International Journal of Academic Pedagogical Research (IJAPR)
ISSN: 2643-9123
Vol. 5 Issue 3, March - 2021, Pages: 17-23

English Vocabulary Retrieval And Recognition Based On Fpga And Machine Learning.

Nigora Abdiyeva Muzaffar qizi[1], Jurabek Abdiyev[2]

[1]Samarkand State Institute of Foreign Languages, Uzbekistan, Samarkand, Bustonsaroy street 93, 140104 Samarqand
[2]Physical-technical Institute of NPO "Physics – Sun" of Uzbekistan Academy of Sciences Uzbekistan, Tashkent, Chingiz Aitmatov street 2B.
Corresponding author: fiztexabdiev@gmail.com (J. Abdiyev)

Abstract: *Digital audio spoken vocabulary retrieval and recognition have been developed in recent years. This growth has prompted extensive research into technology and reliable indexing. The English vocabulary content retrieval requires a combination of audio and speech processing technology and information retrieval. A study of English content search initially investigated units of planning to have a similar audio signal structure, and the subsequent focus outside the audio, naturally, more informal oral content. Shifted to generate conversational settings and the volume provided by this study outlines the component technical discipline of the relationship between lexical speech signal recognition and user interaction issues. The aim is for researchers with a background in speech technology to seek a deeper understanding of whether these fields are integrated into supporting research and development to solve the core challenges of vocabulary and content search. This study describes a content-based search method for Machine Learning based Neural Network classifiers that retrieve relevant documents efficiently and accurately. This overcomes the system limitations of characterization based arrangements regarding restricted vocabulary and preparing information accessibility. To analysis the system, it has designed a one-time learning neural network for classification and achieves better accuracy results.*

Keywords – Vocabulary Retrieval, Recognition, Machine Learning, Field Programmable Gate Array (FPGA), Neural Network (NN).

1. INTRODUCTION.

English language vocabulary retrieval and recognition depend on the signal processing of the voice. Language securing is an exceptionally amazing cycle; however, there is proof that users create language dangers that are a measure of approaching contribution from grown-ups around them. The language is automatically developed for most users. It verified that specific psychological capacities, for example, working memory, discourse preparing methods, and morphological insight, were straightforwardly connected to user vocabulary development. Environ-mental factors and their cognitive abilities are responsible for the user early vocabulary growth. In any case, this advancement originates from social and educational chances and collaborations. The standard framework can't appropriately advance all networks, those experiencing different inabilities.

The most used languages lack the word's sound when it appears on the system, so people with heavy accents in their native language often misread the word. Due to changes in word pronunciation, audio description systems tend to misunderstand commands and produce unwanted reactions. Altogether, for these frameworks to be proficient and helpful for a wide range of users, these frameworks need to comprehend the complement and consequently create the desire for a reaction to the order. This article analyzes and extracts unique Gujarati accents, considers negative British accents and English characteristics and classifies Indian accents. Using this tag set for British and Indian accents, the classification model is trained in classification accents. It is used to train models such as art learning machines, support vector ma-chines and classification of random forest algorithms, and to test the state of their test sets. Since it is a binary classification question of whether the accent is India or not, its purpose is the distance used to separate the planes that separate the two classes and the NN that maximizes the distance from the boundary.

2. LITERATURE SURVEY.

In this work, performance characteristics are the most important issues and need to be improved. Finally, samples with reduced inputs are identified using Multi-Layer Perceptron (MLP) and Bayesian regularity [1,2]. Recently, Convolutional Neural Networks (CNNs) have proven particularly capable of learning from image data, their powerful learning abilities. In this work, the company's online handwriting function's learning ability is investigated by building various CNN architectures.

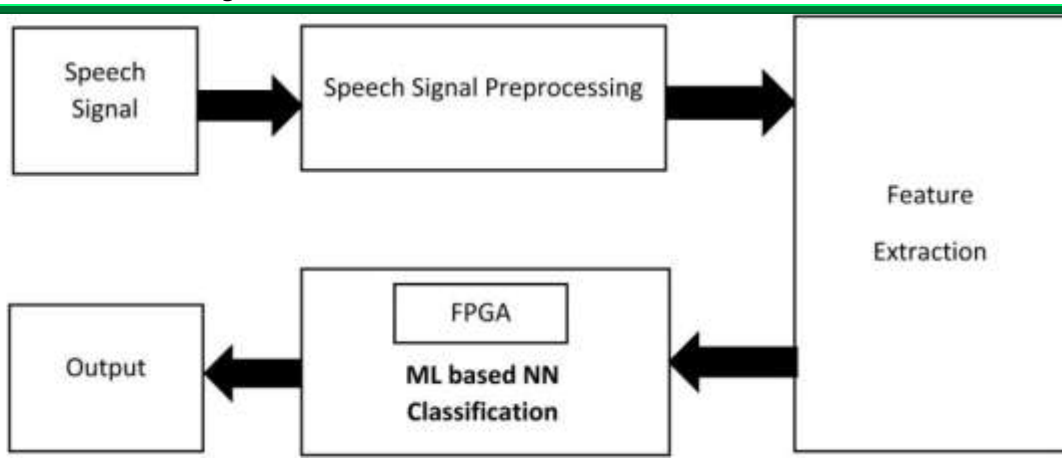

Fig. 1. Proposed block diagram.

Well-developed cellular neural networks can use architecture to directly process online handwriting, unlike online hand-writing, which transforms existing works into signals [3,4].

In current voice colleague frameworks, words expressed by individuals with substantial accents impacted by their first language are regularly misconstrued. AI calculations, particularly Space Vector Ma-chines (SVMs) and irregular woods applications, play an essential role in classifying accents [5,6] on the right training set. Taking in another vocabulary from a setting, for the most part, requires numerous contacts, meanwhile, where the importance of a word can be looked at from memory or surmised from the specific situation. It compares the effects of short-term and long-term retention of memory acquisition and contextual inference [7,8].

It describes an efficient and accurate search method based on the classification of related documents. It uses classification search terms by SVMs and believes that classification-based solutions may be better than the arrangements in numerous functional circumstances. This conquers the down to earth impediments of [9,10] classifier-based structures regarding restricted jargon backing and preparing information accessibility. Test effects refer to discovering that search practices lead to better long-term storage in textbooks than further research. In this examination, it explored whether this finding could be reached out to primary school vocabulary learning. It likewise controls the learning word setting [11,12].

Term identification and organization are the most important steps. At the same time, the domain knowledge of any domain is organized because it is a term that defines the basis of conceptual knowledge and their interrelationships [13,14]. It proposes a new method of measuring discriminatory confidence based on the evolutionary strategy. Our evolutionary algorithm has evolved directly to optimize such objective functions, based on a particular class of coding and some traditional discriminant models. Optimizes classification errors, which is a significant advantage. Master development blueprint and support vector ma-chine [15,16]. As the demand for communication in English increases, different learning support techniques are proposed and given to the user. The numerous understudies think it's troublesome, read the composition, and communicate in English [17,18].

3. MATERIALS AND METHODS.

The proposed Machine Learning based Neural Network system learns the framework, analysis, and classification of typical language accents. Accent expressions focus on a specific vocabulary with various blends of consonants and vowels that influence how to express the language. It helps to understand the properties that can be used to identify good accents.

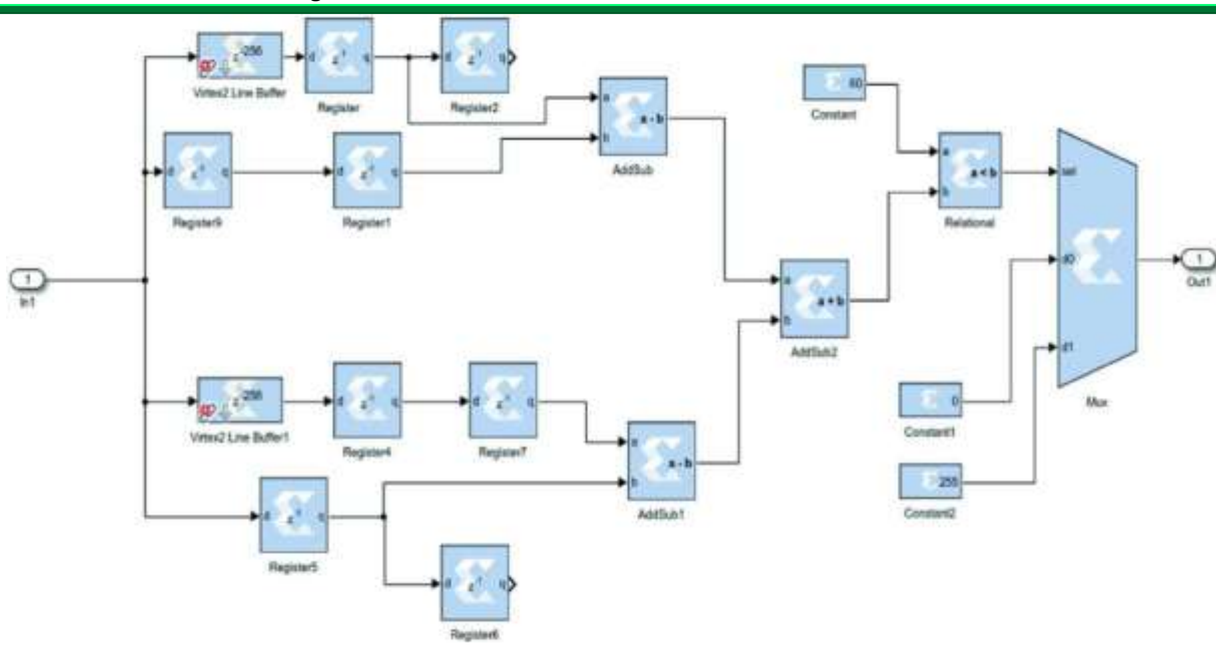

Fig. 2. Proposed signal preprocessing circuit.

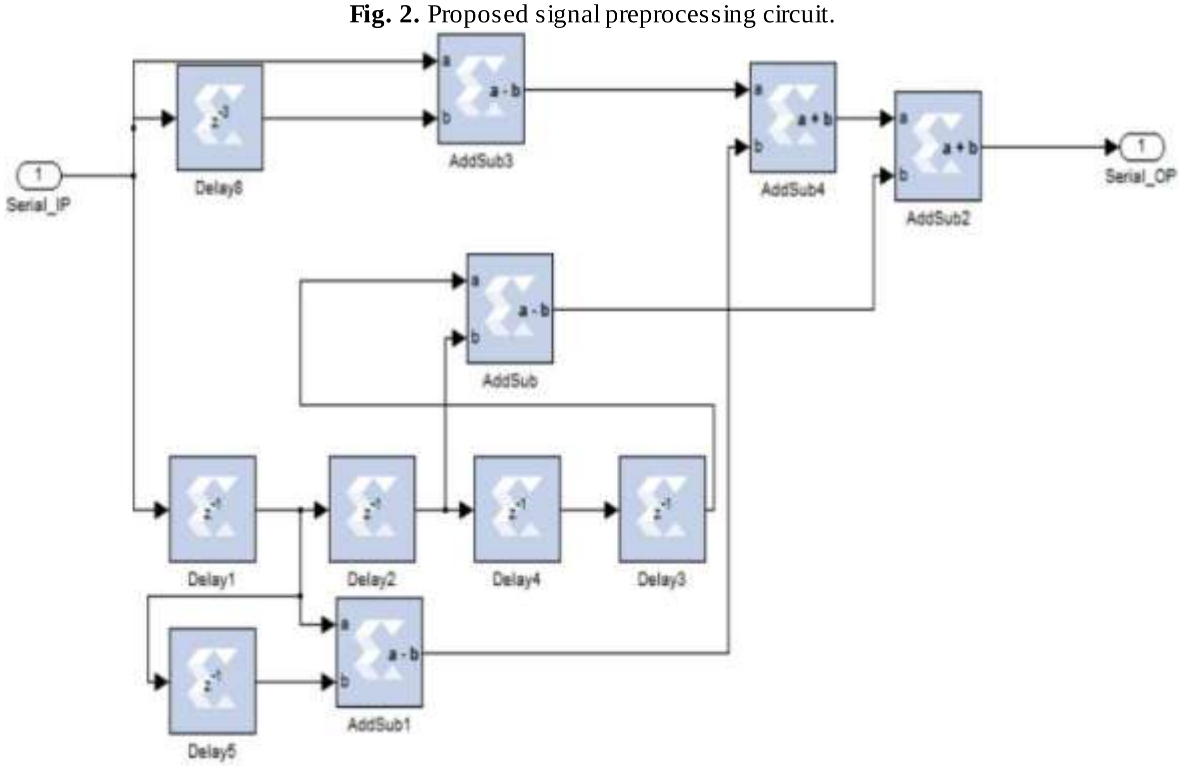

Fig. 3. Feature extraction circuit diagram.

Still, it uses complicated vowel/consonant combinations, so the work is different from previous work in accent analysis.

Fig. 1 gives the details of the proposed system structure. The relevant features obtained are believed to be incorporated into speech recognition and can be used to improve accents' speech intelligibility and generate appropriate responses based on the speech digital assistant system.

3.1. Speech signal preprocessing

Preprocessing a voice signal, separation of the silent region of the captured signal from the silent/unvoiced area is usually described as an important step in developing reliable speech or speaker recognition systems. Most of the speech, speakers, and certain properties are present in the audio signal's voiced part. Also, silent and non-speech areas, which are marked and removed with computational complexity at a later stage, lead to a significant reduction in the extracted speech signal's vocalization. Audio signals are divided into silent/unvoiced and voiced regions and are used to isolate the fundamental frequency estimation of the audio signal, formal extraction or syllable marking, stop sound identification, endpoint detection description, and other applications and the preprocessing circuit is shown in Fig. 2.

Step 1: Initialize the set of input signal data.

Step 2: Find the values of each to attribute the system. Analysis value of the input speech signal.

$$x(n) = s(n) + d(n) \quad (1)$$

Where, $s(n)$ is speech clean signal; $d(n)$ is ambient noise; $x(n)$ is output speech signal

Step 3: Compute the overlap area of the function by using to find the speech Signal data.

Step 4: For each correlation coefficient, value more significant than the input speech signal.

Step 5: The output speech Signal value is increased, so the output speech signal's quality is enhanced.

Step 6: End.

Compared to simply a saved pattern feature array, when that hap-pens, they find a perfect match, and if the voice pattern remains un-changed. However, speech recognition's main problem is that the speech signals are very diverse to different speakers at different rates, different content, and various acoustic conditions. The task is to deter-mine that it is not related to speech recognition changes related to changes in speech.

3.2. Feature extraction.

In principle, it has the option to perceive discourse legitimately from the digitized waveform. Nonetheless, it is fitting to play out some component extractions that decrease fluctuation because of huge changes in the sound signal. In particular, it excludes various sources, such as whether such sounds are voiced or unvoiced. If voice, it calcu-lates the short-term spectrum, excluding the period or interval of the excitation signal and the fundamental frequency. The speech analysis reveals a non-linear frequency scale. The approximation of this scale is approximately a logarithm, followed by a linear frequency signal. Therefore, it will perform feature extraction, a very common spectrum calculation, and the frequency axis's bending frequency. Feature extraction is the process of a small amount of audio data extracting that can represent each word in the future. Feature matching involves the actual process of identifying new words by comparing features extracted from his / her voice input with a set of words called them.

The Feature Extraction Circuit diagram is shown in Fig. 3. The first one is the measurement parameter of the input signal with the help of the sensor. Second, the analog to digital converter is used to convert the signal and given top priority encoder while the other is the testing phase. It has a wide impact on system performance as it is to design an important task for speaker recognition to extract the best parameter characteristics of the acoustic signal. In highlight extraction, the infor-mation waveform is changed into a voice include vector arrangement.

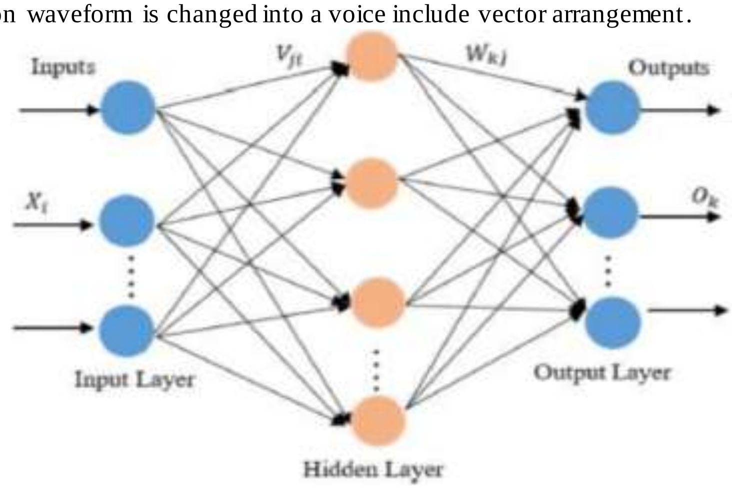

Fig. 4. Proposed neural network diagram.

Every vector speaks to data about the sign in a little league window.

3.3. Neural networks classification.

In the Neural Network plan, the quantity of conceivable organization structures develops exponentially with the number of layers. The objective is to locate the best neural organization structure that limits the blunder rate. The cost

International Journal of Academic Pedagogical Research (IJAPR)
ISSN: 2643-9123
Vol. 5 Issue 3, March - 2021, Pages: 17-23

work is NN's accuracy utilizing fixed-length paired strings to speak to each organize structure and propose exqui-site coding of chromosomes. The NN string speaks to the convolution layer. Every age uses standard hereditary designing of new people to dispose of the powerless. Its acknowledgment precision surveys singular quality got when devouring preparing the organization and assessing it in the approval set. With the current widespread use of convolutional neural networks, it has become possible to solve problems in the field, implicit information that it is not easy to express knowledge explicitly in the original data. The NN layer is shown in Fig. 4.

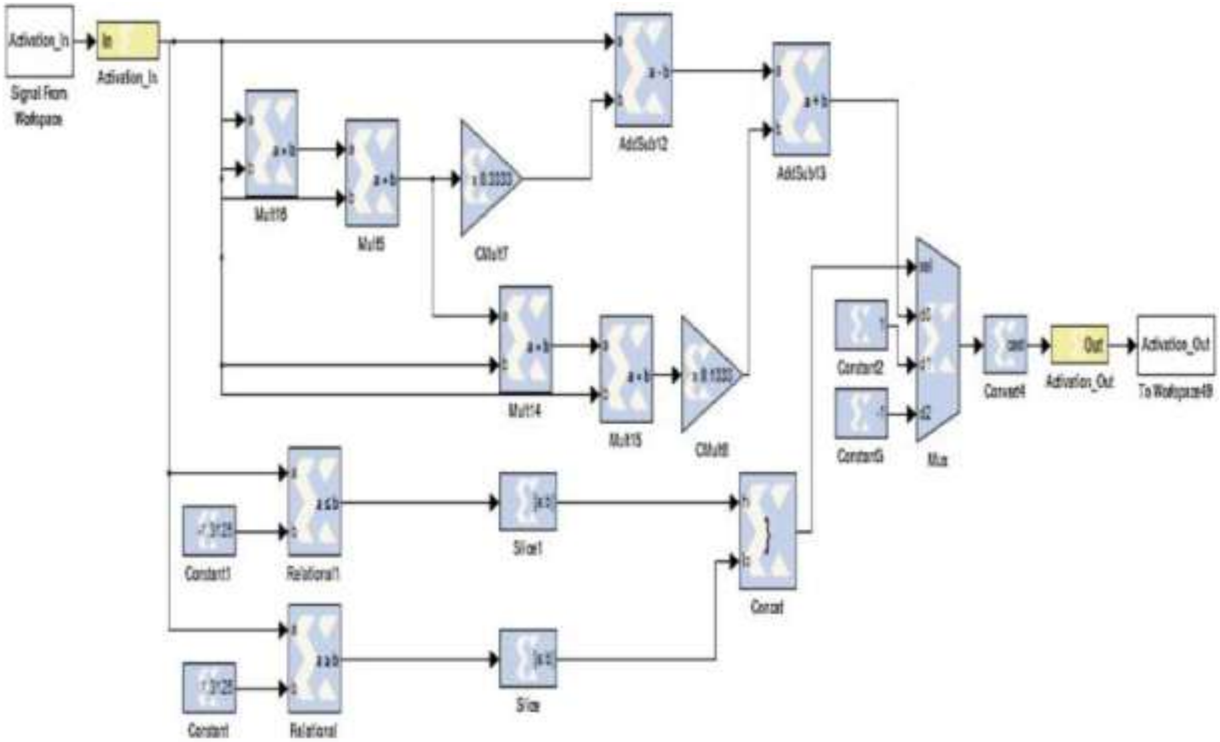

Fig. 5. FPGA processing element circuit diagram.

This is primarily possible with the excellent ability of deep learning to find complex structures in high-dimensional data. The solutions to various problems in fields such as science and business may have been unimaginable for years, despite the best attempts in the AI community.

3.3.1. Field programmable gate array (FPGA).

FPGAs are off-the-rack programmable gadgets that give an adaptable stage to empowering custom equipment abilities at low advancement costs. These fundamentally incorporate many programmable legitimate units called configurable rationale blocks (CLBs), a programmable interconnect network and many programmable info/yield units that encompass the gadget. Likewise, they are Digital signal preparing (DSP) blocks (queries) used to perform number juggling escalated activities, for example, duplication and capacity, RAM, lookup tables, flip-flops, and so forth. It has a rich arrangement of installed parts, and so on FF), clock the executives unit, fast I/O interface, and so forth.

FPGAs are widely regarded as accelerators and compute-intensive applications because they can be parallel to flexible fine-grained. Related business models for broadcast and collective response. Data flows through associative calculations and is presented by the application's accelerated FPGA calculation model, using and including standard hardware structures such as high parallel memory access, such as first- in-first-out (FIFO) buffers, stacks, and priority queues.

FPGAs benefit from boosting power utilization per watt and enormous diminishing scope working expenses are shown in Fig. 5. This is a magnificent decision for battery-fueled gadgets for quickening agents and cloud administrations on workers since they are huge. FPGAs have, as of late, been broadly utilized for profound learning quickening given the adaptability of the execution architecture, mainly leading to high execution speeds in parallel.

4. RESULTS AND DISCUSSION.

MATLAB R2017a is utilized in this research work of speech signal recognition based on Neural Network (NN) with the Field Programmable Gate Array (FPGA) help. The MATLAB is a graphical user interface tool that is used for training and simulation.

Fig. 6 and Table 2 give the performance analysis of the proposed Machine Learning based Neural Network (NN) system with the existing method.

Fig. 7 and Table 3 gives the false ratio, and it's compared to the existing method. The proposed Machine Learning based on the false ratio reduces the neural Network method.

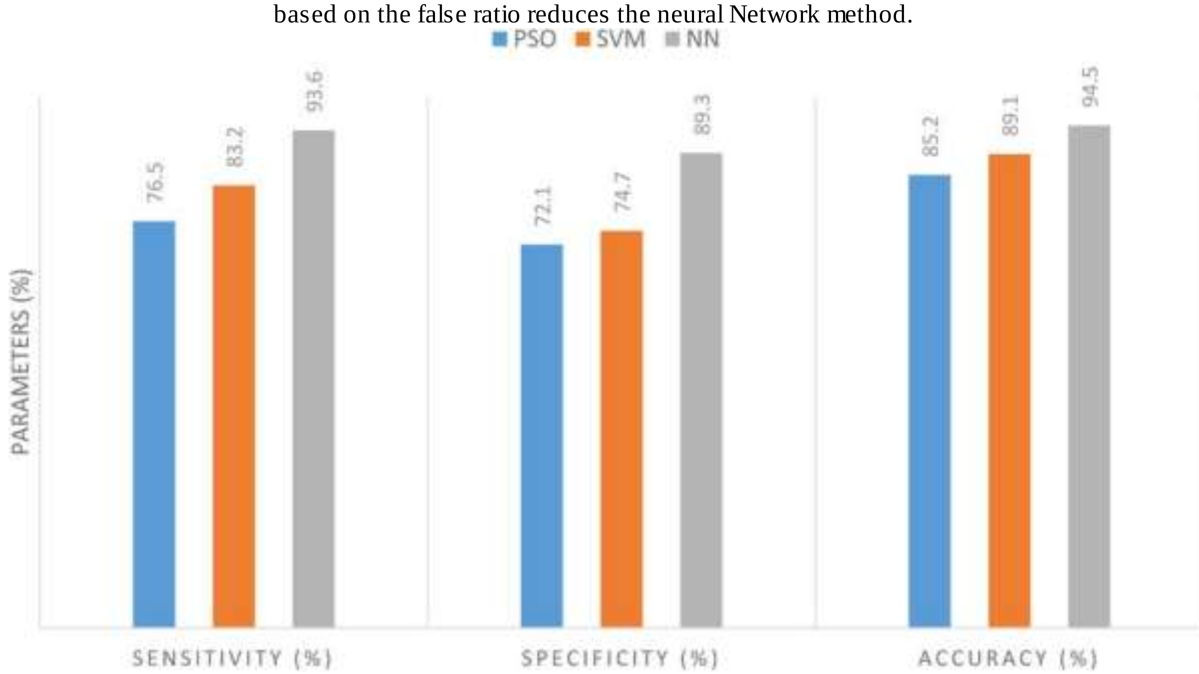

Fig. 6. Proposed machine learning based neural network (NN) comparison result.

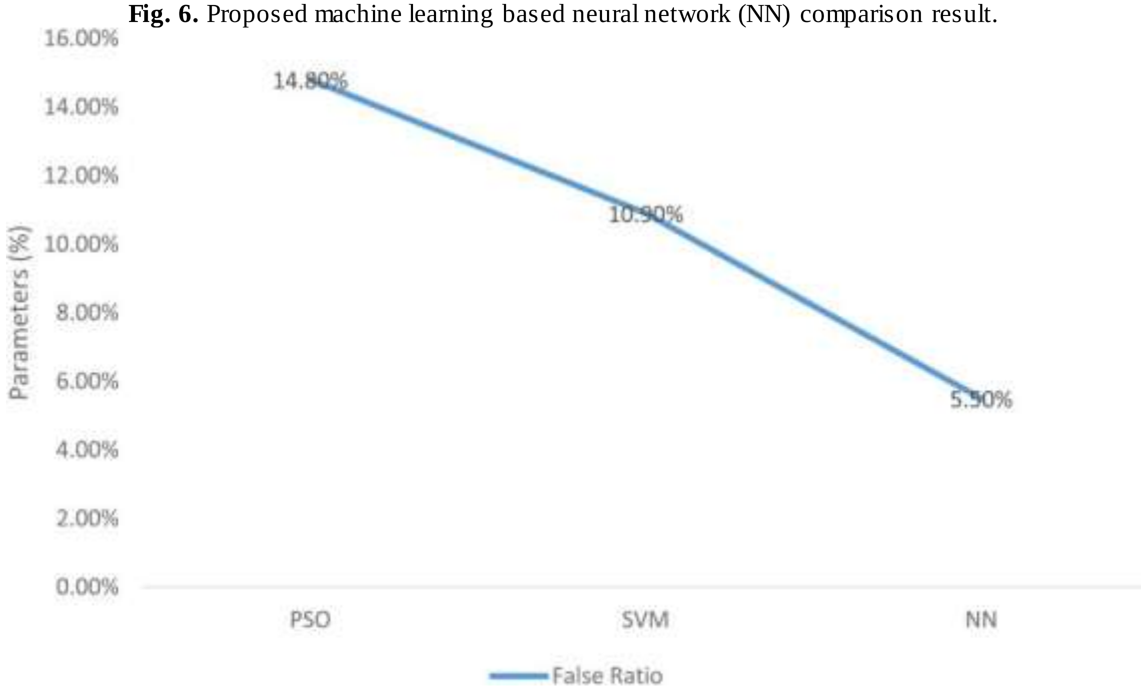

Fig. 7. Comparison chart false ratio.

5. CONCLUSION.

The proposed Machine Learning based Neural Network (NN) system is obtained point out that pitch and high peak frequencies feature for the classification of native English accents. It should be noticed that the speaker's normal estimation in the user pronunciation is higher than the language emphasis, and the speaker's recurrence estimation is a higher priority than the language highlight. Limited to a specific part of the speech's word frequency is considered high value in English with a user accent other than English vocabulary. These parameters in generating accent commands can effectively improve these systems' intelligibility when incorporated into the learning model of speech recognition sys-tems. Network training has been tested and completed with real-time voice datasets from speakers' variable speaker recognition database. The proposed Machine Learning based Neural Network (NN) gives a better performance of accuracy is (94.5%) of this system.

References

[1] N. Kharibam, Automatic Speaker Recognition from Speech Signals Using Self Organizing Feature Map and Hybrid Neural Network, Microprocessors and Microsystems 79 (2020).

[2] X. Yongqiang, Real time speech recognition algorithm on an embedded system based on continuous Markov model, Microprocessors and Microsystems 75 (2020).

[3] S. Mandal, Mahadeva, Exploration of CNN Features for Online Handwriting Recognition, in: 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 2019, pp. 831–836, https://doi.org/ 10.1109/ICDAR.2019.00138.

[4] A. Sandipa, A new Assistive Technology in Android Platform to aid Vocabulary Knowledge Acquirement in Indian Sign Language for better Reading Comprehension in L2 and Mathematical Ability, in: 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2019, pp. 408–413, https://doi.org/10.1109/SPIN.2019.8711619.

[5] M. Thakkar, S. Elias, Speech Recognition Learning Framework for Non-Native English Accent, in: 2019 International Conference on Data Science and Engineering (ICDSE), Patna, India, 2019, pp. 84–89, https://doi.org/10.1109/ ICDSE47409.2019.8971486.

[6] M. Pleva, Yuan-Fu, Towards Slovak-English-Mandarin Speech Recognition Using Deep Learning, in: 2018 International Symposium ELMAR, Zadar, 2018, 151–154, https://doi.org/10.23919/ELMAR.2018.8534661.

[7] A. Gesa, Contextual Richness and Word Learning: Context Enhances Comprehension, but Retrieval Enhances Retention, Journal of Research in Language Studies (2018) 546–585, https://doi.org/10.1111/lang.12285.

[8] T. Nakata, Effects of retrieval formats on second language vocabulary learning, 54, IEEE, 2016, pp. 257–289.

[9] G. Viresh, Document Retrieval with Unlimited Vocabulary, in: Winter Conference on Applications of Computer Vision, 2015, pp. 741–748, https://doi.org/10.1109/ WACV.2015.104.

[10] J. Schmidtke, Second language experience modulates word retrieval effort in bilinguals: evidence from pupilometer, frontiers in psychology (2014), https://doi. org/10.3389/fpsyg.2014.00137.

[11] G. Nicole, The Effect of Retrieval Practice in Primary School Vocabulary Learning, Applied Cognitive Psychology (2013), https://doi.org/10.1002/acp.2956.

[12] P. Krishnan, Jawahar, Bringing Semantics in Word Image Retrieval, in: International Conference on Document Analysis and Recognition, 2013, 733–737, https://doi.org/10.1109/ICDAR.2013.150.

[13] T. Mukhtar, H. Afzal, Vocabulary of Quranic Concepts: A semi-automatically created Terminology of Holy Quran, in: International Multi-topic Conference, 2012, pp. 43–46, https://doi.org/10.1109/INMIC.2012.6511467.

[14] X. Cui, J. Huang, Multi-View and Multi-Objective Semi-Supervised Learning for HMM-Based Automatic Speech Recognition, IEEE Transactions on audio, speech, and language processing (Sept. 2012) 1923–1935, https://doi.org/10.1109/ TASL.2012.2191955.

[15] J. Tejedor, D. Wang, An evolutionary confidence measurement for spoken term detection, IEEE, 2011.

[16] Y.-X. Yue-Jie, Multi-Feature Representation for Web-Based English-Chinese OOV Term Translation, in: International Conference on Machine Learning and Cybernetics, 2011, https://doi.org/10.1109/CBMI.2011.5972537.

[17] J. Kazuyuki, Construction of English Composition Support System Based on Conveying Emotion, in: 2009 International Conference on Natural Language Processing and Knowledge Engineering, Dalian, 2009, pp. 1–8, https://doi.org/ 10.1109/NLPKE.2009.5313802.

[18] L. Yu, Vocabulary Recognition and Memorization: A Comparison of Two Methods, School of Teacher Education English, 2011.