(w_o|w_i) = \frac{\exp(v_{wi}.v_{wo}^T)}{\sum_{w=1}^{V}\exp(v_{wi}.v_w^T)} \qquad \text{...} (1)$$

Where, v_{wi} = input word vector
v_{wo}=output word vector
V =number of words in the vocabulary

SoftMax function also referred as Softargmax function or multi-class logistic regression is a function that converts a vector of *'k'* real number into a vector of' '*k*' real values that add upto 1. The input values could be positive, negative, zero, or even greater than one, but the SoftMax transforms them into values between 0 and 1, which makes it possible for them to be interpreted as probabilities.

The SoftMax function is given by.

$$s(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \qquad \text{.....(2)}$$

Where,
$x_{i=}$ input vector
$x_{j=}$ output vector
$e^{xi=}$ standard exponential function for input vector
$e^{xj=}$ standard exponential function for output vector

SoftMax function is a generalization of logistic regression that can be employed to carry out multiclass classification, and its graph very closely resembles that of the Sigmoid function. A sigmoid function is a mathematical function having a characteristic 'S-shaped' curve or sigmoid curve.

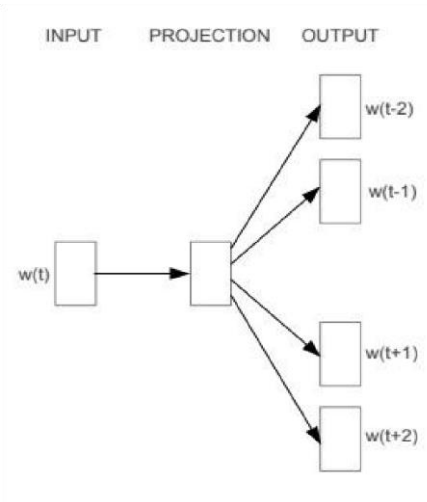


Fig. i Skip- gram model

In this work, Skip gram model (see Fig. i) is employed because it does not involve dense matrix multiplications, resulting in extremely efficient training. An optimized single-machine implementation can train more than 100 billion words in a day [5].

The main objective of skip-gram model is to predict the context of central words. So therefore, training the model means maximizing the objective function. If given a sequence of training words w_1, w_2, w_3……,w_N. By invoking a Naïve Bayes assumption of conditional independence, the probability of each center word is given by

$$P(w_{i-t}, .. w_{i-1}, w_{i+1}, ... w_{i+t}|w_i) = \sum_{t=1}^{N} \sum_{-a\leq i\leq a, i\neq 0} P(w_{i+t}|w_i) \quad \text{.....(3)}$$

Hence, the **objective function** is given by.

$$\frac{1}{N}\sum_{t=1}^{N} \sum_{-a\leq i\leq a, i\neq 0} \log P(w_{i+t}|w_i) \quad \text{.....(4)}$$

Where,
a =size of the training context and w_i is the central word.
w_{i-a}……, w_{i-1}, w_i, w_{i+1}……..., w_{i+a}

Higher accuracy can be achieved by using larger values of because they generate more training instances, but doing so also adds to the computational complexity. This is equivalent to minimising the cross-entropy average over the corpus, which appears to be the loss function.

Loss-function,

$$E = -\frac{1}{}\sum_{t=1}^{N} \sum_{-a\leq i\leq a, i\neq 0} \log P_N (w_{i+t}\backslash w_i) \quad \text{.....(5)}$$

Cross – Entropy Loss function is also called logarithmic loss, log loss or logistic loss. The predicted class probability is compared to the actual desired output, and a logarithmic penalty is calculated. A large score is given for significant differences close to 1, and a small score for minor differences tending to 0. [6] The aim is to minimize the loss, i.e., the smaller the loss the better the model. A perfect model has a cross-entropy loss of 0. [6]

1.3. Verb Chunker

Parsing is a crucial step in Natural Language Processing, extracting meaning from text by identifying speech parts, phrases, and clauses. It is essential for machine translation and preserving semantic and syntactic knowledge of a natural language [7]. The task of parsing is firstly initiated by POS tagging [8] and the taggers used reduces the ambiguity of the parser's input sentence which results in less ambiguous results. [9]. Particularly, Nepali POS tagging [8] has many applications such as it gives information about the word and its neighbouring words which can be further useful for higher level NLP tasks such as semantic analysis, machine translation and so on. [1] . In the given sentence भाईलाईखेल्नलगाइयो, the verb chunker is खेल्नलगाइयो.

2. STATE OF THE ART TECHNIQUES IN SENTIMENT ANALYSIS FOR NEPALI TEXT

A. Machine learning algorithmsfor sentiment classification SVM and Naive Bayes are promising sentiment classification algorithms, but face challenges in Nepali language adaptation due to language characteristics and limited datasets.

B. Lexicon-based approaches for sentiment analysis Develop Nepali sentiment lexicons, use machine learning for accurate analysis, explore rule-based methods for sentiment classification, and integrate lexicon-based and machine learning approaches for improved accuracy and efficiency [12].

C. **Deep learning methods** for sentiment analysis Convolutional Neural Networks revolutionize sentiment analysis by capturing local patterns and dependencies. deep learning models have been shown better than conventional machine learning models and lexicon-based approach in sentiment analysis [12].

The classification of verb tenses in natural language text is a crucial task in natural language processing (NLP) and computational linguistics. Researchers have explored various machine learning algorithms and methods to improve our understanding of temporal aspects within language.

1. **Early Approaches and Rule-Based Systems:** In the early stages of research, rule-based systems dominated the landscape of tense classification. These approaches relied heavily on linguistic rules and heuristics, often struggling to adapt to the complexities and nuances of natural language. While they provided some initial insights, they proved to be inadequate for handling large-scale, diverse corpora [13].

2. **Supervised Learning and Feature-Based Models:** A significant shift occurred with the adoption of supervised learning techniques. Researchers began to employ feature-based models, where engineered linguistic features such as part-of-speech tags and syntactic structures were used to train classifiers. This approach yielded improvements in performance, but it still faced limitations in capturing contextual information effectively.
3. **Statistical and Probabilistic Models:** The advent of statistical and probabilistic models, including Conditional Random Fields vs. Hidden Markov Models brought about significant progress in tense classification [14]. These models allowed for the integration of probabilistic reasoning and sequential dependencies in language, resulting in more accurate tense assignments.
4. **Deep Learning and Neural Networks:** The most revolutionary leap in tense classification research came with the application of deep learning and neural network-based models. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and more recently, Transformers, have demonstrated state-of-the-art performance in numerous NLP tasks, including tense classification. These models can automatically learn complex temporal dependencies and contextual information from large corpora, significantly advancing the field's capabilities.
5. **Design Challenges for Low-resource Cross-lingual Entity Linking:** Researchers have also explored tense classification in low-resource languages and cross-lingual scenarios[15]. Adapting machine learning models to languages with limited digital resources presents unique challenges related to data scarcity and linguistic diversity. Innovative transfer learning techniques and cross-lingual models have shown promise in addressing these challenges.

3. NEPALI DATASET

In this research the dataset available in kaggle [11] for Nepali language Sentiment Analysis has been used to perform the experiment. A total of 4,200 sentences were used during the experiment. Out of which 30% (1260) were test sentences and 70% (2940) were training sentences for the sentiment analysis.

Also, for tense classification nearly 48.5% (2038) were test sentences and 51.5% (2162) were training sentences.

We have used following structure for the classification using MLP.

For Sentiment Analysis (SA)
Input layer:150 Neurons
Hidden Layer1:700 Neurons
Hidden Layer 2:100 Neurons
Hidden Layer 3:10 Neurons
Output layer 1Neuron

For tense classification
Input Neurons: 150
Hidden Layer 1: 1000
Hidden Layer 2: 600
Hidden Layer 3: 100
Hidden Layer 4: 20
Hidden Layer 5: 10
Output Layer: 1 Neuron

A few significant challenges associated with Nepali languages are discussed below.

a. A slight variation of words in Nepali language text can influence the polarity of the word. For example, Line 1) भाईखेल्दैछ ("brother is playing") and Line 2) भाईखेल्दैन ("brother does not play")

The Line 1) refers to positive sentence and Line 2) refers to negative sentence.

b. There is absence of well-annotated corpus which makes the Sentiment Analysis task in the Nepali Language a challenging one.

4. METHODOLOGY

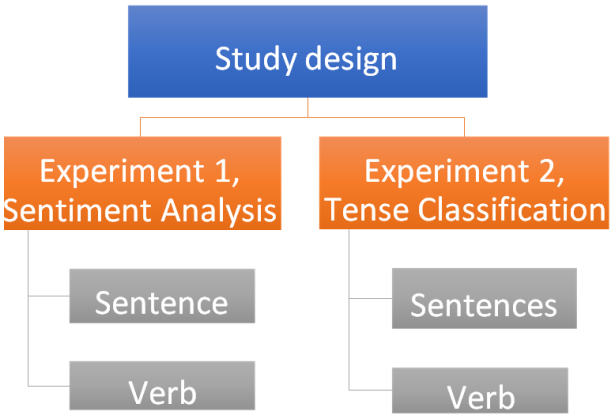


Figure: ii Study design

Experiment-I–Sentiment Analysis

Two experiments were conducted for sentiment analysis using skip gram model.

A. Using Sentences
B. Using Verb Chunks

Part A- Sentiment analysis for sentences:

Using the data from Kaggle [11] the sentences were encoded using skip-gram and MLP classifier. The resulting accuracy was noted. The errors arising out if it were recorded and classified into the following four types.

Type 1: Negative words are not available but still negative.
Type 2: Positive words are not available but still positive.
Type 3: Negative word is available but still positive.
Type 4: Interrogative sentences whose meaning is particularly challenging for the machine to understand.

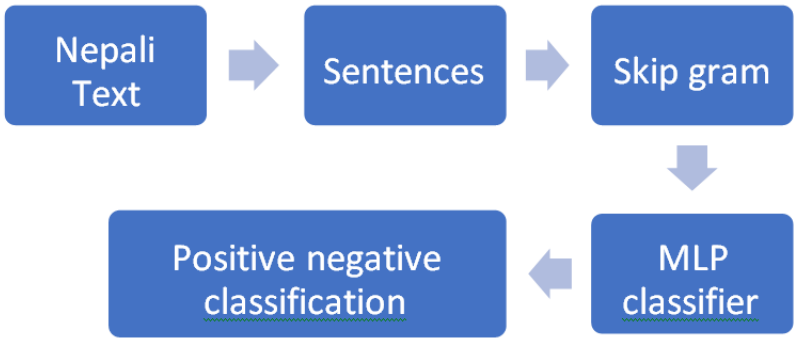


Fig. iii. Flow Diagram of the Methodology

Part B- Sentiment analysis for Verb Chunks

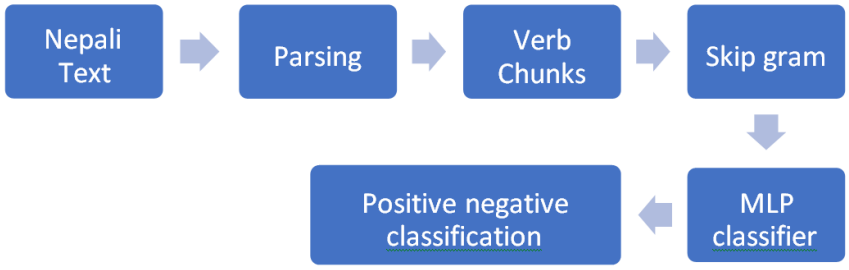


Fig. iv. Flow Diagram of the Methodology

As shown in fig. 2, the words are represented as vectors using the Skip-gram model [4], and the verb chunks and sentences are identified as positive or negative using an MLP classifier. In this section the methodology involved in tense classification has also been presented in the form of a flow chart (See Fig. 2).

Experiment-II–Tense Classification

Two experiments were carried out for tense classification, C.Using Verb Chunks
D.Using Sentences

Part C- Tense classification for Verb Chunks

In experiment C, using parser technique, Verb chunks were extracted, and the tense classification was done using Skip-gram model and MLP classifier.

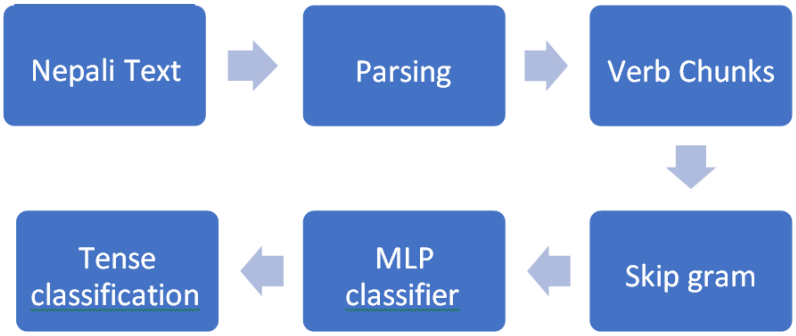


Figure v: Tense Classification for verb chunks

Part D- Tense classification for Sentences

In experiment D, Tense classification was done directly on Sentences using Skip-gram model and MLP classifier.

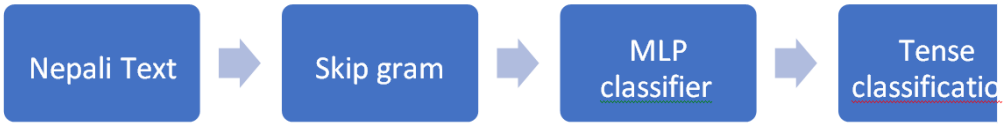


Figure vi:Tense classification for Sentences

5. RESULTS AND DISCUSSION

Experiment I A (Sentiment Analysis using sentences)

The results of tense Classification using Skip gram model on sentences was carried out in experiment I A, are shown in figure 1. Experiment I resulted in an accuracy of 68%. As mentioned in previous section Part- A, Methodology. However, there were a lot of errors along with it. The errors were analysed and were categorised into four types. Out of which Type 3 was, given more emphasis for the Experiment II.

However, some errors were encountered during the Experiment I, which were broadly classified into four types.

In Type 1 error: भाइगीतगाईरहेकोछ (“Brother is singing song”) is detected as negative sentence but negative words are not available.

In Type 2 error: घिनलाग्दोदानब (“Forlorn Demon”) is detected as positive sentence but positive words are not available.

In Type 3 error: रामबजारगएन ("Ram did not go to the market") is detected as positive sentence but negative word गएन ("did not go") is available.

Another example of Type 3 error: रामलाईबजारजानलगाइएन ("Ram wasn't made to go to the market") was detected earlier as positive sentence even if negative word (लगाइएन) ("wasn't made to go") is available.

In Type 4 error: the classification of positive and negative sentences depends on the meaning it forms, which is particularly challenging for the machine to understand. खानाखानेभएयस्तैगर्थे? ("Would have done this while eating the food") is wrongly detected as negative sentence and आखखरदेशघिकासकोबाधककोरहेछत? ("Afterall who is the obstacle to the country's development") is wrongly detected as positive sentence.

Sentiment Analysis was done on sentences using skip gram model and the resulting confusion matrix is shown in the figure below.

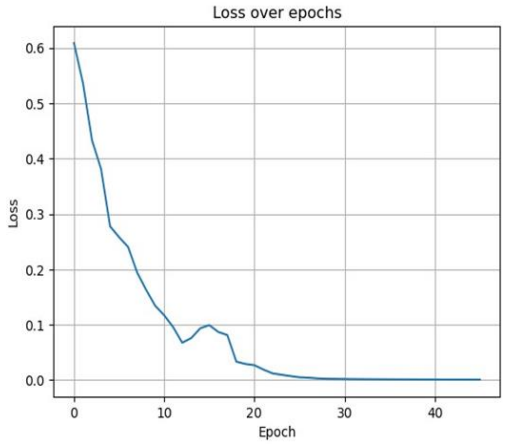


Figure 1: Loss over epoch,
Train Accuracy: 1.0
Test Accuracy: 0.6486

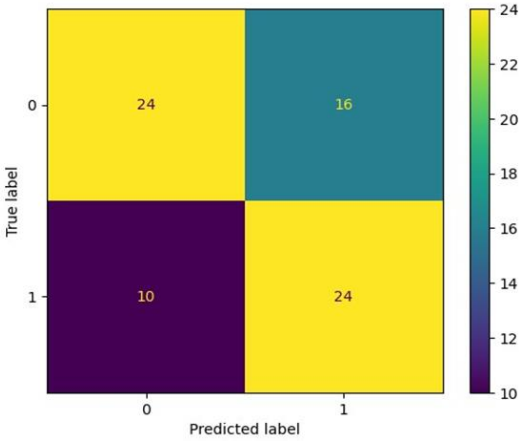


Figure 2: Confusion matrix for sentence
F1 score= 0.6486

The Experiment I, the percentage of total errors occurred during the testing for sentiment analysis of sentences are Type 1 is 30.76%, Type 2 is 7.69%, Type 3 is 53.84% and Type 4 is 8.69%. Experiment I resulted in an overall accuracy of 68%.

Experiment I B (Sentiment Analysis using verb chunks)

The results of tense Classification using Skip gram model on sentences was carried out in experiment I B, are shown in figure 3. The Experiment II was performed using the verb chunks to improve Type 3 errors encountered in Experiment I, and an improvement of 90% was achieved. Making the final accuracy of the experiment as 85%.

Sentiment Analysis was done on verb chunks using skip gram model and the resulting confusion matrix is shown in the figure below.

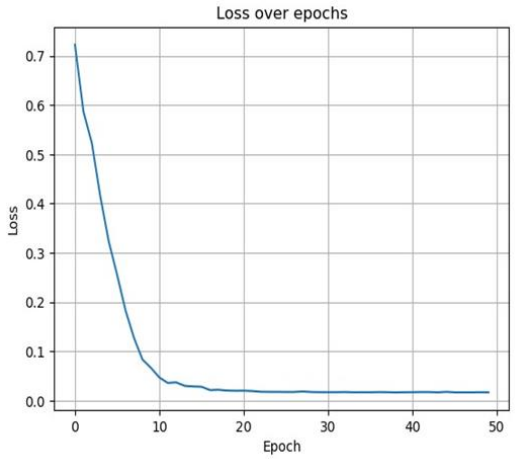


Figure 3: Loss over epoch,
Train Accuracy: 0.9898
Test Accuracy using MLP: 0.6890
Test Accuracy using RBF: 0.6697

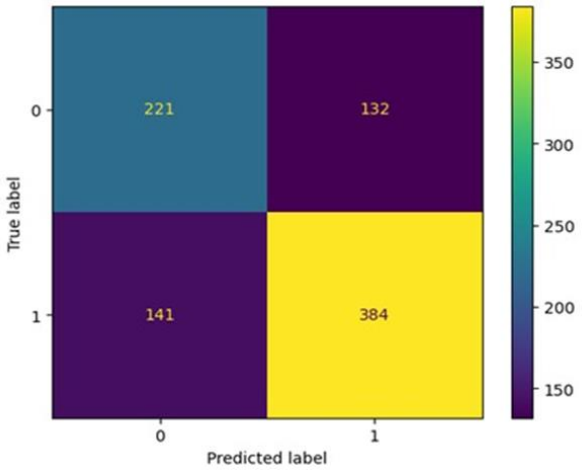


Figure 4: Confusion matrix for verbs
F1 score= 0.6779

Fig. 5 shows that the loss over epochs is least after 30 epochs. Hence it shows that 48 epochs used in this study were enough to train the data. In our study, Training data accuracy was 98.98%

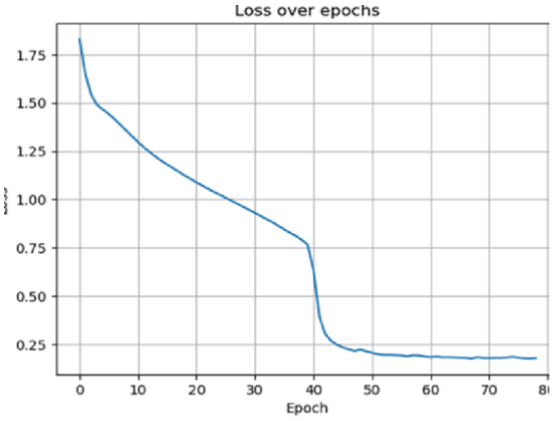
Fig. 6 shows that the accuracy and correctness of results in 878 verb chunks. The experiment showed that out of 878 verbs tested, True negative (0-0, green box) and True positive (1-1, yellow box) were 384+221 = 605. So, Test accuracy using MLP was 605/878 = 0.6890

The Experiment II was performed using the verb chunks to improve Type 3 errors and an improvement of 90% was achieved. Making the final accuracy of the experiment as 85%. So, Chunker based sentiment analysis using Verb Chunks is proven to be better than sentiment analysis using sentences.

Experiment II C (Tense classification using verb chunks)

The results of tense Classification using Skip gram model on sentences was carried out in experiment II C, are shown in figure 5. It shows that the loss over epochs is least after 78 epochs. Hence it shows that 78 epochs used in this study were enough to train the data. In our study, Training data accuracy was 90.94%.

Tense classification was done on verb chunks using skip gram model and the resulting confusion matrix is shown in the figure below.



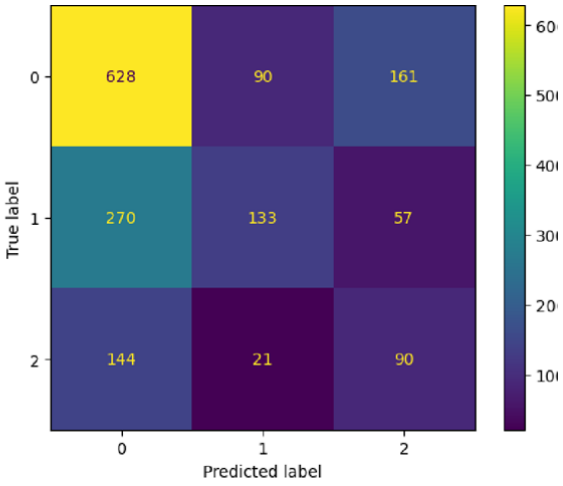


Figure 5: Loss over epoch,
Train Accuracy: 0.9094
Test Accuracy using MLP: 0.54

Figure 6: Confusion matrix for verbs Test Accuracy using RBF: 0.52.

	Precision	Recall	F1-score	support
0	0.60	0.71	0.65	879
1	0.55	0.29	0.38	460
2	0.29	0.35	0.32	255
Accuracy			0.53	1594
Macro avg	0.48	0.45	0.45	1594
Weighted avg	0.54	0.53	0.52	1594

Fig. 6 shows that the accuracy and correctness of results in 1594 verb chunks. The experiment showed that out of 1594 verbs tested, True positive (0-0,1-1,2-2 boxes) were 628+133+90 = 951. So, Overall Test accuracy using MLP was 851/1574 = 54%.

The accuracy was better for Past tense (65%) and very low for Present and Future Tense (38% & 32% resp.)

The Experiment II C showed that verb chunks had very poor overall accuracy. Hence experiment II D was conducted for Tense Classification using Sentences.

Experiment II D (Tense classification using sentences)

The results of tense Classification using Skip gram model on sentences was carried out in experiment II D, are shown in figure 7. It shows that the loss over epochs is least after 87 epochs. Hence it shows that 87 epochs used in this study were enough to train the data. In our study, Training data accuracy was 99.27%.

The tense classification has been categorised into 0= past tense, 1=present, 2=future tense. The resulting confusion matrix shown in Fig. 8. It shows that the accuracy and correctness of results in 2038 sentences. The experiment showed that out of 2038 sentences tested, True positive (0-0,11,2-2 boxes) were 1730+42+39 =1811. So, Test accuracy using MLP was 1811/2038 = 54%. The accuracy was better for Past tense (0-0 box) (94%) and very low for Present (1-1 box) and Future Tense (2-2 box) (35% & 52% resp.)

The Experiment II D showed that sentence-based tense classification had good accuracy for past tenses only.

Tense classification was done on sentences and the resulting confusion matrix is shown in the figure below.

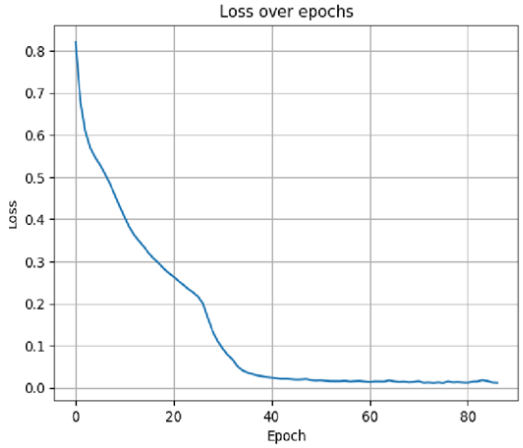


Figure 7: Loss over epoch,

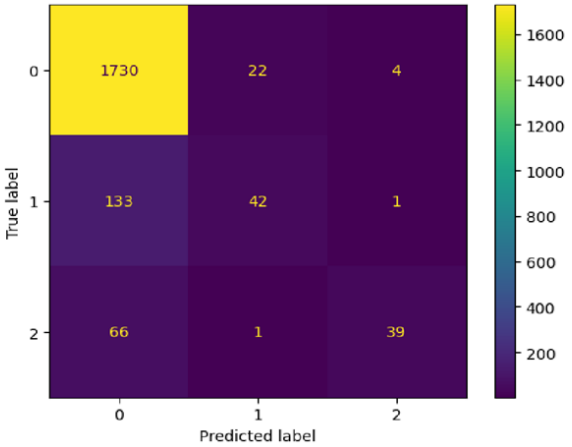


Figure 8: Confusion matrix for Sentences Train Accuracy: 0.9927

Test Accuracy using MLP: 0.8910
Test Accuracy using RBF: 0.8601

	Precision	Recall	F1-score	support
0	0.90	0.99	0.94	1756
1	0.65	0.24	0.35	176
2	0.89	0.37	0.52	106
Accuracy			0.89	2038
Macro avg	0.81	0.53	0.60	2038
Weighted avg	0.87	0.89	0.87	2038

So, we concluded that, Sentence based Tense Classification is proven to be better than Chunker based Tense Classification using Verb Chunks, while using Skip-gram model.

6. CONCLUSION

The article represents the sentiment analysis for Nepali Language using Skip-gram model. The identification of the tags of positive-negative is conducted in Nepali dataset available in Kaggle. The Experiment I, carried out for sentiment analysis of sentences, resulted in an overall accuracy of 68% and F1 score of 0.6486 (Fig- 4). The percentage of total errors occurred during the testing for sentiment analysis of sentences are Type 1 is 30.76%, Type 2 is 7.69%, Type 3 is 53.84% and Type 4 is 8.69%.

In Experiment II, Type 3 errors have been improved and an improvement of 90% was achieved. The highest F1 score of 0.6779 (Fig-6) is achieved for the positive –negative classification using skip-gram encoding technique employed on verb chunks with an accuracy of 85%. It has also been found that Multilayer Perceptron (MLP) classifier has performed better than Radial Basis Function (RBF) model for the Skip-gram model.

So, we can conclude that Chunker based sentiment analysis using Verb Chunks, has been proven to be better than sentiment analysis using sentences.

The article also represents the Tense Classification for Nepali Language using Skip-gram model. The experiment showed that out of 1594 verbs tested, True positive were 628+133+90 = 951. So, Overall Test accuracy using MLP was 851/1574 = 54%

The accuracy was better for Past tense (65%) and very low for Present and Future Tense (38% & 32% resp.)

The Experiment II C, Tense Classification experiment on verb chunks, showed that verb chunks had very poor overall accuracy. Hence experiment II C was conducted for Tense Classification using Sentences.
The Experiment II D, Tense Classification experiment on sentences, showed that out of 2038 sentences tested, True positive were 1730+42+39 =1811. So, Test accuracy using MLP was 1811/2038 = 54%

The accuracy was better for Past tense (94%) and very low for Present and Future Tense (35% & 52% resp.)

So, we concluded that, Sentence based Tense Classification is proven to be better than Chunker based Tense Classification using Verb Chunks, while using Skip-gram model.

7. SUMMARY OF THE FOUR EXPERIMENTS

Sentiment Analysis	**Experiment -I A**	**Experiment –I B**
Principle	Sentences encoded using skipgram and MLP classifier.	verb chunks encoded using skipgram and MLP classifier
Overall Accuracy	68%.	85%
MLP Test accuracy	0.6486	0.6890
F1 Score	0.6486	0.6779
Type 3 error	53.84%	Improved 90%

Tense Classification	**Experiment -II C**	**Experiment –II D**
Principle	verb chunks encoded using skipgram and MLP classifier	Sentences encoded using skipgram and MLP classifier.
Overall Accuracy	54%	89%
Training accuracy	0.90	0.99
MLP Test accuracy	0.54	0.89
F1 Score	0.65 (past tense)	0.94 (past tense)
	0.38 (present)	0.35 (present)
	0.32 (future)	0.52 (future)

ACKNOWLEDGEMENTS

We would like to acknowledge and express our sincere gratitude to the "Department of Science and Technology, Government of India", for sponsoring the project entitled "Study and develop a natural language parser for Nepali language "reference no. SR/CSRI/- 28/2015(G)" under the "Cognitive Science Research Initiative (CSRI)" to carry out this work. We also acknowledge the "TMA Pai University (Sikkim Manipal University)" research grant for supporting this work.

REFERENCES

[1] K. &. K. D. S. Shrivastava, "A Sentiment Analysis System for the Hindi Language by Integrating Gated Recurrent Unit with Genetic Algorithm.," *The International Arab Journal of Information Technology.,* no. 17. 954-964. 10.34028/I ajit/17/6/14., (2020).

[2] S. Ghosh, "Multitasking of sentiment detection and emotion recoignition in code-mixed Hinglish data.," *Knowledge-Based Systems.,* no. 260.110182.10.1016/j.knosys.2022.110182, (2022).

[3] B. K.Bal., "Structure of Nepali Grammar (1st.ed.)"., ,Nepal. : Madan PuraskarPustakalaya, 2004.

[4] T. Mikolov, K. Chen, G. Corrado and a. J. Dean., "Efficient estimation of word representations in vector space.," *ICLR Workshop,* 2013.

[5] T. Mikolov, I. Sutskever, K. Chen, G. Corrado and J. Dean., "Distributed representations of words and phrases and their compositionality.," *NIPS ,,* 2013.

[6] K. Kiprono Elijah Koech, "https://towardsdatascience.com/cross-entropy-loss-functionf38c4ec8643e," 20 Oct 2020. [Online]. Available: https://towardsdatascience.com/cross-entropy-lossfunction-f38c4ec8643e.

[7] A. Pradhan, A. Yajnik and a. Prajapati., "A Conceptual Graph Approach to the Parsing of Projective Sentences.," *International Journal of Mathematics and Computer Science,,* no. 15(1) 199–221, (2020).

[8] A. Pradhan and A. Yajnik, "Parts-of-speech tagging of Nepali texts with Bidirectional LSTM, Conditional Random Fields and HMM," *Multimedia Tools and Applications,* 2023.

[9] A. Pradhan and A. Yajnik, "Probabilistic and Neural Network Based POS Tagging of Ambiguous Nepali text:," *A comparative Study. ISEEIE, Association for Computing Machinery, Seoul,Republic of Korea.,* no. https://doi.org/10.1145/3459104.3459146., (2021).

[10] A. MacKinlay., "The effects of Part –Of-Speech Tagsets on Tagger Performance (Bachelor's thesis) Master's thesis", .University of Melbourne ,Australia., 2005.

[11] https://www.kaggle.com/datasets/aayamoza/nepali-sentiment-analysis

[12] Piryani, Rajesh & Piryani, Bhawna & Singh, Vivek & Pinto, David. "Sentiment analysis in Nepali: Exploring machine learning and lexicon-based approaches". Journal of Intelligent and Fuzzy Systems. 1-12. 10.3233/JIFS-179884. (2020).

[13] Tung, A.K.H. Rule-Based Classification. In: Liu, L., Özsu, M. (eds) Encyclopedia of Database Systems. Springer, New York, NY. https://doi.org/10.1007/978-1-4899-7993-3_559-2, (2017).

[14] Akbar Karimi., "Hidden Markov Models vs. Conditional Random Fields"., June 13, 2023., https://www.baeldung.com/cs/hidden-markov-vs-crf.

[15] Xingyu Fu, Weijia Shi, Zian Zhao, Xiaodong Yu, and Dan Roth., "Design challenges for lowresource cross-lingual entity linking"., arXiv preprint arXiv:2005.00692., 2020.

AUTHORS

Dr Archit Yajnik, is an Additional Professor in the Department of Mathematics at Sikkim Manipal Institute of Technology. He is a member of DPRC Committee and Coordinator of Workshops and Seminar Committee. His area of interest is Wavelet Analysis and Natural Language Processing and has published several papers in Indexed Journals in NLP.

Ms. Sabu Lama Tamang is a PhD. Research Scholar at Sikkim Manipal Institute of Technology under the guidance of Dr. Archit Yajnik. She has completed her M.Sc. Mathematics from the prestigious NIT Durgapur. Her area of interest is NLP.