Clickbait Detection for Amharic Language Using Deep Learning Techniques

[1]Rajesh Sharma R, [2]Akey Sungheetha, [3]Mesfin Abebe Haile, [4]Arefat Hyeredin Kedir, [5]Rajasekaran A and [6]Charles Babu G

[1,2]Department of Computer Science and Engineering, Alliance College of Engineering and Design, Alliance University, Bangalore, India.

[3,4]Department of Computer Science and Engineering, School of Electrical Engineering and Computing, Adama Science and Technology University, Adama, Ethiopia.

[5]Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai, India.

[6]Department of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India.

[1]sharmaphd10@gmail.com, [2]sun29it@gmail.com, [3]esfin.abebe@astu.edu.et, [4]arefat.hyeredi@gmail.com, [5]arajasekaran139@gmail.com, [6]charlesbabu.grient@gmail.com

Correspondence should be addressed to Mesfin Abebe Haile : mesfinabha@gmail.com

Article Info

Journal of Machine and Computing (http://anapub.co.ke/journals/jmc/jmc.html)

Doi: https://doi.org/10.53759/7669/jmc202404058

Received 05 Decmber 2023; Revised from 12 March 2024; Accepted 04 June 2024

Available online 05 July 2024.

Abstract – Because of, the increasing number of Ethiopians who actively engaging with the Internet and social media platforms, the incidence of clickbait is becomes a significant concern. Clickbait, often utilizing enticing titles to tempt users into clicking, has become rampant for various reasons, including advertising and revenue generation. However, the Amharic language, spoken by a large population, lacks sufficient NLP resources for addressing this issue. In this study, the authors developed a machine learning model for detecting and classifying clickbait titles in Amharic Language. To facilitate this, authors prepared the first Amharic clickbait dataset. 53,227 social media posts from well-known sites including Facebook, Twitter, and YouTube are included in the dataset. To assess the impact of conventional machine learning methods like Random Forest (RF), Logistic Regression (LR), and Support Vector Machines (SVM) with TF-IDF and N-gram feature extraction approaches, the authors set up a baseline. Subsequently, the authors investigated the efficacy of two word embedding techniques, word2vec and fastText, with Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) deep learning algorithms. At 94.27% accuracy and 94.24% F1 score measure, the CNN model with the rapid Text word embedding performs the best compared to the other models, according to the testing data. The study advances natural language processing on low-resource languages and offers insightful advice on how to counter clickbait content in Amharic.

Keywords – Clickbait Detection, Artificial Neural Networks, Natural Language Processing, Machine Learning Techniques, Deep Learning Techniques, Amharic Language, Social Media.

I. INTRODUCTION

According to Loewenstein's [1], clickbait stimulates the reader's curiosity by creating a cognitive gap that needs to be filled between what we already recognize and what we would like to recognize. Thus, they sense the impulse to click the link. This theory serves as the major foundation for clickbait studies. Clickbait appears in different forms, ranging from partial facts, exaggerated headlines, and catchy wordings [2] to enticing advertisements, and misinformation campaigns [3]. Clickbait also clogs up social media and news feeds while also violating journalistic codes of ethics [4]. Clickbait is a growing concern on the Internet, particularly within social networks, such as Facebook, Twitter, and YouTube [5]. It is very challenging to identify these clickbait contents manually due to their vast number and wide distribution. It is essential to detect clickbait content on the Internet using automated means [3]. The extensive usage of the Internet and the social media led to the growing of the acceptance of the online advertisements, which is complemented by a disturbing spread of clickbait content [3]. Advertisement revenue is one of the driving factors that content creators focus on, chasing how many clicks they get while depreciating the quality of content they produce and share. This behavior and practice lead to users feeling deceived, irritated, and unsatisfied [6]. Most of the content on websites is supported by online advertising, which is an inevitable part of the current Internet. The economy of online advertising has become a de-facto standard. They are run under an economic model known as Clickthrough rate (CTR). Clickbait detection is a task that falls under the broad umbrella of Natural Language Processing (NLP). Clickbait detection involves using NLP techniques to automatically identify and classify clickbait content in online social media posts, news articles, and headlines [7]. It has recently gained traction in academic research due to the growing prevalence of clickbait content in the digital world and its potential to manipulate public opinion and spread misinformation. In Ethiopia, social media and Internet usage is growing rapidly every year. Clickbait is being extensively used in these digital platforms; mediums that share clickbait hide behind a financial goal achieved through reaching more people. The contents on these platforms are exaggerated and miss a foundational context. It has been weakening media credibility and promoting the online spread of rumors and false information [8].

The only script with African roots is the Ethiopian script used in Amharic, which is called Ethiopic or Fidel. After Arabic, it is currently the second most spoken Semitic language worldwide. The official working language of the Ethiopian government is Amharic [9]. It has unique characteristics, notably a subject-object-verb (SOV) word order, which is probably a result of extended contact with Cushitic languages [10]. Amharic, on the other hand, is a low-resource language, meaning that it has restricted access to the tools required for NLP and other computational linguistic applications [9]. This paper presents an artificial neural network method for identifying clickbait in Amharic. The goal is to fulfill the demand for a sophisticated computational linguistic model that can identify and categorize Amharic clickbait in texts found on social media. To the best of our knowledge, this study is the first to investigate clickbait detection and categorization for Amharic, and it aims to address the dearth of available literature in the field. We created the Amharic clickbait dataset in addition to establishing baselines and contrasting the effectiveness of several neural network designs for clickbait detection. The remaining part of the paper is structured as follows. Section 2 discussed the related works of the study. Section 3 explained in details about the dataset, methodology and proposed work. Section 4 discussed the experimental findings and results. Finally, Section 5 presents the conclusion and the gaps identified for future study.

II. RELATED WORKS

Early clickbait detection studies heavily relied on feature engineering methods. [11] discussed that to identify clickbait, different techniques can be employed such as lexical features, image-based features, and user-behavior features. Machine learning approaches have been employed by multiple research groups to identify clickbait. used a machine-learning classifier to identify clickbait [12]. Word patterns, N-grams, and linguistic analysis were among the fourteen sets of components utilized to train the classifier. For the study of clickbait detection, numerous studies have proposed a variety of methodologies utilizing a range of traditional machine learning algorithms, such as Logistic Regression [2], [13], Support Vector Machine [12], Decision Tree [7], Random Forest [14], as well as deep learning techniques like Recurrent Neural Network [15], LSTM [16], Gated Recurrent Unit [17], and CNN [2], [18]. As described in [2], a general end-to-end Convolutional Neural Network based method that automatically detects clickbait was presented. Without depending on any auxiliary features, it was able to induce several beneficial qualities for the final work. Agrawal et al. [19] were also among the first to use a CNN-specific deep learning technique that works well for classifying headlines as clickbait or not, with an accuracy of 0.90 on the English Twitter dataset. In a similar vein, [20] presented a Recurrent Neural Network (RNN)-based neural network architecture that produced satisfactory clickbait detection results. A deep learning approach for clickbait detection was presented by Naeem et al. [4]. The authors utilized an LSTM decision-making device known as POSAM (Part of Speech Analysis Model) to train a model for information discovery by identifying the fundamental structure and qualities of clickbait. [2] There has been several research projects published targeting other languages recently. Some notable clickbait detection studies in non-English languages include Filipino [15], Thai [7][21]. Indonesian [22], Chinese [5], Telugu [23], and Arabic [3]. In the work of Klairith et al. [7], the authors proposed six different sets of neural networks: BiLSTMs with word-level embedding models for detecting clickbait in Thai language. They crowdsourced 30,000 headlines, and the model performed well achieving an accuracy rate and F1 score of 0.98. In the Philippines, [15] collected English and Filipino headlines from social media and classified clickbaits using a neural network architecture and Word2Vec for word embedding. They were able to achieve a 91.5% accuracy level in their experiments. [24] presented a web application using BERT and a RESTFUL API that divested the computation sources required to training the model on the cloud server. Furthermore, a more recent and broadened work that discussed overcoming NLP challenges in resource-poor languages for clickbait detection, while using Telegu as the target language was done by [23]. They stated that clickbait-related tasks cannot be effectively handled with machine translation as a cross-language experiment because the sense of clickbait content in the source language may alter, and the interest component might be lost in translation [23] .

In their study on Amharic text classification, [25] employed machine learning techniques and investigated the impact of operations like stemming and POS tagging on classification performance. They utilized a medium-sized, manually annotated Amharic corpus and observed that stemming did not significantly affect the performance of text classification in Amharic. The study [26] conducted a study on automated Amharic news categorization using a neural network approach with Learning Vector Quantization on a dataset containing 1,762 items across nine categories. [27] created annotation tools and classification models to aid in their work, and they also researched sentiment analysis for material found on Amharic social media. A general-purpose Amharic corpus (GPAC) [28], a novel Amharic fake news detection dataset

(ETH_FAKE), an Amharic fastText word embedding (AMFTWE), and an Amharic false news detection model are among the other contributions that are provided [9]. sought to solve the issue of Amharic fake news detection by utilizing deep learning techniques and a recently assembled dataset. Bi-GRU and CNN outperformed other recurrent and attention-based models, resulting in a f1-score of 94% and an accuracy of 93.92%, respectively. Using a freshly curated dataset of Amharic fake news received from Facebook, [8] has investigated lexicon-stance based Amharic fake news detection. The experimental fallouts demonstrated that the integration of mixture features (combining lexicon and stance) yields significant improvements over previous lexicon-based detection approaches.

III. RESEARCH METHODOLOGY

Data Gathering

The dataset gathering of the clickbait detection is based the following three steps:

- Collecting Amharic tweets, titles, and posts from a social media through API calls or scraping, includes Twitter tweets, YouTube video titles, and Facebook posts.
- Cleaning each record to minimize the adverse effects of impurities on the model.
- Annotating the data and consolidating it into a dataset file.

When it comes to selecting publishers from the platforms (YouTube, Twitter and Facebook), engagement metrics are a critical factor. Publishers with a large number of followers or subscribers and high engagement rates such as likes, comments, replies, and shares are more likely to generate clickbait content or authentic content. On the other hand, publishers with a smaller following and a lower engagement rate may be more likely to produce legitimate content. Therefore, it is crucial to take a holistic approach to publisher selection, considering factors such as topics and themes, frequency of posting, and overall credibility and reputation. The study applied both purposive and systematic sampling methods to collect a representative dataset.

The data collection process resulted in a total raw dataset instance of 53,227 gathered from the mentioned platforms. The statistics of the data instances across the various data sources and the clickbait nature is show in the **Table 1.**

Table 1. Distribution of the Dataset

Data Source	Number of Clickbait	Number of non-clickbait	Total
YouTube	14,378	11,061	25,439
Twitter	7,136	9,432	16,568
Facebook	5,164	6,056	11,220
Total	26,678	26,549	53,227

During the data collection process for clickbait, distinguishing genuine public pages from clickbaiting ones poses a specific challenge. To streamline the analysis of extensive social media data for potential clickbait content, specific criteria were employed to select pages for inclusion in the dataset. These criteria included language, timeframe, flagged pages, redirecting links, content focus, follower count and content volume. Overall, the dataset preparation process for an Amharic clickbait dataset involves cleaning the data such as removing missing fields, standardizing its format, removing duplicate fields, addressing the non-Amharic text, the **Table 2** is showing that some of sample text of Amharic with English translation.

Table 2. Amharic Clickbait Samples with English Translation

Amharic Clickbait Sample Statements with English Translations
ከሽንብራ በጣም የሚያስገርም ጥቅም በቀላል ወጪ ፊቶት ላይ ላለዉ ችግር መፍትሄ አያምልጦት!
Don't miss out on the amazing benefits of chickpeas for the problems on your face at a simple cost.
በየቀኑ ክሊክ የሚደረጉ 95 ስራዎች አሉት። እያንዳንዳቸው 67.00 ብር ያስገኛሉ።
It has 95 jobs to be clicked every day. They reward 67.00 Birr each.
የኢትዮጵያ አንድ ብር 2,000 ብር እየተሸጠ ይገኛል። ሳንቲሙ እጃችሁ ላይ የሚገኝ አሁኑኑ ሽጡ። አድራሻ ሊንኩ ላይ አለ።
One Ethiopian Birr is being sold at 2,000 Birr. Sell the coin now if you have it. Address is at the link.
ልታዩት የሚገባ ምርጥ የአማረኛ ፊልም እንዳያመልጦ
Don't miss the best Amharic movie you should watch
አስደንጋጭ ዜና... ከአይሮፕላን ላይ ወድቆ ሞተ
Shocking news... he died after falling from the plane
ስለ ፍሬህይወት ታምሩ አስገራሚና ያልተሰሙ እውነታዎች
Surprising and unheard facts about Firehiwot Tamiru
የሚስቴን ጉድ ስራ ቦታ ድረስ ሄጄ አየሁ። አይታቹህ ፍረዱኝ።
I went to my wife's workplace and saw the shame. Watch and judge for yourself.

Data Annotation

Sentence annotation for the purpose of identifying clickbait in Amharic is a subjective technique. We have defined the attractiveness of clickbait headlines as their capacity to arouse curiosity, create attention, and persuade readers to click on the link in the instructions sent to the annotators. We have also used automated techniques to label instances automatically following certain rules.

This study explored some options for the annotation tools to use, that include using a spreadsheet in a tabular format to label, and a social media-based annotation tool using Telegram chatbot. The level of agreement between annotators is computed using the kappa (κ) statistic. It is important to note that a κ value of 1 represents perfect agreement, while 0 indicates chance agreement [29]. The Fleiss's kappa score, a measure of inter-annotator agreement, was determined to be 0.94.

Data Preprocessing

While it is crucial to perform preprocessing steps, it is important to exercise caution when applying all commonly used techniques, as they may have unintended consequences for this model.

Stopword Removal

The use of stop word removal as a preprocessing step in clickbait detection experiments can be a topic of debate. Stop words are frequently occurring words in a language that are often removed to reduce noise and improve computational efficiency in NLP tasks [8]. It is advisable to conduct experiments with and without stopword removal to evaluate the impact on the detection performance and make an educated decision based on the results obtained, where in our case several irrelevant stopwords are removed. Other pre-processing procedures such as cleaning, normalization and tokenization are also applied.

Data Cleaning

The cleaning process involves removing special characters, symbols, punctuation, emojis, and whitespace from the dataset. Non-Amharic text is also eliminated to ensure the dataset consists solely of Amharic content. Additionally, irrelevant numerals are removed, and long sentences are filtered out to maintain a fair document length.

Normalization

To achieve normalization of Amharic words, [30] discussed types of normalization issues in Amharic word tokens. The first issue involves identifying and replacing shorthand representations of words written with forward slashes ('/') or periods ('.'). For instance, "ም/ቤት" is replaced with "ምክር ቤት," and "ዓ.ም" is replaced with "ዓመተ ምህረት." The second normalization issue addresses the presence of homophone Fidels or spellings within the Amharic writing system. These Fidels have the same pronunciation but different symbols. Although they convey different connotations in Ge'ez (the parent language of Amharic), they have been used interchangeably in Amharic. Examples of such Fidels include አ and ዐ, ፀ and ጸ, ሰ and ሠ, ሀ, ሐ, and ኀ. For instance, the word "Artist" could be written as አርቲስት, ዐርቲስት, ዓርቲስት - all with the same pronunciation but different orthography.

Feature Extraction

As this is a newly organized dataset and first of a kind experiment on Amharic clickbait, we considered the traditional feature representations techiques such as: To set up the experiment, use TF-IDF (Term Frequency-Inverse Document Frequency) and BoW (Bag of Words). By employing word clouds to visualize the data, the findings on the most frequently occurring terms in both clickbait and non-clickbait groups can be further supported.

Fig 1. Amharic Clickbait Terms Word Cloud.

As it can be observed in the **Fig 1** above, the most common Amharic clickbait words predominantly revolve around current, trending, and hot topics. These words often employ hyperbolic language, provocative subjects, and sensationalized claims to capture the readers' attention. Word embedding, which represents words as dense vectors in a continuous multi-dimensional space, is another well-liked method for feature extraction. By capturing the semantic links between words based on their contextual usage, it enables algorithms to use this data for a range of natural language processing tasks, such as the detection of clickbait. The clickbait detection models can increase the accuracy and efficacy of the detection process by utilizing word embeddings such as Word2Vec and fastText, which enable them to capture the semantic and contextual meaning of words in the language.

Dataset Splitting

Every data point is guaranteed to be used for both training and testing thanks to this procedure. We used a 5-fold cross-validation dataset splitting technique, in which the model is tested on the fifth fold of the dataset and is trained on the first four. To provide a qualitative analysis and a quantitative metric for evaluating text classification tasks such as clickbait detection models, there are various evaluation metrics. The commonly used metric to measure effectiveness is accuracy, which measures how often the classifier predicts the correct label (e.g., whether a given text content is clickbait). Additional evaluation criteria employed in the study were the confusion matrix, recall, precision, and F1-score.

Exploratory Data Analysis

As an exploratory data analysis phase, the clickbait dataset was examined before undergoing cleaning and preprocessing steps. Several features are extracted from the gathered dataset to gain insights into the lexical tones present in the dataset. These features included analyzing whether a title contained a question mark or an exclamation mark, determining the number of words in each headline, and identifying whether a title included numeric characters or not.

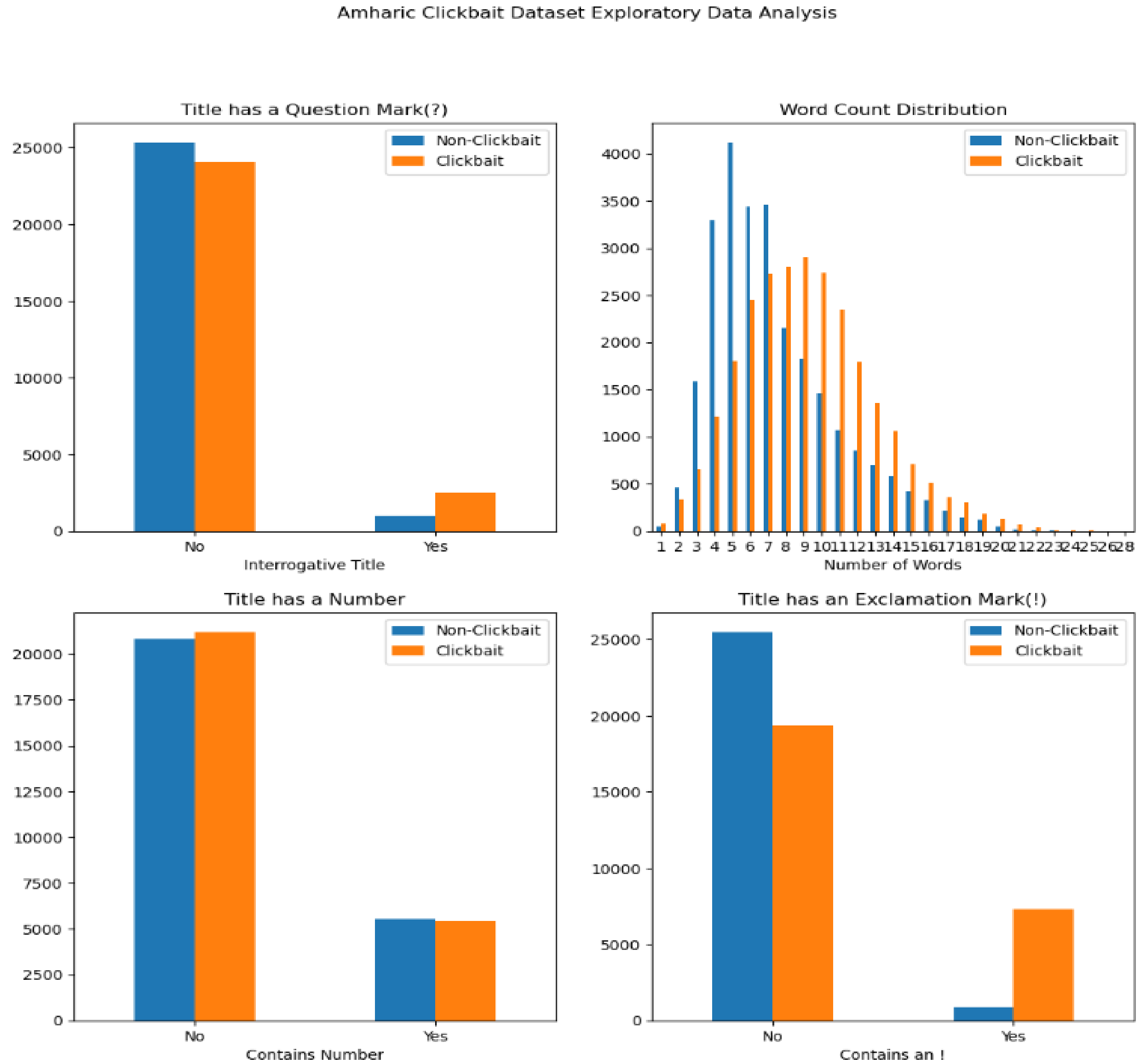

Fig 2. Exploratory Data Analysis Results of the Amharic Clickbait Dataset.

Fig 2 illustrates a notable disparity among clickbait and non-clickbait items where the number of interrogatory words and question mark are used. Clickbait titles tend to employ a higher number of questions, capitalizing on the human brain's inherent cognitive bias and evoking a sense of curiosity. Similarly, clickbait title exhibits a distinct sentence structure characterized by a greater number of words and a composition that incorporates an abundance hyperbolic language signified using the "!" exclamation mark, enticing readers with sensational, provoking, and alluring language. Both clickbait and non-clickbait items exhibit a similar presence of numbers, indicating that numerical information is found in both types of content to a comparable extent.

Problem Definition

As discussed by [31], The definition of the clickbait categorization problem is as follows: Let C be the collection of two classes, {clickbait, non-clickbait}, where {clickbait} sentences are represented by one class and `non-clickbait} sentences by another. Given a dataset *D* consisting of sentences, our objective is to train a model that can classify each sentence into one of the two classes. The training set *t* consists of labeled sentences $\{t,C\}$, where each sentence is associated with the corresponding class label $\{t,C\} \in D \times C$. The goal is to learn a function f that can effectively map the sentences in the dataset to their respective classes $f : D \rightarrow C$, enabling accurate classification of clickbait and non-clickbait sentences.

Establishing Baseline

This study experimented with Logistic Regression, Support Vector Machine Random Forest and XGBoost, with TF-IDF and n-gram as features to establish baseline.

Logistic Regression (LR)

Several studies have shown the effectiveness of LR in clickbait detection. [32] used Logistic Regression and reported good results, where [33] utilized LR with TF-IDF and n-gram features to establish baseline and they achieved competitive results.

Support Vector Machine (SVM)

The model produces a hyperplane to separate data points into two or more classes and classifies new data points by using the hyperplane's orientation. Previous studies, such as works by [12], [13] have demonstrated the effectiveness of SVM in clickbait detection.

Random Forest (RF)

It leverages ensemble learning techniques and the strength of decision trees to effectively classify clickbait and non-clickbait instances. Its ability to handle noisy data and capture non-linear relationships makes it a suitable candidate for our research. RF models have shown promise in clickbait detection, as highlighted in the study by [14], [34] where they achieved the best results.

XGBoost

Study carried out by [34] demonstrates the efficacy of XGBoost and other Gradient Boosting algorithms like LightGBM [23] in accurately identifying clickbait content. It is particularly adept at handling unstructured text data by using techniques such as bag-of-words or TF-IDF representations.

Keyword Embedding

The suggested model initially gathers texts that are both clickbait and non-clickbait in order to generate the data corpus. Following that, word embeddings of these textual contents are created, and they are then used as input for various neural network models. [35] utilized word2vec to detect Amharic hate speech in social networks and achieved sound results. Word2vec word-level embedding creates a 300-dimensional vector using a continuous bag of words architecture. The fastText model enables us to construct representations for uncommon words by treating each word as a bag of character n-grams. The authors of a study [9] investigated the application of fastText for Amharic fake news identification. They found that fastText outperformed traditional word embedding methods in capturing the semantics of Amharic words, even with limited training data [36].

Proposed Architecture

Fig 3 illustration of the proposed approach high-level architecture of the Amharic clickbait detection model and the general procedures applied in the design and implementation of the various models.

Model Selection

Two types of word embedding vectorization approaches were used in the study's experiments: convolutional neural networks and recurrent neural networks, such as LSTM and GRU. This example makes use of word2vec and fastText word-level embeddings.

Long Short-Term Memory (LSTM)

LSTMs are a kind of recurrent neural network fine suited for catching lengthy term dependances among chunks of text information contained within a single news article title or post description to distinguish between clickbaiting words among all other words present in the text. LSTMs are particularly helpful when dealing with highly visible topics like entertainment celebrities because they capture relationships across different types of contextual information contained within an individual title or post description quickly and accurately. In a study by [16], LSTM models were employed for clickbait pattern detection in news headlines. Additionally, [15] explored the use of Recurrent Neural Networks like Bi-LSTM for clickbait exposure in Filipino and English language.

Gated Recurrent Units (GRU)
GRU is intended to help classic RNNs overcome the vanishing gradient issue that frequently arises and to enable it to successfully capture the long-term dependencies in sequential data. In contrast to conventional RNNs, GRU has gating features that allow it to selectively update and reset data inside the hidden state, increasing its efficiency and ability to gather pertinent contextual information. By analyzing the sequential nature of textual information, GRU can effectively capture important patterns and dependencies within the text, enabling accurate classification. We observed from the works of [7], [17] where they used GRU and BI-GRU models in clickbait detection and yielded reliable results.

Convolutional Neural Networks (CNN)
Multiple weight matrices are combined using a convolutional neural network (CNN), which then creates a new vectorized representation of the input. A fully connected linear layer is then applied to this new representation to perform classification or regression. This architecture, combined with the utilization of word embeddings and attention mechanisms, allows CNNs to effectively capture both local and global dependencies in text data as it was demonstrated with the works of [2], [18], leading to robust and accurate clickbait classification results. Recent study done by Abebaw et al., (2022) demonstrated the usefulness of CNN text classification approach with word2vec embedding for detecting Amharic hate speech, which lead us to choose the CNN model as the primary proposed model for Amharic clickbait detection.

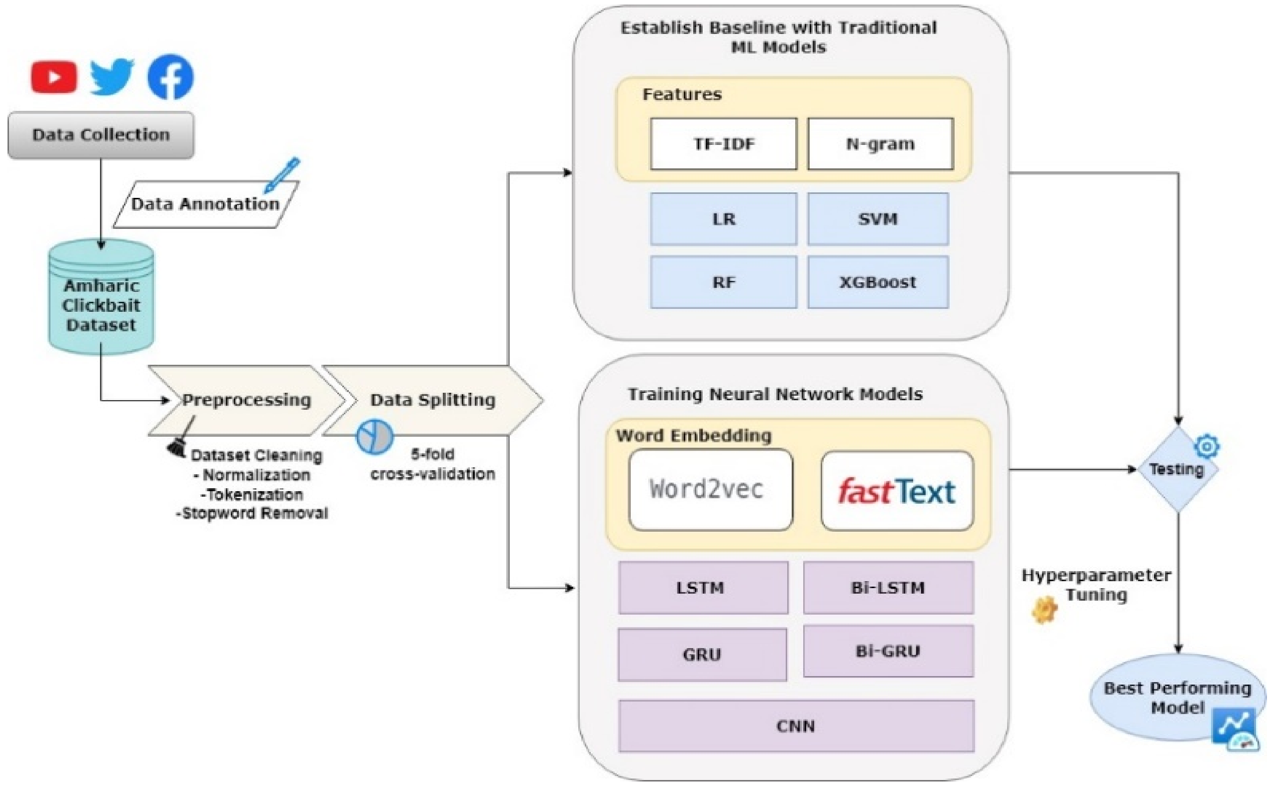

Fig 3. Architecture of the Proposed Clickbait Detection Model.

Based on the works of [21], we framed the feature extraction from the clickbait titles as vector representing the local features within a sequence of word embeddings through the convolution step, we apply one-dimensional max pooling. By taking into account pairs of subsequent words, irrespective of their precise placement within the wider input sequence, a high-level feature representation with a pooling size of 2 or 3 is produced. This allows us to discard less-relevant local information **Fig 4** shows that feature extraction method.

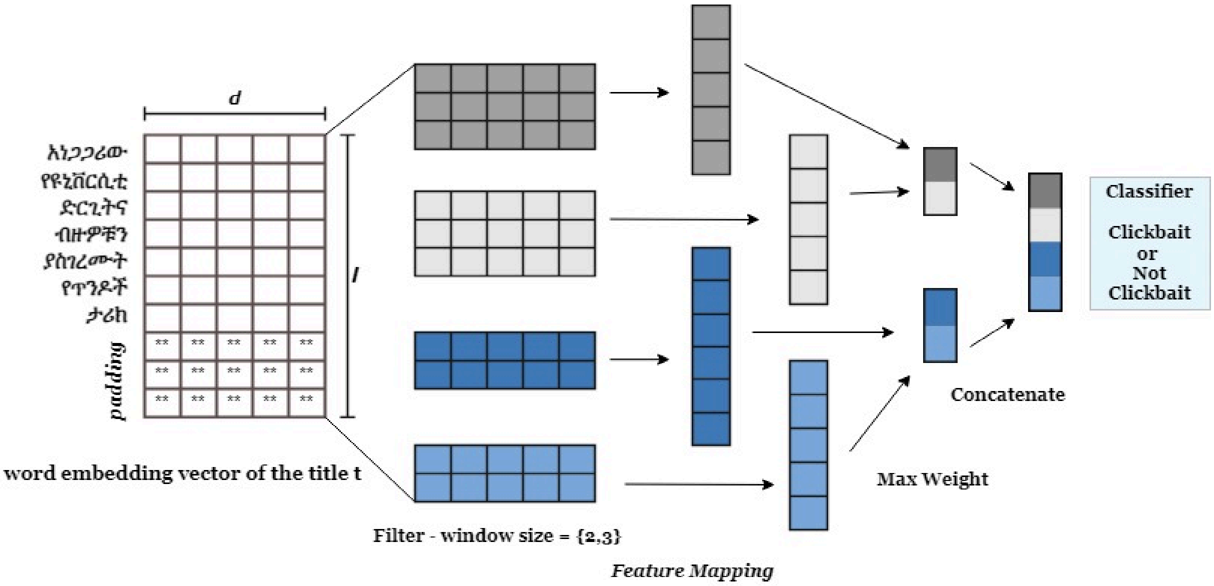

Fig 4. CNN Architecture for Feature Extraction of Amharic Clickbait Detection.

After extracting the local features, we perform global max pooling over the entire sequence of filters, processing them one by one. Finally, a sigmoid initiation task is applied in the output neuron to classify a given text as clickbait or non-clickbait.

IV. RESULTS AND DISCUSSION

Baseline Results

Our first experimental result established a baseline using TF-IDF with unigram and bigram features. It was observed that the TF-IDF feature performed well when using unigrams but not with bigrams. This can be attributed to the nature of Amharic and the characteristics of clickbait content. Unigrams capture individual words, allowing the model to identify specific terms and their importance in distinguishing clickbait from non-clickbait. On the other hand, bigrams combine pairs of adjacent words, which may not capture the nuanced patterns and linguistic structures that are prevalent in clickbait headlines. Higher order n-grams like trigram have not yielded any comparable results.The Random Forest experiment yielded promising results after conducting tuning with GridSearch, using a maximum depth of 300 and 900 estimators. From the results showed in **Table 3**, we can see that the unigram-based features outperformed the bigram-based features. Across all models, including Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and XGBoost, the accuracy, precision, recall, and F1-score were consistently higher when using unigram features combined with TF-IDF. The XGBoost model achieved the highest accuracy of 90.86%, among traditional ML models when using unigram features with TF-IDF, indicating its strong performance in classifying clickbait and non-clickbait instances in Amharic. XGBoost also exhibited high recall and F1-score values, consistently surpassing other models.

Table 3. Baseline Model Performance Results using Traditional Machine Learning Techniques

Feature Engineering	Model used	Model Accuracy	Model Precision	Recall Model	F1-score Model
Unigram + TF-IDF	L R	0.8801	0.8847	0.8778	0.8812
	SVM	0.8831	0.8661	0.9029	0.8841
	R F	0.8922	**0.9161**	0.8740	0.8946
	XG-Boost	**0.9086**	0.9100	**0.9095**	**0.9098**
Bigram + TF-IDF	L R	0.8425	0.8433	0.8362	0.8402
	SVM	0.8307	0.8217	0.8294	0.8255
	R F	0.8683	0.8652	0.8517	0.8643
	XG-Boost	0.8617	0.8595	0.8778	0.8685

The rationale behind using XGBoost in this experiment lies in its capacity to manage complex text patterns and capture non-linear associations. By leveraging the power of ensemble learning and combining multiple weak classifiers, Gradient Boosting models such XGBoost and LightGBM have been proven to effectively model intricate clickbait detection patterns, leading to accurate predictions and improved performance [34]. It is notable that the Random Forest (RF) classifier emerged as the second-best performing traditional model with a precision of 91.61%. It is essential to note that the Arbitrary Forest pattern's performance was further enhanced through parameter tuning using grid search. By optimizing the number of estimators with a value of 900 and maximum depth up to 300, the model achieved improved precision and overall performance.

While the bigram features combined with TF-IDF still achieved reasonably good accuracy and performance metrics across all models, they were consistently lower than those obtained using unigram features. This suggests that for Amharic clickbait detection, unigram features provide more informative and discriminative signals for classification, capturing the important lexical nuances present in the language.

Experimental Results

As per the proposed architecture, after establishing baseline - we experimented on the two variants of recurrent neural networks; LSTM and Bi-LSTM which mostly depends on the historical context of inputs rather than the last input. We used the two word embedding vectors we prepared, word2vec and fastText to feed the vector information along with the text sequence.

The LSTM model while using the word2vec embedding classified 9813 titles as clickbait and non-clickbait effectively, from the total 10,646 testing set. It conceded accuracy of 92.18% and f1-cut of 92.24%. The second variation of the LSTM model we experimented on, the bi-directional LSTM (Bi-LSTM), which is an extension of traditional LSTMs is assumed to enhance model performance as it can retain information from different occurrences. We also explored the use of GRUs (Gated Recurrent Units), which are computationally efficient models similar to LSTMs, but with a different architecture. The GRU model using the word2vec embedding vectors correctly classified 9841 instances of titles as True Positives and True Negatives, while misclassifying 805 instances as False Negatives and False Positives out of the averaged fold size of 10646. The model succeeded an accuracy of 92.44% and an f1-score of 92.50%.

Moreover, the bidirectional GRU model with fastText embeddings performed somewhat better, with a f1-score of 93.13% and an accuracy of 93.07%. These findings demonstrate how well Amharic clickbait detection works when fastText embeddings are used with GRU and Bi-GRU models. Additionally, we evaluated the effectiveness of Convolutional Neural Network (CNN) models utilizing Word2Vec and fastText, two distinct word embedding methods, for Amharic clickbait detection. Using word2vec embeddings, the CNN model obtained 93.60% accuracy. With a few fine-tuned settings, the precision and f1-score were relatively high, at 93.90% and 93.66%, respectively.However, the CNN model that made use of fastText embeddings performed even better. It obtained a 94.20% accuracy rate. The model's accuracy in distinguishing between clickbait and non-clickbait text was demonstrated by the precision and f1-score for fastText embedding, which were 93.82% and 94.23%, respectively. **Fig 5**'s model accuracy graph, which starts at about 87% and rises steadily with each epoch, depicts the training accuracy's continuous rising trend. The model's ability to accurately classify the training data improves significantly, reaching a high of 94% on the 4th epoch. This indicates that the model is effectively learning the underlying patterns and features in the dataset.

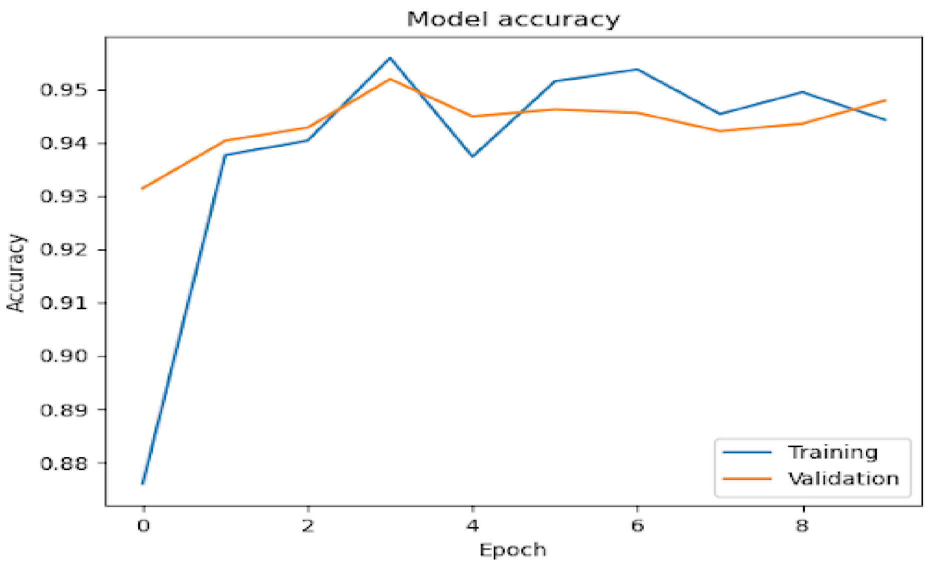

Fig 5. CNN Model Accuracy Graph for Amharic Clickbait Detection Model.

Similarly, the validation accuracy values demonstrate a positive trend throughout the training process. The validation accuracy reaches a peak of 94% on the 10th epoch. This suggests that the model can successfully distinguish between clickbait and non-clickbait content because it performs well not only on the training data but also on fresh, unknown data. The CNN model that used fastText embeddings had the highest accuracy of 94.20%, which means that it accurately recognized the greatest percentage of titles that were clickbait and those that weren't. This outcome shows how well the CNN architecture captures key clickbait trends and characteristics. In terms of precision and recall, the CNN model with fastText embeddings likewise produced the greatest f1-score of 94.23% when analyzed. This suggests striking a balance between reducing false positives (precision) and accurately categorizing clickbait (recall). The high f1-score highlights the CNN model's excellent performance in precisely identifying clickbait content in Amharic when it uses fastText embeddings.

Model Comparison

The precision values among the neural network models, the Bi-GRU model using fastText embeddings achieved the highest precision score of 93.85%. This implies that it had the lowest number of false positives, effectively distinguishing non-clickbait text from clickbait ones as it had similarly been observed in the work of [17]. **Table 4** shows the classification of Performance Results of Neural Network model.

Table 4. Classification Performance Results of Neural Networks Model

Feature Engineering	Model Used	Model Accuracy	Model Precision	Model Recall	Model F1-score
word2vec	LSTM	0.9218	0.9104	0.9347	0.9224
	BiLSTM	0.9276	0.9360	0.9198	0.9278
	GRU	0.9244	0.9327	0.9174	0.9250
	BiGRU	0.9295	0.9221	0.9382	0.9301
	CNN	0.9360	0.9390	0.9343	0.9366
fastText	LSTM	0.9265	0.9193	0.9352	0.9272
	BiLSTM	0.9260	0.9359	0.9177	0.9267
	GRU	0.9282	0.9333	0.9246	0.9289
	BiGRU	0.9307	**0.9385**	0.9243	0.9313
	CNN	**0.9420**	0.9382	**0.9465**	**0.9423**

When the outcomes of the LSTM and Bi-LSTM models are compared, it becomes clear that the Bi-LSTM model performs marginally better, attaining a higher f1-score and accuracy. The Bi-LSTM's bidirectionality enables it to consider words that come before and after, providing a more thorough comprehension of the textual context. The results highlight the impact of the choice of embedding technique on the performance of neural network models. fastText embeddings, which capture subword information, appear to be more effective in capturing the nuances of Amharic language, resulting in higher accuracy and f1-score compared to word2vec embeddings.

In general, the experimental findings indicate that the CNN model with fastText embeddings achieved the highest accuracy and f1-score among the neural network models, which proves its effectiveness in Amharic clickbait detection. From **Fig 6**. The Bi-GRU model showed the highest precision, emphasizing its ability to minimize false positives.

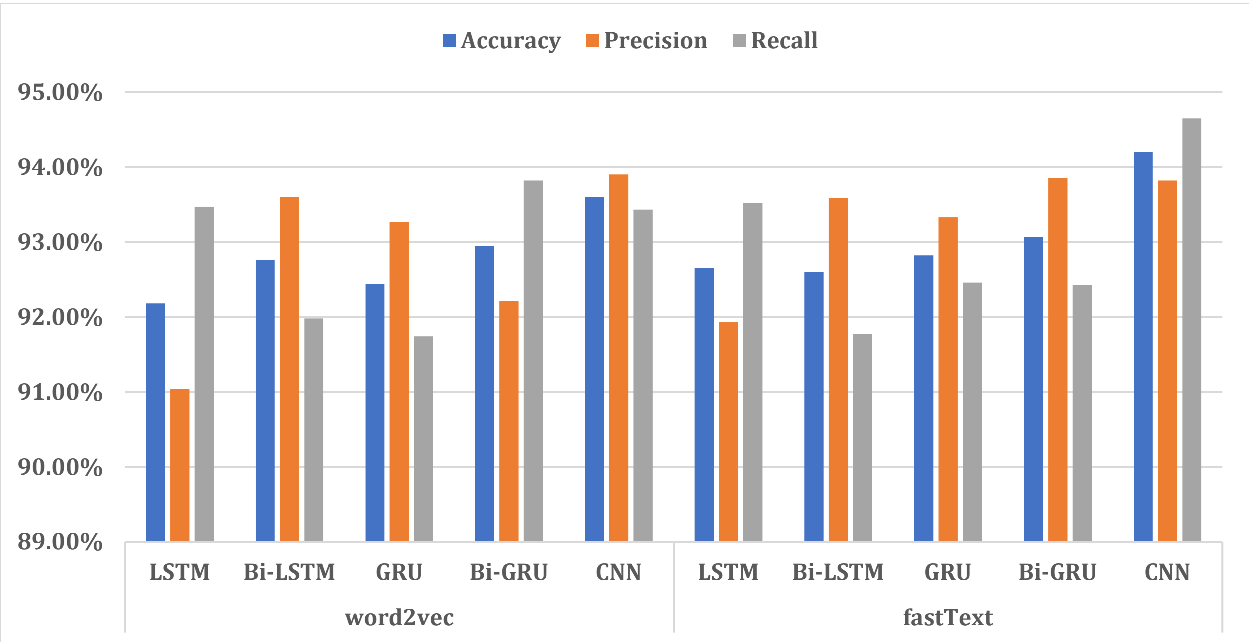

Fig 6. Comparison of various Models.

Hyperparameter Tuning

As a binary classification task, it is practical to choose Sigmoid function as an output layer activation function. Tanh and sigmoid activation functions are better suited for recurrent layers like LSTM because they have a bounded output range (-1 to 1 for tanh and 0 to 1 for sigmoid). On the other hand, ReLU activation function is commonly used in CNNs because it introduces non-linearity and helps in capturing complex patterns in the vector data. ReLU does not suffer from the vanishing gradient problem and can provide faster convergence during training [31].

Table 5. Hyperparameter Tuning (Activation Function & Epoch) Results

Activation Function Used	Model Used	Epochs Size	Model Accuracy
Sigmoid	LSTM + word2vec	10	91.74%
		20	92.18%
	LSTM + fastText	10	92.17%
		20	92.65%
Tanh	LSTM + word2vec	10	90.87%
		20	91.82%
	LSTM + fastText	10	92.33%
		20	92.69%
ReLU	CNN + word2vec	10	93.28%
		20	93.60%
	CNN + fastText	10	93.86%
		20	**94.20%**

The results shown in **Table 5** gives the accuracy when the hyperparameter tuning of different activation functions with different epoch is performed at the hidden layers of LSTM and CNN models. Despite testing different batch sizes, ranging from 16 to 64, the resulting performance metrics remained relatively stable. This suggests that the choice of batch size did not have a substantial impact on the model's ability to learn and make accurate predictions. Generally, it was showing consistency, indicating that the default batch size produced satisfactory results without the need for further adjustment. We explored tuning the learning rate within the range of 0.001 to 0.5. The results showed that slower learning rates (0.001, 0.002, 0.003) produce better performance, however slower learning rates may require running more epoch cycles. After a certain threshold around 0.01, the performance of the model begins to deteriorate. This performance aligns with our expectations, as high learning rates can cause the weights to diverge and prevent the network from effectively learning and converging to an optimal solution.

For our CNN modeling technique, we utilized the different hyperparameter settings to investigate the changes. Specifically, we experimented with static, non-static, and random variations. As word2vec provides a fixed 300 dimension for all vectors, it was ideal for tuning the model setting with word2vec embedding rather than fastText. The non-static modeling technique achieves an accuracy of 93.6%. This can be attributed to the feature that non-static modeling enables the network to capture the subtle changes in word meanings and associations that may be specific to the clickbait [21]. By updating the word embeddings during training, the model can better capture the evolving semantics and linguistic nuances present in the Amharic clickbait dataset. The other two CNN hyperparameters we tuned were window size and dimension of the word embedding vector. The window size, which determines the number of surrounding words considered as an n-gram, was explored using values of 2, 3, and 4. This allowed for an investigation into the impact of different context sizes on the performance of the models. Additionally, the dimensions of the fastText word vectors were examined, with options of 100 and 300 dimensions, corresponding to the available pretrained models. word2vec only provides the 300 dimensions in a pretrained format, which has been experimented on earlier.

Table 6. Different Window size and Vector Dimension Tuning Result

Vector dimension	Window size	Accuracy	F1-score
100	2	92.85%	92.82%
	2,3	93.08%	93.11%
	2,3,4	92.27%	92.23%
300	2	94.20%	94.23%
	2,3	**94.27%**	**94.24%**
	2,3,4	93.12%	93.07%

The numbers of n-gram are {2,3,4}, and vector dimensions 100 and 300 were utilized in our experiment to determine an appropriate window size and vector dimension for the word embedding. **Table 6** contains the results of the experiment. The result shows that the most appropriate window size is n = {2,3} and the 300-vector dimension which gives 94.27% accuracy, the highest accuracy we achieved with any of the experiments we have done on the Amharic clickbait dataset.

In summary, through the process of hyperparameter tuning, several optimal settings were identified to enhance the performance of the CNN model for Amharic clickbait detection. Specifically, by combining a window size of n={2,3}, along with 300 dimensions of fastText word embedding, an epoch size of 20, and employing the non-static model setting of the CNN model with ReLU as the activation function, the highest accuracy of 94.27% and the highest f1-score of 94.24% were achieved.

V. CONCLUSION

The findings primarily directed on a content-based approach, considering textual components like teaser messages, titles, and headlines. The study proposed and implemented various neural network models for Amharic clickbait detection. These models incorporated word sequence and word-level embeddings using Amharic word2vec and fastText. Notably, when employing fast Text embedding, the CNN model succeeded in an accuracy of 94.27% and an F1-score of 94.24%. Incorporating attention models within deep learning algorithms can potentially improve the performance of the detection system, especially when combined with a more comprehensive dataset. Considering social-context and user-based features, along with engagement metrics, can provide valuable insights for clickbait detection. Exploring multimodal clickbait detection in images such as analyzing thumbnails can expand the scope of the research. Developing user-interactive components such as browser extensions can facilitate the identification and filtering of clickbait content. Extending the study of clickbait classification in multilingual sense which considers other local languages can contribute to a more comprehensive understanding of clickbait in diverse linguistic contexts.

Data Availability

The datasets are available at following link.
https://drive.google.com/file/d/1CczBDCplru1Wegq2YtgOpx3kltlTI0W/view?usp=sharing

Conflict of Interest

The authors declare that there are no conflicts of interest.

Acknowledgements

This research study was funded by Adama Science and Technology University under the grant number: ASTU/SM-R/851/23. The authors would like to express their gratitude for the assistance received from the institute.

References

[1]. G. Loewenstein, "The psychology of curiosity: A review and reinterpretation.," Psychological Bulletin, vol. 116, no. 1, pp. 75–98, 1994, doi: 10.1037//0033-2909.116.1.75.
[2]. J. Fu, L. Liang, X. Zhou, and J. Zheng, "A Convolutional Neural Network for Clickbait Detection," 2017 4th International Conference on Information Science and Control Engineering (ICISCE), Jul. 2017, doi: 10.1109/icisce.2017.11.
[3]. M. Al-Sarem et al., "An Improved Multiple Features and Machine Learning-Based Approach for Detecting Clickbait News on Social Networks," Applied Sciences, vol. 11, no. 20, p. 9487, Oct. 2021, doi: 10.3390/app11209487.
[4]. B. Naeem, A. Khan, M. O. Beg, and H. Mujtaba, "A deep learning framework for clickbait detection on social area network using natural language cues," Journal of Computational Social Science, vol. 3, no. 1, pp. 231–243, Feb. 2020, doi: 10.1007/s42001-020-00063-y.
[5]. C. Zhang and P. D. Clough, "Investigating clickbait in Chinese social media: A study of WeChat," Online Social Networks and Media, vol. 19, p. 100095, Sep. 2020, doi: 10.1016/j.osnem.2020.100095.
[6]. P. Mowar, M. Jain, R. Goel, and D. K. Vishwakarma, "Clickbait in YouTube Prevention, Detection and Analysis of the Bait using Ensemble Learning," *arXiv preprint arXiv:2112.08611*, 2021.
[7]. P. Klairith and S. Tanachutiwat, "Thai Clickbait Detection Algorithms Using Natural Language Processing with Machine Learning Techniques," 2018 International Conference on Engineering, Applied Sciences, and Technology (ICEAST), Jul. 2018, doi: 10.1109/iceast.2018.8434447.
[8]. I. N. Awol and S. M. Gashaw, "Lexicon-Stance Based Amharic Fake News Detection," *researchgate.net*, May 2022, Accessed: May 10, 2023. [Online]. Available: https://www.researchgate.net/profile/Ibrahim-Awol/publication/369203279_Lexicon-Stance_Based_Amharic_Fake_News_Detection/links/64105d84a1b72772e4f9308a/Lexicon-Stance-Based-Amharic-Fake-News-Detection.pdf
[9]. F. Gereme, W. Zhu, T. Ayall, and D. Alemu, "Combating Fake News in 'Low-Resource' Languages: Amharic Fake News Detection Accompanied by Resource Crafting," Information, vol. 12, no. 1, p. 20, Jan. 2021, doi: 10.3390/info12010020.
[10]. I. Zitouni, Ed., Natural Language Processing of Semitic Languages. Springer Berlin Heidelberg, 2014. doi: 10.1007/978-3-642-45358-8.
[11]. Y. Chen, N. J. Conroy, and V. L. Rubin, "Misleading Online Content," Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection, Nov. 2015, doi: 10.1145/2823465.2823467.
[12]. A. Chakraborty, B. Paranjape, S. Kakarla, and N. Ganguly, "Stop Clickbait: Detecting and preventing clickbaits in online news media," 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Aug. 2016, doi: 10.1109/asonam.2016.7752207.
[13]. A. Geckil, A. A. Mungen, E. Gundogan, and M. Kaya, "A Clickbait Detection Method on News Sites," 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Aug. 2018, doi: 10.1109/asonam.2018.8508452.
[14]. M. Potthast, S. Köpsel, B. Stein, and M. Hagen, "Clickbait Detection," Advances in Information Retrieval, pp. 810–817, 2016, doi: 10.1007/978-3-319-30671-1_72.
[15]. P. K. Dimpas, R. V. Po, and M. J. Sabellano, "Filipino and english clickbait detection using a long short term memory recurrent neural network," 2017 International Conference on Asian Language Processing (IALP), Dec. 2017, doi: 10.1109/ialp.2017.8300597.
[16]. S. Manjesh, T. Kanakagiri, P. Vaishak, V. Chettiar, and G. Shobha, "Clickbait Pattern Detection and Classification of News Headlines Using Natural Language Processing," 2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), Dec. 2017, doi: 10.1109/csitss.2017.8447715.

[17]. L. M. Bantelay, M. Abebe, R. Sharma Rajendran, A. Sungheetha, and S. N, "Heuristic Pneumonia and Tuberculosis Detection in X-Ray Images Using Convolutional Neural Networks," 2023 Annual International Conference on Emerging Research Areas: International Conference on Intelligent Systems (AICERA/ICIS), Nov. 2023, doi: 10.1109/aicera/icis59538.2023.10420329.
[18]. H.-T. Zheng, J.-Y. Chen, X. Yao, A. K. Sangaiah, Y. Jiang, and C.-Z. Zhao, "Clickbait Convolutional Neural Network," Symmetry, vol. 10, no. 5, p. 138, May 2018, doi: 10.3390/sym10050138.
[19]. A. Agrawal, "Clickbait detection using deep learning," 2016 2nd International Conference on Next Generation Computing Technologies (NGCT), Oct. 2016, doi: 10.1109/ngct.2016.7877426.
[20]. A. Anand, T. Chakraborty, and N. Park, "We Used Neural Networks to Detect Clickbaits: You Won't Believe What Happened Next!," Advances in Information Retrieval, pp. 541–547, 2017, doi: 10.1007/978-3-319-56608-5_46.
[21]. M. Ali Nur, M. Abebe, and R. S. Rajendran, "Handwritten Geez Digit Recognition Using Deep Learning," Applied Computational Intelligence and Soft Computing, vol. 2022, pp. 1–12, Nov. 2022, doi: 10.1155/2022/8515810.
[22]. R. Sharma R*, A. Sungheetha, and J. Nuradis, "Brain Tumor Classification by EGSO Based RBFNN Classifier," International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 5, pp. 3005–3012, Jan. 2020, doi: 10.35940/ijrte.e6073.018520.
[23]. M. Marreddy, S. R. Oota, L. S. Vakada, V. C. Chinni, and R. Mamidi, "Clickbait Detection in Telugu: Overcoming NLP Challenges in Resource-Poor Languages using Benchmarked Techniques," 2021 International Joint Conference on Neural Networks (IJCNN), Jul. 2021, doi: 10.1109/ijcnn52387.2021.9534382.
[24]. M. N. Fakhruzzaman and S. W. Gunawan, "Web-based Application for Detecting Indonesian Clickbait Headlines using IndoBERT," Feb. 2021, doi: 10.48550/arxiv.2102.10601.
[25]. E. Tilahun, M. Abebe, R. Rajesh Sharma, A. Sungheetha, and N. Sengottaiayn, "Culture Reflecting Artistic Fashion Design using Deep Learning and Assisting Custom Algorithm," 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), Oct. 2023, doi: 10.1109/iccams60113.2023.10525953.
[26]. W. Kelemework, "Automatic Amharic text news classification: A neural networks approach," *Ethiopian Journal of Science and Technology*, vol. 6, no. 2, pp. 127–137, 2013, Accessed: May 17, 2023. [Online]. Available: https://www.ajol.info/index.php/ejst/article/view/117217
[27]. S. M. Yimam, H. M. Alemayehu, A. Ayele, and C. Biemann, "Exploring Amharic Sentiment Analysis from Social Media Texts: Building Annotation Tools and Classification Models," Proceedings of the 28th International Conference on Computational Linguistics, 2020, doi: 10.18653/v1/2020.coling-main.91.
[28]. E. N. Hailemichael, "Fake news detection for amharic language using deep learning," *academia.edu*, 2021, Accessed: May 17, 2023. [Online]. Available: https://www.academia.edu/download/84664801/ERMIAS_20NIGATU.pdf
[29]. R. Sharma, A. Sungheetha, and P. Marikkannu, "Three-dimensional MRI brain tumour classification using hybrid ant colony optimisation and grey wolf optimiser with proximal support vector machine," International Journal of Biomedical Engineering and Technology, vol. 29, no. 1, p. 34, 2019, doi: 10.1504/ijbet.2019.10017861.
[30]. B. Gambäck, F. Olsson, A. Argaw, and L. Asker, "Methods for Amharic part-of-speech tagging," *First Workshop on Language Technologies for African Languages*, Mar. 2009, Accessed: May 17, 2023. [Online]. Available: https://www.diva-portal.org/smash/record.jsf?pid=diva2:1042595
[31]. C. Kiran et al., "Cyber Physical System Centred Protective Laboratory for Industries," Advances in Microelectronics, Embedded Systems and IoT, pp. 365–374, 2024, doi: 10.1007/978-981-97-0767-6_30.
[32]. X. Cao, T. Le, J. (Jiasheng,) Zhang, and D. Lee, "Machine Learning Based Detection of Clickbait Posts in Social Media," Oct. 2017, Accessed: Apr. 06, 2023. [Online]. Available: https://arxiv.org/abs/1710.01977v1
[33]. P. Adelson, S. Arora, and J. Hara, "Clickbait; Didn't Read: Clickbait Detection using Parallel Neural Networks," 2017, Accessed: May 16, 2023. [Online]. Available: http://cs229.stanford.edu/proj2017/final-reports/5231575.pdf
[34]. K. Shu, S. Wang, T. Le, D. Lee, and H. Liu, "Deep Headline Generation for Clickbait Detection," 2018 IEEE International Conference on Data Mining (ICDM), Nov. 2018, doi: 10.1109/icdm.2018.00062.
[35]. R. Rajesh Sharma and P. Marikkannu, "Hybrid RGSA and Support Vector Machine Framework for Three-Dimensional Magnetic Resonance Brain Tumor Classification," The Scientific World Journal, vol. 2015, pp. 1–14, 2015, doi: 10.1155/2015/184350.
[36]. Z. Abebaw, A. Rauber, and S. Atnafu, "Multi-channel Convolutional Neural Network for Hate Speech Detection in Social Media," Advances of Science and Technology, pp. 603–618, 2022, doi: 10.1007/978-3-030-93709-6_41.