International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Efficient Input File Classification by Applying Natural Language Processing and Hybrid Deep Learning Technique

Ayushi Chaudhary[1]*, Prof. Ashish Sharma[2]

Submitted: 18/01/2024 **Revised**: 27/02/2024 **Accepted**: 04/03/2024

Abstract: The digital era's advancements have prompted the adoption of communication as the primary medium for the corporate industry. Formerly, business discussions, profiles, conferences, purchasing, and settlements were all carried out in person. But the modern era has made everything digitized. In the past few years, it's been observed there is an exponential increase in the count of complicated manuscripts and writings that need a better recognizing of machine learning methodologies to successfully detect languages in various purposes. Several Artificial Intelligent approaches have produced outstanding achievements in processing natural languages. The ability of various machine learning and deep learning to realize complex models and non-linear associations within data is critical to their efficiency. Learning applicable frameworks, architecture, and algorithms for input classification, such as text files, audio, and video files, on the other hand, is a challenging task. Objective**:** This study aims at Natural Language Processing in the identification of text, voice messages, smart records, and chatbots. Hybrid deep learning approach for the classification of the inputs that are in the form of text, voice, and video records. Problem Statement: As interaction becomes more crucial to business, firms have designed sophisticated NLP programs. These NLP take human wishes and satisfy them quickly through messages, telephone calls, digital records, and chatbots. The ease of communication and connection has shown a stronger impact on customer preferences, aspirations, and demands. Contemporary service providers today utilize email, messaging, telephone calls, digital records, and chatbots as primary points of contact for practically all of their transactions, client inquiries, and preferable trade channels. Method**:** The study uses text content, voice message, and audio as part of Natural Language Processing and Hybrid Deep Learning approaches to demonstrate how input is processed depending on user reactions, replies to text messages, and audio record identification during communications.

Keywords: *Artificial Intelligence, Natural Language Processing, Machine Learning, Deep Learning, chatbots.*

1. Introduction

Natural Language Processing, or NLP, [1] 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. 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.

[1]*GLA University*
ayushichaudhary11@gmail.com
[2]*GLA University*
ashish.sharma@gla.ac.in

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 Deep Learning Approach:

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. 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.

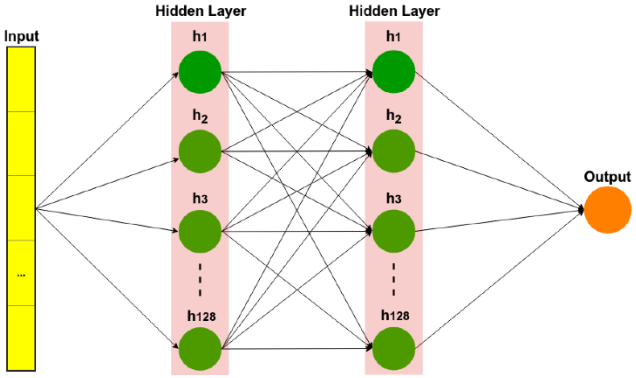

Fig 1: General Architecture of Deep Learning Approach

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.

Convolution Neural Networks (CNN):

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. 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. Figure 2 depicts an instance of a CNN model.

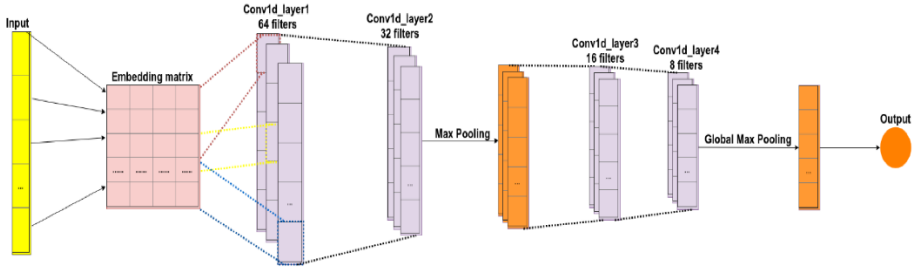

Fig 2: Architecture of Convolutional Neural Network

The data collected is pre-processed in order to restructure it for use with the embedded network. The diagram depicts an input-embedded pattern that has been processed by 4 convolution operations and two layers with

maximum pooling. The initial two convolutional use sixty-four and thirty-two layers to train distinct characteristics, respectively; those would be preceded by a layer for maximum pooling, which is intended to minimize the complexities of the outcome and avoid overfitting the input. The next two convolutional layers, which are followed by a maximum pooling layer, contain 16 and 8 filters.

Recurrent Neural Network (RNN):

Recurrent neural networks [7] are neural networks in which neuron connections establish a unidirectional cycle, resulting in feedback mechanisms inside the RNN. RNN's primary role is to analyze contextual data using internal storage collected by targeted cycles. Long short-term memory (LSTM) is a form of RNN that can use long memory as the input for the training algorithms in the hidden neurons. Hochreiter and Schmidhuber (1997) [8] first proposed this. Figure 4 depicts an LSTM architectural model. The input information is pre-processed in order to restructure it for use in the embed matrices. The LSTM, which has 200 cells, is the following layer.

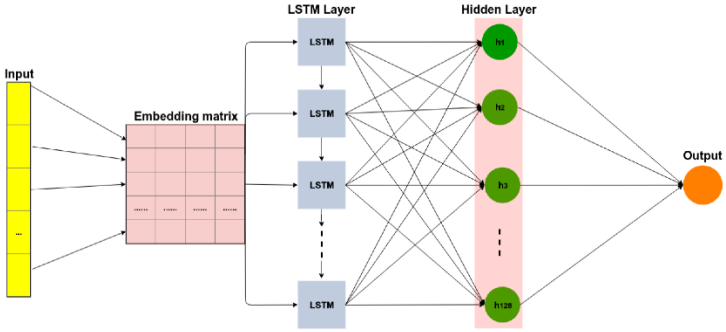

Fig 3: Architecture of Long Short Term Memory

2. 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.

3. 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:

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 o is nothing but data collection. The obtained data can then be saved in the workspaces as Dataframes, 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 analyzing 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.

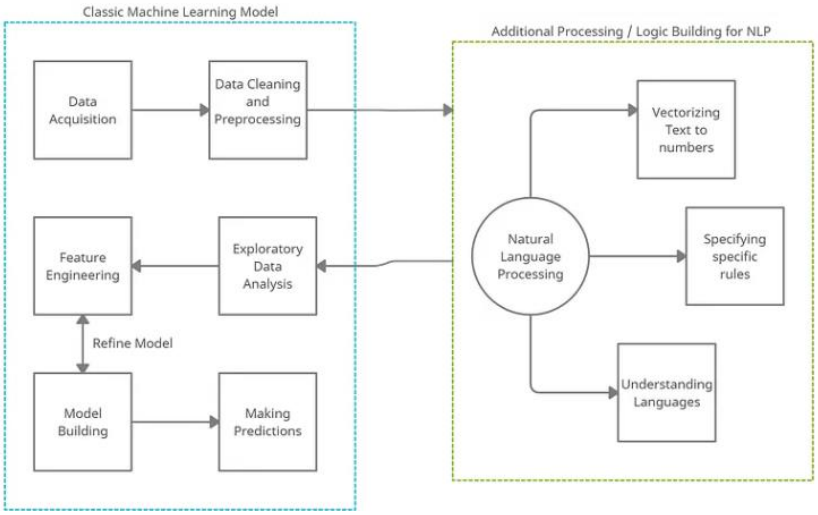

Fig 4: Architecture of Data Processing in the Hybrid Model

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.

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.

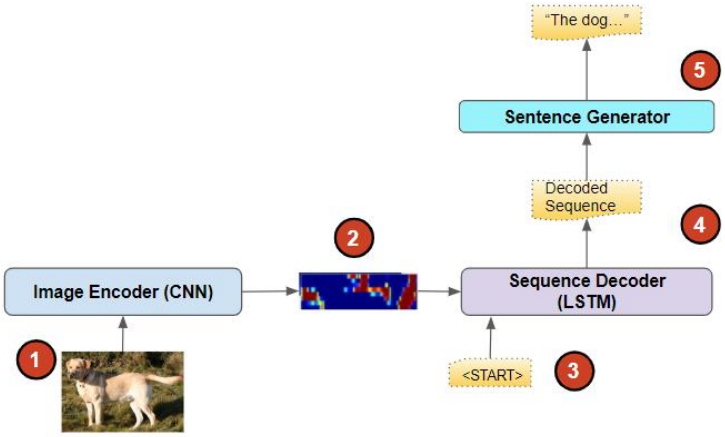

Fig 5: Simple scenario NLP with proposed hybrid model to classify input

References:

[1] Wilson, Akinola Antony. "Natural Language Processing: Machine Learning and Modelling 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.; Schmid Huber, 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 behavioural 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.