[1]Ayushi Chaudhary

[2]Prof. Ashish Sharma

Automatic Classification of Stegano files using the Features of Natural Language Processing and Hybrid Deep Learning Algorithms

Journal of Electrical Systems

Abstract: - The corporate sector has adopted communication as its principal medium due to the improvements of the digital era. In the past, in-person meetings, purchases, settlements, business talks, and profile-taking were all done. However, everything has gone digital in the present period. Over the last several years, there has been a noticeable exponential rise in the amount of complicated texts and documents that require a deeper comprehension of machine learning techniques in order to recognize languages in a variety of applications. Outstanding results in the processing of natural languages have been obtained by a number of Artificial Intelligent techniques. A key component of machine learning and deep learning techniques' efficacy is their capacity to understand intricate models and non-linear connections in data. Choosing the right architecture, frameworks, and algorithms for classifying input, like text files, managing audio and video files, however, is a difficult undertaking. The proposed work's goal is to identify text, audio messages, chatbots, and smart records using natural language processing. Using a hybrid deep learning approach, text, voice, and video recordings are used as inputs to be classified. Problem Synopsis: With contact becoming more and more important to business, companies have created advanced NLP programs. Through chatbots, digital records, phone conversations, and messages, these NLP swiftly fulfil human desires. Customer preferences, desires, and demands have been found to be more strongly impacted by the ease of communication and connection. Today's service providers use chatbots, digital records, messaging apps, email, and phone conversations as their main methods of communication for almost all of their trade channels of preference, client queries, and transactions. Proposed Method: The study shows how input stegano file is processed automatically based on user reactions, text message replies, and audio record identification during communications using text content, voice messages, and audio using the features of Natural Language Processing and hybrid Deep Learning Algorithms

Keywords: Automatic Classification, Natural Language Processing, Deep Learning, CNN, LSTM.

I. INTRODUCTION

Natural Language Processing, or NLP, [1] is a branch of artificial intelligence that focuses on training computers to interpret and comprehend natural language. It serves as the basic model for text interpretation, voice recognition systems, and other scenarios in AI when humans communicate with computers. Machines can comprehend humans and appropriately respond to machines when Linguistics is employed as a device for different applications, unlocking huge possibilities in many domains. NLP integrates computing semantics languages with analytical, deep learning models and machine learning models. These techniques, when combined, allow machines to analyze natural language as a type of textual or speech input and 'interpret' its correct interpretation. Natural Language Processing enables computer programs to translate texts from one language to another, interpret verbal commands, and quickly summarise vast amounts of text in real-time.

Natural language processing is essential in the future because it allows users to design computational models and algorithms that accept data fragments as input that is given in the form of speech, words, or a combination of both forms and modify them according to the methodology used in the machines. Different approaches in order to process the input in the form of text, audio and video by Natural Language Processing along with Deep Learning and Machine Learning are LSTM (Long Short Term Memory) [2], Named Entity Recognition framework, Word Embedding framework, Sequence 2 Sequence model, Feature-based extraction Model using Fuzzy inference rules, Google Neural Machine Translation, Word Recognition Acoustic framework, Neural Machine Translation, Phase based machine translation.

Many NLP activities [3] split up human text and speech input in ways that help the machine sound right about what it's receiving. Among these activities are as follows:

Voice recognition, often known as voice-to-text, is the process of reliably translating audio signals into textual data. Voice detection is necessary for any software that accepts voice speech instructions or responds to spoken queries.

[1,2]GLA University

ayushichaudhary11@gmail.com

ashish.sharma@gla.ac.in

Speech recognition is very challenging because the way people talk may differ in many ways for example pronunciation during the speech, incorrect grammar, accents, etc.

Part of grammar tagging: The practice of identifying the verbal component of a certain phrase or textual content depending on its application and environment is known as linguistic component tagging. The word 'make' is used as an action in the sentence 'I can make a cardboard aircraft,' and as a noun in the sentence 'What make of aircraft do you own?'

Lexicon differentiation is the process of determining the significance of a phrase with several interpretations using lexical analysis to discover which phrase is most relevant in the present situation. Lexicon differentiation, for example, assists in distinguishing the meanings of the verb 'made' in 'made the grade' (accomplish) vs.' made a bet' (place).

The task of determining whether two phrases pertain to an identical item is called co-reference assessment. The most frequent instance is establishing the person or thing to whom a given pronoun refers (e.g.,' she' = 'Marie'), although it may also include detecting metaphors or idioms in the text (e.g., a case in which 'mountain lion' belongs to a giant furry man rather than a mammal). Sentiment classification seeks to derive emotional characteristics from texts, such as opinions, feelings, humour, bewilderment, and suspicions. Natural language translation can be thought of as the inverse of voice recognition or way of speaking; it is the effort of converting structured data into natural speech [3].

1.1 Applying the concept of Deep Learning:

A deep neural network [4] is a subset of machine learning method that consists of multiple neurons, some of which consist of hidden layers shown in Figure 1. Deep neural networks apply advanced mathematical models to interpret information in a wide range of ways. A neural network is an adaptable framework of outcomes as combinations of input data that comprises many levels: input nodes that include data input; convolution neurons that include computational nodes known as neurons; and output neurons that include one or more neurons.

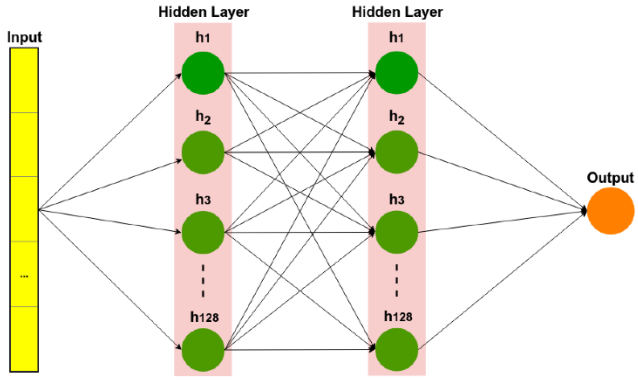

Figure 1: General Architecture of Deep Learning Approach

Deep learning [5] applies a multi-layered technique to the neural network's hidden layers. Deep learning models, train and retrieve characteristics intelligently, resulting in higher precision and efficiency. In principle, the hyper-parameter of different classifiers is also automatically measured. Deep learning and artificial neural networks are presently providing the optimal answers to several challenges in the recognition of images and speech, in addition to the processing of natural language.

The research idea is motivated from three aspects:

- To make the framework to be independent of input files [audio, video and image].
- To reduce the manual classification of input files. Move towards the automatic classification by introducing deep learning algorithm.
- To enhance the existing security level from encryption to blockchain. For enhancing the security blockchain technology is introduced. The first level of security is provided by encryption and second level through blockchain.

Research Gaps

- The existing steganography process are limited to one category of input type. That means existing model is designed to work on one type of input either image or video or audio.
- Existing models provide the highest level of security is provided in existing steganography process is through encryption which is treated as one level security to stegano documents.

Problem Statement

- Aim is to find a solution for utilizing the concept of artificial intelligence in steganographic processes. In many sectors, organizations opt for manual classification of documents due to the unavailability of the technology.
- Existing stegano documents are stored using single level of security model using the encryption techniques and proposed work objective is to apply the multi-level security model.
- For the purpose of security enhancement is to extend up to second level, that is find a solution for enhancing from encryption to blockchain technology.

II. LITERATURE REVIEW

Alharbi, Ahmed Sulaiman M et. al. [9] in their work offer a neural network framework that takes into account user behavior within a specific article (tweet). The author employs a Convolutional Neural Network (CNN) as its neural network. The proposed system is tested using two sets of data from the SemEval-2016 Training session. The suggested model improves existing approaches such as Naive Bayes (NB) and Support Vector Machines (SVM). The proposed model demonstrates the analysis beyond the text of a document (tweet) which is advantageous in emotion classification.

Collobert et al. [11] in their work explore a basic deep learning methodology that gives the best results in most state-of-the-art techniques in numerous NLP tasks, including named-entity recognition (NER), semantics role labelling (SRL), and Part of speech tagging. With these results, numerous complex deep training approaches have been designed to address challenging NLP issues. The author also examines significant deep learning-related algorithms and techniques used in natural language problems, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and recursive neural networks. The author also addresses memory-enhancing tactics, cognitive techniques, and how unsupervised algorithms, reinforcement learning techniques, and, more recently, deep generative models have been used for linguistic applications.

The research [12] concentrated on the fundamental issue in SA, which is sentiment polarity classification. It makes use of the Amazon ratings and reviews set of data. The Supervised ML Support Vector Machine (SVM), Ensembled Random Forest (RF), and Supervised ML Nave Bayesian (NB) approaches were used, and the results were better.

The scholars of [13] used the WEKA evaluation instrument to contrast SVM and NB approaches for Islamic Twitter message and text analysis. The TF-IDF and trigonometric functions approaches were employed in the weighting strategy and written statement among the similarity calculation, respectively. The results of the experiments show that NB was conducted well enough in both precision and duration.

The study [14] used CNN on micro-blogging responses to determine the behavior and viewpoints of online consumers regarding special occasions. The CNN methodology was used because it overwhelms the extraction of features and impliedly adapts the data via training. A catalog of 1000 micro-blogging commentaries was created and labeled with three distinct tags.

III. PROPOSED METHODOLOGY:

Natural Languages generally contain all types of linguistics that differ in accent, perception, and significance. These natural languages vary depending on the individuals, making it challenging to design and develop a model that accurately understands these notions. As a result, Machine Learning cannot be used as a stand-alone NLP approach. ML models are effective for recognizing overall input features or recognizing the properties contained in the input data. Yet, they conflict when it comes to identifying patterns or relating responses to specific items or topics. As a result, in the proposed model rules for the Hybrid Deep Learning model with the NLP technique are explained. These constraints and features assist the algorithms in much more accurately linking classifications to human

perception. In the proposed model Convolutional Neural Networks with Long Short Term Memory approaches are applied in a hybrid Deep Learning Technique.

The LSTM deep learning model is used to retrieve the meaning of words at a higher level to extract past text analysis within manuscripts. For huge manuscripts, the LSTM algorithm achieves long-term relationships between lexical items effectively. The proposed hybrid deep neural network, consisting of CNN and LSTM, extracts relevant features and trains long-range relationships that aid in effectively acquiring the information required. The proposed methodology can help in the prediction of different types of input, and the classification of different patterns in the input files which are in the form of text, audio, and video files in a more efficient way.

Steps involved in training the proposed model:

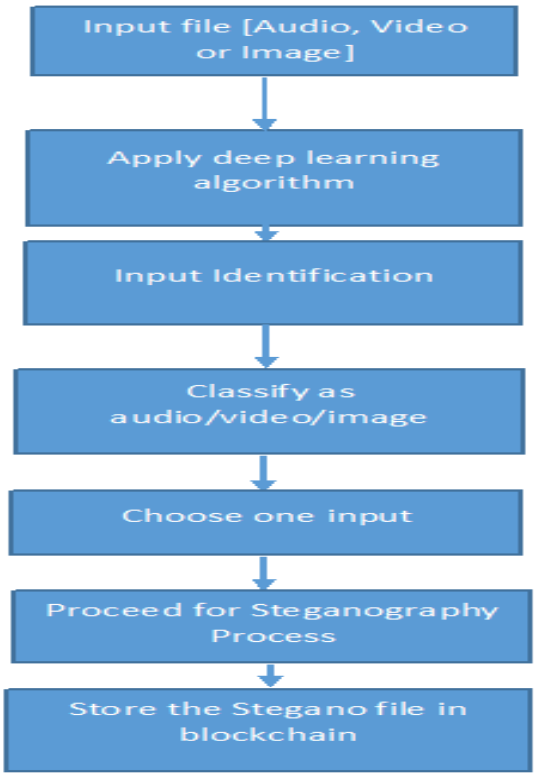

Figure 4: Flow diagram of the proposed work

Data Acquisition: All analyses require data. The procedure for gathering or collecting input data from a variety of sources, including web crawls, data warehouses, live data streams like audio and video, and so on is nothing but data collection. The obtained data can then be saved in the workspaces as Data frames, which operate smoothly with Python's statistics components.

Pre-processing and data cleaning: In most circumstances, unprocessed data contains details of the input data collected that do not benefit in interpretation and instead deprecate predictions. As a result, once the data is collected, pre-processing and data cleaning is carried out such as eliminating anomalies, discriminating among labels and characteristics, standardizing the attributes, discarding column which will not add to the evaluation, applying attribute values, etc.

Exploratory Data Analysis (EDA): The EDA phase is significant at the commencement of the analysing the model since it gives details about the data that we will be working with. EDA assists in data interpretation by identifying structures, classifications, and mathematical inferences such as average, percentile, sample variance, and deviation.

Training the proposed hybrid model: As all of the datasets have been standardized and are prepared for evaluation, we begin developing the relevant hybrid deep learning model. The goal of developing a model is to provide it with

a large amount of data from which to train. The algorithm learns about data characteristics and trains itself to perform on unlabelled data sets of a similar data type.

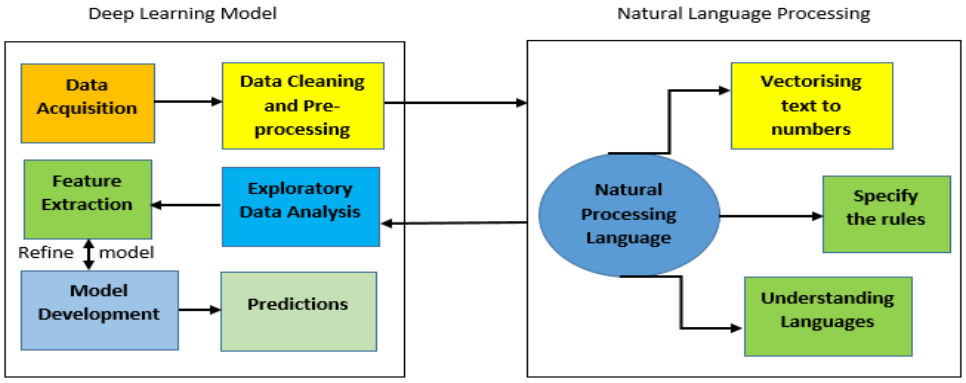

Figure 5: Architecture of Data Processing in the Hybrid Model

Prediction of the outcome of the proposed model: The final result is the predictions made by the hybrid DL model on unlabelled data. After being trained with input data sets, the algorithm knows how distinct kinds of data inside the context are interrelated with each other. And during the classification stage, an input value and its dependency variable are sent into the algorithms. This data is fed into the proposed algorithm, which predicts the appropriate result.

Algorithms used in the proposed model

```
Algorithm 1:
        Finding the best bit location in the host image for a secret data set.
        Input: Images selected from different sources.
        Output: Embedding capacity of each host image.
        Begin
            Divide the host image into four equal parts.
            Check the size of the host image and the size of secret data extracted using
            the CNN algorithm embedded in the image.
            Start sub iteration
                Implement the CNN-LSTM algorithm to Scan each part of the
                host image.
            End sub iteration
        Set the embedding capacity of the image.
End
```

```
Algorithm 2: Least Signified Bit (LBS) data hiding based on Blockchain PSO algorithm
and hash function
Input: Grey-scale host images, secret data extracted using the CNN algorithm.
Output: Stego images, hash for each block of secret data.
Begin:
    Convert secret data into binary.
    Cut secret data into blocks according to the host images' size.
    Start sub iteration 1:
        Calculate the hashing for the block of secret image data as the hash of the
        current secret data.
        Set the hash of each next block of secret data.
        Set the hash of the last block of secret data as N.
        Set the number of blocks of secret data.
        Set the number for each host image used for embedding this block of secret
        data.
        Set the number of the first host image used for meddling this block of secret
        data as the Genesis image.
        Implement the Blockchain PSO algorithm using Particle = [Direction X-offset, Y-
        offset, bit-planes, X-side length, Y-side length, data block number, host image
        number, Genesis image number, HC-SD, HN-SD, HL-SD].
        Scan each part of the host image based to hide the block of secret data.
        Hide the secret data of all particles in the last row of the host image.
  End sub iteration
Save all the hashes (Genesis, 2, 3, ..., N) in the ledger.
End
```

A convolutional neural network, or CNN, is a variation of a feed-forward architecture that was first used in video processing, recommendation systems, and the processing of natural languages. It is a sophisticated neural network

design that typically consists of pooling and convolutional or resampling layers that provide inputs to a fully-connected classification model. CNN models retrieve characteristics by filtering their input; the outcomes of many layers can be merged. The sub-sampling stages diminish pattern quality, which can improve CNN sensitivity to distortion and noise. Classifier tasks are carried out by fully connected layers.

Implementation of Natural Language Processing into the hybrid deep learning model

To integrate Natural Language Processing (NLP) in the hybrid Deep learning sequence, data processing, and data cleaning phases are added. But the learning, training, and testing approaches operate in a similar way as the DL approaches. The essential fundamental modification in the process is the conversion of textual into numerical values that the hybrid deep learning algorithm can use. There are numerous aspects that must be addressed throughout this process. For instance, the number of times a phrase appears may aid in determining the subject being discussed.

The dataset published from [15] first article to publish the standard IMDB data sets on film reviews for sentiment classification. This data set includes 40,500 reviews that have been binary labelled. The dataset is split up into training and testing in the ratio 80:20 cases respectively. With each data subset, the label distribution is balanced. 10% of the data set labelled from training examples were used as the testing datasets. Firstly the datasets were normalized by removing the irrelevant tags from the HTML.

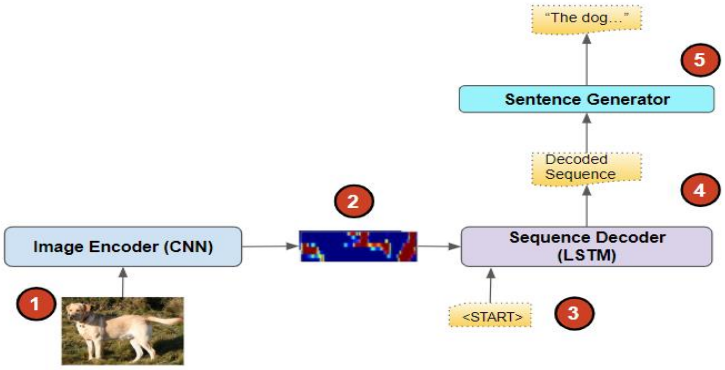

Figure 6: Simple scenario NLP with a proposed hybrid model to classify input

The pre-processing of the data sets was carried by the removal of the space, punctuation, unrelated words, and tokenization of the words. The analysis was done on the 2500 movie reviews, with 50% of them being positive and 50% being negative reviews in the sample data sets. Out 2000 samples were taken for training the model and 500 samples were taken to test the model. From the taken data set '1' is considered as the positive comment and '0' is considered as the negative comment. The proposed CNN-LSTM model take the input in the vector forms and changes the text into vectors by applying word2vec.

The performance of the proposed CNN-LSTM model was measured by using evaluation metrics as shown in the below equation. In the proposed model Adam optimizer is used to calculate the accuracy of the proposed hybrid model.

$$Precision\ (P) = \frac{\text{True Positive}}{\text{True Positive+False Positive}}$$

$$Recall(R) = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$F - measure = \frac{2 * (\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}}$$

IV. RESULT AND DISCUSSION

The proposed work developed a hybrid framework using three machine learning algorithms to automate the classification step. The fundamental idea involved is to automate the process of identifying the input category out

of three types; image, audio, or video. For this purpose, the deep learning algorithm concept is proposed so that the framework can work for all three categories of inputs; audio, video, and image.

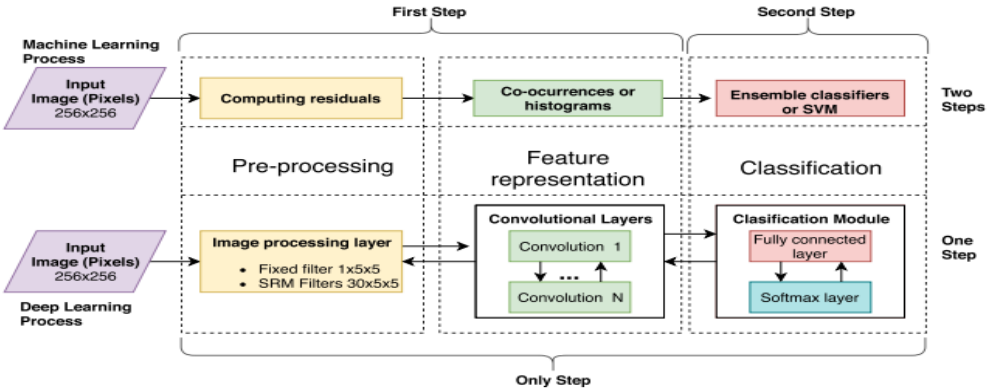

Figure 7: steganographic process of images using deep learning

V. CONCLUSION:

CNN assists in learning to retrieve data features. Moreover, numerous convolution layers are needed to record long-term correlations, and as the input sequence increases from the neural network expands capturing the correlations becomes very difficult. In this article, the Hybrid CNN-LSTM mode was proposed for the analysis of sentiment. Accuracy measurements, the proposed Approach hybrid CNN-LSTM model outperformed single CNN and LSTM models on two standard film customer testimonials datasets. In contrast to conventional deep learning as well as machine learning models, the proposed approach CNN-LSTM model successfully accomplished 91% accuracy.

REFERENCES

[1] Wilson, Akinola Antony. "Natural Language Processing: Machine Learning and Modeling Linguistics in the Domain of Sentiment Analysis." PhD diss., University of Nottingham, 2020.

[2] Sullivan, Florence R., and P. Kevin Keith. "Exploring the potential of natural language processing to support microgenetic analysis of collaborative learning discussions." *British Journal of Educational Technology* 50, no. 6 (2019): 3047-3063.

[3] https://www.ibm.com/in-en/topics/natural-language-processing

[4] Basu, Saikat, Supratik Mukhopadhyay, Manohar Karki, Robert DiBiano, Sangram Ganguly, Ramakrishna Nemani, and Shreekant Gayaka. "Deep neural networks for texture classification—A theoretical analysis." *Neural Networks* 97 (2018): 173-182.

[5] Dang, Nhan Cach, María N. Moreno-García, and Fernando De la Prieta. "Sentiment analysis based on deep learning: A comparative study." Electronics 9, no. 3 (2020): 483.

[6] Zhang, L.; Wang, S.; Liu, B. Deep learning for sentiment analysis: A survey. WIREs Data Min. Knowl. Discov. 2018, 8, e1253.

[7] Britz, D. Recurrent Neural Networks Tutorial, Part 1–Introduction to Rnns. Available online: http://www.wildml.com/2015/09/recurrent-neural-networkstutorial-part-1-introduction-to-rnns/.

[8] Hochreiter, S.; Schmidhuber, J. LSTM can solve hard long time lag problems. In Proceedings of the Advances in Neural Information Processing Systems, Denver, CO, USA, 2–5 December 1996; pp. 473–479.

[9] Alharbi, Ahmed Sulaiman M., and Elise de Doncker. "Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information." Cognitive Systems Research 54 (2019): 50-61.

[10] Collobert, Ronan, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. "Natural language processing (almost) from scratch." Journal of machine learning research 12, no. ARTICLE (2011): 2493-2537.

[11] Long, Haixia, Bo Liao, Xingyu Xu, and Jialiang Yang. "A hybrid deep learning model for predicting protein hydroxylation sites." International Journal of Molecular Sciences 19, no. 9 (2018): 2817.

[12] Fang X, Zhan J (2015) Sentiment analysis using product review data. J Big Data 2(1):5

[13] Elghazaly T, Mahmoud A, Hefny HA (2016) Political sentiment analysis using twitter data. In Proceedings of the ACM International Conference on Internet of things and Cloud Computing, pp.11

[14] Yanmei L, Yuda C (2015) Research on Chinese Micro-Blog Sentiment Analysis Based on Deep Learning. In IEEE 8th International Symposium on Computational Intelligence and Design (ISCID), pp. 358-361

[15] Singh J, Singh G, Singh R (2017) Optimization of sentiment analysis using machine learning classifiers. Human-centric Comput Inf Sci 7(1):32