Neutrosophic Sets and Systems, Vol. 55, 2023
University of New Mexico

# Neutrosophic speech recognition Algorithm for speech under stress by Machine learning

D. Nagarajan[1], S. Broumi[2,3] , Florentin Smarandache[4]

[1] Department of Mathematics, Rajalakshmi Institute of Technology, Chennai, Tamil Nadu, India.
[2]Laboratory of Information Processing, Faculty of Science Ben M'Sik, University Hassan II, B.P 7955, Sidi Othman, Casablanca, Morocco,

[3]Regional Center for the Professions of Education and Training(C.R.M.E.F), Casablanca-Setat, Morocco.

[4] The University of New Mexico Mathematics, Physics, and Natural Science Division, 705 Gurley Ave., Gallup, NM 87301, USA;

E-mail: *dnrmsu2002@yahoo.com, broumisaid78@gmail.com, smarand@unm.edu

Corresponding author:D.Nagarajan,dnrmsu2002@yahoo.com

Abstract
It is well known that the unpredictable speech production brought on by stress from the task at hand has a significant negative impact on the performance of speech processing algorithms. Speech therapy benefits from being able to detect stress in speech. Speech processing performance suffers noticeably when perceptually produced stress causes variations in speech production. Using the acoustic speech signal to objectively characterize speaker stress is one method for assessing production variances brought on by stress. Real-world complexity and ambiguity make it difficult for decision-makers to express their conclusions with clarity in their speech. In particular, the Neutrosophic speech algorithm is used to encode the language variables because they cannot be computed directly. Neutrosophic sets are used to manage indeterminacy in a practical situation. Existing algorithms are used except for stress on Neutrosophic speech recognition. The creation of algorithms that calculate, categorize, or differentiate between different stress circumstances. Understanding stress and developing strategies to combat its effects on speech recognition and human-computer interaction system are the goals of this recognition.

Keywords: speech recognition, categorization of stress in speech, linguistic technology, Neutrosophic, Machine learning

## 1.Introduction

In order to produce speech, a series of intricately synchronised articulator movements, respiratory system airflow, and timing of the vocal system physiology are all required. While the posture of the articulator changes to create speech, not all utterances made by a speaker will be identical in every way. This is due to the fact that the subject is frequently experiencing some sort of emotional stress, which will affect the utterance and cause an error in the articulator motions. Listeners can handle or interpret these subtle variations in human communications much better than the automatic human-machine interface. The features of stress, and its effects on human speech production, perception, and automatic speed systems, are still not fully understood. Speech is therefore a complex signal that contains information about the speaker. The speaker's intent, language history, features of their accent and dialect, and additional paralinguistic information. Stress can cause a change in speech output that can large and will consequently affect how well speech processing apps function [1],[2]. Numerous research has examined how stress affects speech production variability[3],[4], and [5]. Moreover, a stress-based expansion of multi-style training Additionally, token generation has improved anxious speech recognition [6]. Then, five stress-sensitive targeted feature sets are chosen.

stress situations such as the cockpit of an Apache helicopter, anger, clarity, the Lombard effect, loudness, etc. features that are frequently employed Include the cepstral characteristic for speaker identification [7]. When doing cepstral analysis, speaker recognition software often ignores the excitation source data that appears as a high-time component of the cestrum[8]. The Mel-Frequency Cepstral Coefficient, a phonetic characteristic, was retrieved from the voice signals, and the stress was identified using a neural network that was programmed into the system using Python[9]. serve as a resource for decision-makers in many real-world scenarios and application domains, particularly from a technical standpoint, for both academic and business experts[10]. In the research described in this paper, stress during applicant screening interviews is identified via voice analysis. The mean energy, mean intensity, and Mel-Frequency Cepstral Coefficients are employed as classification features in machine learning to identify stress in speech[11]. This study uses an EEG signal to suggest a stress classification system. 35 individuals' EEG signals were analyzed after being collected using a commercially available 4-electrode Muse EEG headgear with four EEG sensors [12]. In this study, it is expected that risk factors would, both cross-sectionally and longitudinally, predict mental health issues after controlling for sociodemographic traits and intent to become pregnant[13]. This investigation uses brain signals to look at how stress levels are affected by English and Urdu language music tracks[14]. This project looks into methods for sensing stress that is used to identify hardware[15]. The high-level features are combined into one unified representation using a proposed model-level fusion technique, which classifies the stress states into baseline, stress, and amusement[16]. The heart rate was measured and classified into three categories of positive, negative, and neutral emotions using the Geneva affective picture database. The support vector machine is a machine learning technique that has been built to predict the mental stress situation from the measured heart rate[17]. The development of a model for measuring stress levels makes use of several sensors, including those that measure body temperature, blood pressure (BP), heart rate, and CO2 concentration[18]. Studies show that combining IoT and AI with deep learning (DL) technology makes it possible to take preventative measures. Recognise stress well before its effects on human health become apparent[19]. In order to assess teaching effectiveness, enhance education, and limit risks from human errors that could occur as a result of workers' stressful circumstances, stress detection is crucial in both education and industry [20]. has good classification performance in this study and is able to gauge the stress levels of kids with accuracy. The growth of students' mental health has a strong foundation thanks to the precise measurement of stress, which also has important practical ramifications[21]. This research explores the concept of the intervention effect of physical activity on college students using an integrated evaluation-based algorithm. College students are used as an example of stress groups. The findings indicate that regular physical activity can significantly reduce college students' stress levels[22]. This study employs Neutrosophic logicto provide a valid ranking of hospital construction assets based on their changeable criticality and to lessen the subjectivity pertaining to expert-driven judgements[23]. This document compiles all research on machine learning mapping.

methods from the sharp number space to the neutrosophic environment. We also talk about contributions and combining single-valued neutrosophic numbers with machine learning methods

Modeling faulty information using (SVNs)[24]. In this paper, a brand-new paradigm for incorporating neutrosophy into deep learning models is given. To further comprehend the feelings, we quantified them using three membership functions as opposed to simply predicting a single class as the outcome. The two components of our suggested model are feature extraction and feature categorization[25]. The proposed framework would be an appropriate progression in the future by eliminating ineffective qualities through feature selection [26]. Stress is a psychological condition that results from an alleged threat or work demand and is accompanied by a variety of feelings. Finding linguistic cues of stress could be one of the verbal signs of stress. verbal indicators of stress are perceived by the listener, markers range in visibility from very visible to invisible. Consciously and unconsciously, these signals are watched continuously [27]. Speech recognition is the ability of a system to recognise the words and phrases of the speech and convert them to readable or written format. Speech recognition is typically carried out through processes including call routing, speech-to-text conversion, voice dialling, voice audibility, and language modelling. Although there are numerous techniques and algorithms for voice recognition, none of them is handling all factors including word length, speaker independence, a wide vocabulary, comprehension of speech, time complexity, noisy surroundings, and conversational speech. Neutrosophic can be integrated to analyse the acoustic signal of an unknown speaker and the decision-making process when indeterminacy occurs, respectively, to solve these issues.

## 2. Preliminaries

$A$ neutrosophic set $\tilde{\mathcal{A}}_N$ in $\mathcal{U}$ (Universe of discourse) is catogorized as functions of a truth membership $T_{\tilde{\mathcal{A}}_N}(g)$, an indeterminacy membership $I_{\tilde{\mathcal{A}}_N}(g)$ and a falsity membership $F_{\tilde{\mathcal{A}}_N}(g)$and is given by

$\tilde{\mathcal{A}} = \{g, \langle T_{\tilde{\mathcal{A}}_N}(g),\ I_{\tilde{\mathcal{A}}_N}(g), F_{\tilde{\mathcal{A}}_N}(g)\rangle \mid g \in \mathcal{U}\}$.

Here $T_{\tilde{\mathcal{A}}_N}(g),\ I_{\tilde{\mathcal{A}}_N}(g), F_{\tilde{\mathcal{A}}_N}(g) \in [0,1]$ and the relation $0 \leq \sup T_{\tilde{\mathcal{A}}_N}(g) \leq \sup I_{\tilde{\mathcal{A}}_N}(g) \leq \sup F_{\tilde{\mathcal{A}}_N}(g) \leq 3$ holds for all $g \in \mathcal{U}$.

*Definition 2.1[27,28 and 29]*

Let $X$ be the universal set, then Neutrosophic set is defined as $S = \{(T_S(x), I_S(x), F_S(x)), x \in X\}$ where $T_S(x), I_S(x), F_S(x) \in [0,1]$ and $0 \leq T_S(x) + I_S(x) + F_S(x) \leq 3$.

## 3.Database

The assessments carried out in this study are based on information previously gathered for speech analysis in noise and stress analysis and algorithm formulation. Because the task at hand entails mapping audio single value Neutrosophic sets(SVNS) to text SVNS for comparison, a dataset that included audio translation was necessary. LibriSpeech dataset was chosen as a result. The following two folders were utilised for the project demonstration: Dev-clean (337 MB) and Train-clean-100 (6.3 GB).

4.Methodology

4.1 Audio

converting.flac audio files to.wav

The dataset could be downloaded in FLAC format. These files had to be converted into.wav format in order to be processed further and have features extracted.

4.2 Features Extraction and Preprocessing

The python feature extraction script was then run on the audio files, extracting 193 features for each audio file. As a result, the npy files X dev.npy (2703 x 193) and X train.npy were created (28539 x 193). Then, sklearn was used to normalise these files.

4.3 Text

Using VADER, analyse the sentiment of translated text.

For each input sentence, the sentiment analysis programme VADER delivers a score for the truth,indeterminacy and falsity. Each audio file's text translation was examined using VADER, and SVNS were produced.

## 5. Speech recognition in to text conversion

*5.1 Algorithm:1*

*Step 1*: Import library
Step 2:Import speech recognition
Step 3:Initialize recognizer class
Step 4; Reading Microphone source
*Step 5*: Convert audio to text
*Step 6*: Adjust for ambient noise.
*Step 7*: Recognize the error
*Step 8*:Type the text.

5.2 Programme for Speech to Text

```
r = sr.Recognizer()
    print("Talk")
```

```
    r.adjust_for_ambient_noise(source, duration=0.2)
    audio_text = r.listen(source)
    print("Time over, thanks")
      print("Text: "+r.recognize_google(audio_text))
        print("Sorry, I did not get that")
```

Once the programe is over, then run the programe. The out put is

Talk

Speak through microphone then it will showing. In this experiment speech word is "very good"

Time over, thanks

The output in the screen is

Text: very good

6. Neutrosophic speech stress analysis

*6.1 Algorithm:2*

*Step 1*: Import SentimentIntensityAnalyzer class

Step 2: Function to print sentiments

Step 3: Score for sentiment speech

Step 4: Which contains Truth, Falsity, Indeterminacy, and compound scores.

Step 5: Decide sentiment as Truth, Falsity and Indeterminacy se.

Step 6: Print the value of the compound score

Step 7: Print overall the stress statement is truth ,falsity or indeterminacy.

6.2 Programme for text to stress analysis by Neutrosophic speech algorithm

```
def sentiment_scores(sentence):

sid_obj = SentimentIntensityAnalyzer()
C = sid_obj.polarity_scores(sentence)
print
("Overall sentiment dictionary is : ", C)
Print
("sentence was rated as ", C['falsity']*100, "% Negative")
Print
("sentence was rated as ", C['indeterminacy']*100, "% Neutral")
Print
("sentence was rated as ", C['Truth']*100, "% Positive")
Print
("Sentence Overall Rated As", end = " ")
```

The following sentence "Very Good.","Not bad","Bad", "happy birth day."
"god bless you.","beautiful."

In this algorithm 2 , include the output of the algorithm 1 statements. Once run the programme.

6.3 The output of the programme

1st statement is Very Good the output of the programme is
{'Falsity': 0.0, 'Indeterminacy': 0.238, 'Truth': 0.762}
0.0 % Falsity, 23.799999999999997 % Indeterminacy, 76.2 % Truth and the speech is not under strees
2nd Statement : Not bad
{'Falsity': 0.0, 'Indeterminacy': 0.26, 'Truth': 0.74}
0.0 % Falsity, 26.0 % Indeterminacy, 74.0 % Truth and the speech is not under stress.
3rd Statement :Bad
{'Falsity': 1.0, 'indeterminacy': 0.0, 'Truth': 0.0}
100.0 % Falsity, 0.0 % Indeterminacy, 0.0 Truth and the speech is under stress.

4th statement :Happy Birthday
{'Falsity': 0.0, 'Indeterminacy': 0.351, 'Truth': 0.649}
0.0 % Falsity, 35.099999999999994 %Indeterminacy, 64.9 % Truth and the speech is not under stress.

5th Statement : god bless you
{'Falsity': 0.0, 'Indeterminacy': 0.169, 'Truth': 0.831}
0.0 % Falsity,16.900000000000002 Indeterminacy, 83.1 Truth and the speech is not under stress.
6th Statement : beautiful
{‘Falsity': 0.0, 'Indeterminacy': 0.0, 'truth': 1.0}
0.0 Falsity, 0.0 % Indeterminacy, 100.0 % Truth the speech is not under stress.

7th Statement :Please help me
{'Falsity': 0.0, 'Indeterminacy': 0.167, 'Truth': 0.833}
0.0 % Falsity,16.7 % Indeterminacy, 83.3 % Truth and the speech is not under stress.

8th Statement :hate
{'Falsity': 1.0, 'Indeterminacy': 0.0, 'Truth': 0.0}
100.0 % Falsity,0.0 % Indeterminacy, 0.0 % Truth and the speech is under stress.

9th Statement :Great
{'Falsity': 0.0, 'Indeterminacy': 0.0, 'Truth': 1.0}
0.0 % Falsity, 0.0 % Indeterminacy,100.0 % Truth and the speech is not under stress.
Fig:1 Stress Analysis using Neutrosophic speech recognition

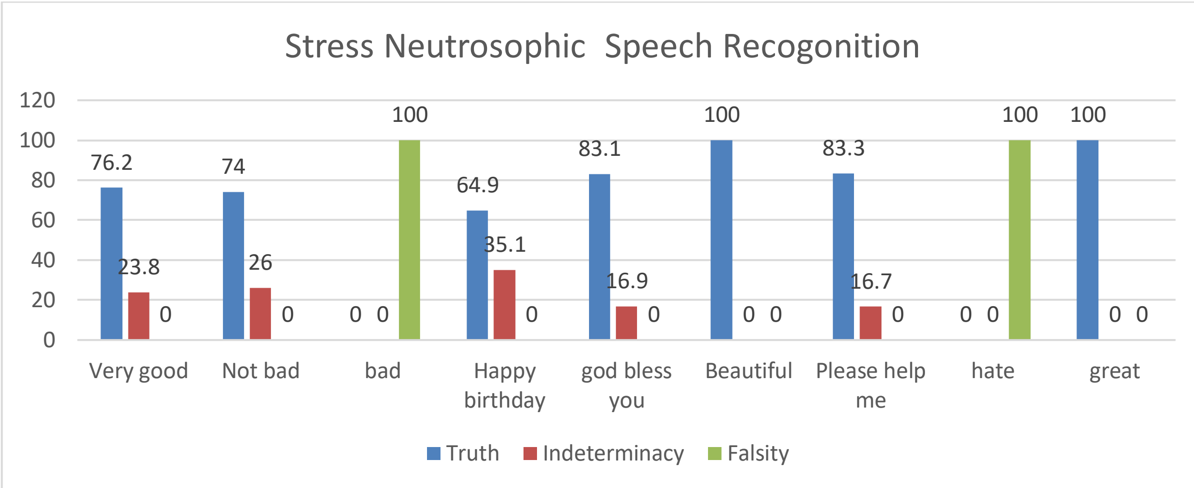


Fig:1 reveal that the percentage of the speech shows the truth, indeterminacy and falsity value .That means the probability of the stress in the speech. The probability value is give the statement is the speech is under stress or not.

Fig:2 Overall rated for Speech

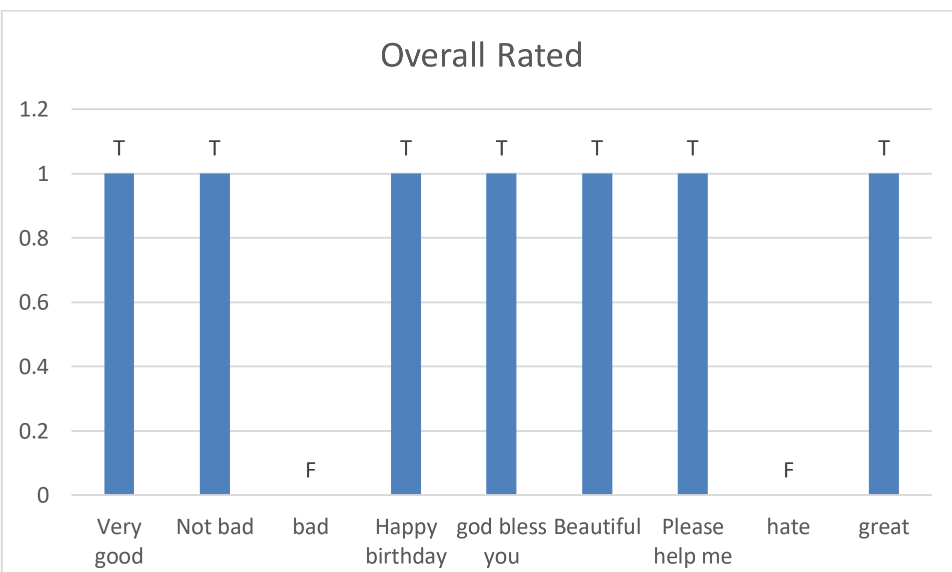

Fig :2 shows that the speech is under stress or not under stress. From the analysis of the speech verygood, notbad, happy birthday, god bless you, beautiful and great is positive speech text. Bad and hate is negative speech .

Conclusion

The requirement to accurately analyse, model, encode, identify, and categorise speech under stress will become increasingly important as speech and language technology develops. The condition of the speaker can be useful information for human-machine and dialogue systems that use voice interaction. This information can be utilised to create speaker and speech recognition technologies, leading to the development of systems that function better in actual multi-tasking environments.The difficulty, though, lies in finding a framework that can effectively analyse and model such speech technologies. The issue of better stress classification utilising targeted speech features has been taken into consideration in this work. categorization of stress The estimation of a probability vector that represents the level of speaker stress is proposed using neurosophic algorithms. Machine learning has demonstrated context-sensitive stress classification. The output stress probability vector can also be used to quantify combinations of speaker stress, such as speech that is both fast and loud. It is claimed that a stress mixture model could be helpful for tasks like sorting emergency phone messages or enhancing the efficiency of traditional speech processing systems. In conclusion, it has been demonstrated that stress classification utilising focused features in Neutrosophic speech recognition algorithm is effective for estimating the level of speaker stress and for providing helpful information for enhancing the performance of a voice recognition algorithm.

**Reference**

[1] J. H. L. Hansen and S. E. Bou-Ghazale, “Robust speech recognition training via duration and spectral-based stress token generation,” ZEEE Trans. Speech Audio Processing, vol. 3, pp. 415421, Sept. 1995.

[2] B. D. Womack and J. H. L. Hansen, “Stress independent robust HMM speech recognition using neural network stress classification,” in P roc. EuroSpeech, pp. 1999-2002.1995.

[3] D. A. Cairns and J. H. L. Hansen, “Nonlinear analysis and detection of speech under stressed conditions,” J. Acous. Soc. Amer., vol. 96, no. 6, pp. 3392-3400, 1994.

[4] J. C. Junqua, “The Lombard reflex and its role on human listeners and automatic speech recognizers,” J. Acous. Soc. Amer., vol. 93, pp, 510-524, Jan. 1993.

[5] B. J. Stanton, L. H. Jamieson, and G. D. Allen, “Robust recognition of loud and Lombard speech in the fighter cockpit environment,” in P roc. ICASSP, May 1989, pp. 675-678.

[6] J. H. L. Hansen and S. E. Bou-Ghazale, “Robust speech recognition training via duration and spectral-based stress token generation,” ZEEE Trans. Speech Audio Processing, vol. 3, pp. 415421, Sept. 1995.

[7] Shaughnessy, D. O. (1986). Speaker recognition. IEEE ASSP Magazine, 3(4), 4–17.

[8] Murty, K. S. R., & Yegnanarayana, B. (2006). Combining evidence from residual phase and MFCC features for speaker recognition. IEEE Signal Processing Letters, 13(1), 52–55.

[9] M. S. Hafiy Hilmy *et al*., "Stress Classification based on Speech Analysis of MFCC Feature via Machine Learning," *2021 8th International Conference on Computer and Communication Engineering (ICCCE)*, 2021, pp. 339-343, doi: 10.1109/ICCCE50029.2021.9467176.

[10] Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN* COMPUT. SCI. 2, 160 (2021). https://doi.org/10.1007/s42979-021-00592-x.

[11] K. Tomba, J. Dumoulin, E. Mugellini, O. A. Khaled and S. Hawila, "Stress detection through speech analysis", *ICETE 2018 - Proc. 15th Int. Jt. Conf. E-bus. Telecommun*, vol. 1, pp. 394-398, 2018.

[12] Nishtha Phutela, Devanjali Relan, Goldie Gabrani, Ponnurangam Kumaraguru, Mesay Samuel, "Stress Classification Using Brain Signals Based on LSTM Network", *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7607592, 13 pages, 2022. https://doi.org/10.1155/2022/7607592.

[13] J. Eichler, R. Schmidt, A. Hiemisch, W. Kiess, and A. Hilbert, “Gestational weight gain, physical activity, sleep problems, substance use, and food intake as proximal risk factors of stress and depressive symptoms during pregnancy,” *BMC Pregnancy and Childbirth*, vol. 19, no. 1, p. 175, 2019.

[14] A. Asif, M. Majid, and S. M. Anwar, “Human stress classification using eeg signals in response to music tracks,” *Computers in Biology and Medicine*, vol. 107, pp. 182–196, 2019.

[15] Hatem S. A. Hamatta, Kakoli Banerjee, Harishchander Anandaram, Mohammad Shabbir Alam, C. Anand Deva Durai, B. Parvathi Devi, Hemant Palivela, R. Rajagopal, Alazar Yeshitla, "Genetic Algorithm-Based Human Mental Stress Detection and Alerting in Internet of

Things", *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4086213, 7 pages, 2022. https://doi.org/10.1155/2022/4086213.

[16] A. Kumar, K. Sharma, and A. Sharma, "Hierarchical deep neural network for mental stress state detection using IoT based biomarkers," *Pattern Recognition Letters*, vol. 145, pp. 81–87, 2021.

[17] N. E. J. Asha and R. Khan, "Low-cost heart rate sensor and mental stress detection using machine learning," in *Proceedings of the 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 1369–1374, IEEE, Tirunelveli, India, June 2021.

[18] N. Padmaja, A. Anusha, M. Dvs, and B. S. Kumar, "IOT based stress detection and health monitoring system," *Helix-The Scientific Explorer| Peer Reviewed Bimonthly International Journal*, vol. 10, no. 2, pp. 161–167, 2020.

[19] D. Raval, "Stress detection using convolutional neural network and internet of things," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 12, pp. 975–978, 2021.

[20] O. Attallah, "An effective mental stress state detection and evaluation system using minimum number of frontal brain electrodes," *Diagnostics*, vol. 10, no. 5, p. 292, 2020.

[21] Li Liu, Yunfeng Ji, Yun Gao, Tao Li, Wei Xu, "A Novel Stress State Assessment Method for College Students Based on EEG", *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4565968, 11 pages, 2022. https://doi.org/10.1155/2022/4565968.

[22] Bo Yang, Zhizhong Ding, "Algorithm for Evaluating the Intervention Effect of Physical Exercise on Stress Groups", *Mobile Information Systems*, vol. 2022, Article ID 1667814, 9 pages, 2022. https://doi.org/10.1155/2022/1667814.

[23] ReemAhmedaFuzhanNasiriaTarekZayed, A novel Neutrosophic-based machine learning approach for maintenance prioritization in healthcare facilities, Journal of BuildingEngineering,Volume 42, October 2021, 102480, https://doi.org/10.1016/j.jobe.2021.102480

[24] Azeddine Elhassouny, Soufiane Idbrahim, F. Smarandache: Machine learning in Neutrosophic Environment: A Survey, Neutrosophic Sets and Systems, vol. 28, 2019, pp. 58-68. DOI: 10.5281/zenodo.3382515.

[25] MayukhSharmaIlanthenralKandasamyW.B.Vasantha, Comparison of neutrosophic approach to various deep learning models for sentiment analysis, Knowledge-Based Systems, Volume 223, 8 July 2021, 107058, https://doi.org/10.1016/j.knosys.2021.107058.

[26] Belal Amin, A. A. Salama, I. M. El-Henawy, Khaled Mahfouz, Mona G. Gafar, "Intelligent Neutrosophic Diagnostic System for Cardiotocography Data", *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6656770, 12 pages, 2021. https://doi.org/10.1155/2021/6656770.

[27] F. Smarandache (2005) A unifying feld in logic. Neutrosophy: neutrosophic probability, set, logic, 4th edn. American Research Press, Rehoboth.

[28] H.Wang , F. Smarandache , Y. Zhang , R. Sunderraman (2010) Single valued neutrosophic sets. MultispMultistruct 4:410–413.

[29]H. Wang , F. Smarandache , Zhang YQ, R. Sunderraman (2005) Interval neutrosophic sets and logic: theory and applications in computing. Hexis, Phoenix.

Received: August 6, 2022. Accepted: January 15, 2023