International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 15, No.1 (May-24)

http://dx.doi.org/10.12785/ijcds/1501116

# Real-time Speech-based Intoxication Detection System: Vowel Biomarker Analysis with Artificial Neural Networks

**Panduranga Vital Terlapu[1,*] and Ram Prasad Reddy Sadi[2]**

[1,*] *Department of Computer Science and Engineering, Aditya Institute of Technology and Management, Tekkali, Srikakulam (Dist), Andhra Pradesh-532 201, India*

[2]*Department of Information Technology, Anil Neerukonda Institute of Technology and Sciences, Sangivalasa, Visakhapatnam, Andhra Pradesh- 531162, India*

*Received 26 Sep. 2023, Revised 22 Mar. 2024, Accepted 29 Mar. 2024, Published 1 May. 2024*

**Abstract:** Alcohol consumption can lead to vocal health risks and long-term health issues for individuals. The paper introduces a novel dataset that analyzes vowel vocalizations to detect early alcohol consumption. This study examines hidden parameters in vowel sounds, such as frequency, jitters, shimmer, and harmonic ratio, which can identify individuals who consume alcohol. It aims to identify subtle vocal patterns that serve as markers for alcohol consumption. This study analyzed 509 vowel vocalizations from 290 records of 46 alcohol-consuming individuals and 219 non-drinkers aged 22–34. The study used intelligent machine learning models and Incremental Hidden Layer Neurons Artificial Neural Networks (IHLN-ANNs) with Back-propagation to identify patterns indicative of alcohol consumption. The Random Forest (RF) model achieved 95.3% accuracy, while the BP-ANNs model showed 99.4% accuracy with five neurons in a hidden layer. The findings could be applied to developing smartphone applications to provide timely alerts and cautionary measures for alcohol consumption, reducing accident risks. The study highlights voice analysis's potential as a non-invasive and cost-effective tool for identifying alcohol consumers, offering potential avenues for future public health initiatives.

**Keywords:** Alcohol Consumers, Voice Parameters, Machine Learning, Neural Networks, ANN

## 1. INTRODUCTION

Alcohol consumption is a widespread and significant public health concern with numerous negative consequences for individuals and society. It can also lead to road collisions and injuries [1]. Accurate identification of alcohol consumption is essential for early intervention and prevention of the adverse effects of excessive alcohol use. Voice analysis has emerged as a promising and non-invasive means of identifying alcohol consumption [2]. According to the World Health Organization (WHO), the average global alcohol consumption in 2019 was 5.8 liters of pure alcohol per capita (age 15 years or older), a slight decrease from 6.1 liters per capita in 2010. Men in the WHO European Region had the highest consumption in 2019 at 15.2 liters per capita, despite a declining trend since 2000. Alcohol has been an integral part of many societies and cultures for centuries, but it is also responsible for three million deaths each year. Of individuals aged between 20 and 39, roughly 13.5% of deaths are attributed to alcohol. There is a clear connection between alcohol use and its adverse effects, such as mental and behavioral disorders, injuries, and non-communicable conditions. Alcohol use also has significant economic and social costs [3][4].

Alcohol consumption is a widely prevalent issue with significant health, social, and economic implications [5]. Alcohol consumption has effects on the voice, with the main ones being dehydration, inflammation, nerve damage, acid reflux, and vocal strain. Alcohol is a diuretic, which means it increases urine production and can lead to dehydration. When the body is dehydrated, the vocal cords become dry, making it harder to produce sound, resulting in a raspy or hoarse voice. Alcohol can cause inflammation in the throat and vocal cords, making it harder to have clear and smooth sounds, resulting in a voice that sounds scratchy, strained, or breathy [6]. Chronic alcohol consumption can cause damage to the nerves that control the muscles involved in speaking and swallowing, leading to a weak or hoarse voice, difficulty speaking loudly or projecting, or difficulty controlling pitch and tone. Alcohol consumption can increase the production of stomach acid

*E-mail address: vital2927@gmail.com, reddysadi@gmail.com*

and lead to acid reflux, which can irritate the vocal cords and cause inflammation or damage. Alcohol can also cause people to speak louder and more forcefully than usual, straining the vocal cords and leading to hoarseness or vocal fatigue. In general, alcohol consumption can affect the voice negatively, especially if consumed excessively or frequently. For people who rely on their voices for their profession, such as singers, actors, or public speakers, it is important to limit alcohol consumption or avoid it altogether to maintain vocal health [7]. Drinking water and avoiding smoking can also help keep the voice healthy. Early identification and intervention can help prevent the negative consequences of excessive alcohol consumption.

This research provides a detailed description of the social and personal impacts of alcohol consumption, as well as an in-depth discussion of alcohol use disorders, particularly those related to voice. It also explains how alcohol affects the acoustic system and how drinkers and non-drinkers can be identified through a vowel voice dataset. Furthermore, it presents a proposed model, Incremental neurons in HL of BP-ANN, that performs better for identifying alcohol consumers. Finally, it provides a detailed analysis of a novel approach and proposed methodology for identifying drinkers, with comparisons to existing machine learning models. This research is a promising novel approach for identifying individuals who consume alcohol based on their voice patterns and could have critical applications in the fields of healthcare and public safety. This approach has several advantages. First, it is non-invasive, requiring no physical tests or samples from the evaluated individual. Second, it is fast and can quickly process large amounts of data, making it ideal for screening large populations. It also has the potential to be highly accurate, as ANNs can learn complex patterns in data and create exact predictions.

**The remaining sections of the paper are presented as follows:**

- Section 2 provides a literature review of relevant background works related to alcohol consumption and voice.
- Section 3 outlines the proposed model and materials, including the construction and description of the dataset, parameters of voices, and description of BP-ANN approaches with GD (Gradient Decent) optimizer. Additionally, performance parameters and confusion matrix for classification analysis are described.
- Section 4 presents the analysis of the results, including the dataset analysis with statistical values, five experimental ML model performances, and simulation results of BP-ANN-based incremental neurons (2 to 5 neurons) HL models.
- Section 5 provides an analysis of comparison, with detailed discussions and comparisons of the proposed model against five experimental ML models and 2 to 5 HLs neurons of BP-ANN, as well as other works related to this research work.
- Finally, Section 6 concludes the paper by reporting the research limitations and suggesting future works on recognizing alcohol consumers utilizing voice datasets.

## 2. LITERATURE REVIEW

This literature survey focuses on background research on the relationship between voice and alcohol consumption. We collect significant research papers and abstracts from various high-quality journals. This survey includes a comprehensive review of relevant literature, including studies on voice analysis for identifying alcohol consumers. Various authors apply different approaches and multiple models to determine alcohol consumption, with some limitations.

Voice analysis has been explored as a non-invasive and cost-effective means of identifying alcohol consumers. Some of the voice parameters are pitch, loudness, timbre, resonance, and range. Pitch is a person's voice's perceived lowness or highness [8]. The resolution is based on vocal cord vibrations' hertz (Hz) frequency. Loudness is the perceived volume of a person's voice. It is defined by the sound wave intensity and is measured in decibels (dB). Timbre is the unique characteristic of a person's voice that differentiates it from others [9]. The harmonics determine a person's voice and can be depicted as bright, warm, nasal, or breathy. Resonance refers to how sound waves vibrate and resonate in the cavities of the head and throat, creating the unique sound of a person's voice. The range is the span of pitches a person can produce comfortably, from their lowest to their highest pitch. Johnson et al. (1990) [10] investigated whether voice recordings can reveal whether a person is intoxicated.

The study found that alcohol consumption affects various aspects of speech, including speech rate, articulation, intonation, and overall speech quality. Specifically, the study found that in the intoxicated state, the individual had a slower speech rate, reduced articulation, decreased pitch variation, and a lower overall speech quality. This analysis found that longer pauses, slower speech rate, and reduced pitch variation characterized the individual in the intoxicated state. The study also found that the individual had a reduced ability to produce specific speech sounds accurately. Wakista et al. (2014) [11] researched the effects of an alcoholic beverage on supra-segmental features of voice, which include stress, intonation, and rhythm. For this analysis, they chose 50 males aged between 21 and 50. The study found that alcohol consumption significantly decreased pitch range, pitch variability, and speech rhythm. There was an increase in the duration of speech pauses, suggesting a step-down in speech fluency. The study also found that the degree of impairment varied depending on the quantity of alcohol and the individual's age and gender.

One limitation of the study is that it only investigated the consequences of alcoholic drinks on supra-segmental features of speech.

Alcohol consumption can have a significant impact on a person's F0, which is the fundamental frequency of their voice. The effects of alcohol on F0 can also vary depending on the amount consumed as well as other factors. Hollien et al. (2001) [12] found that alcohol intoxication affected speech supra-segmentally, with intonation, speech rate, accentuation, and F0 range changes. They chose 35 young (males and females) individuals for this research. They investigated alcohol intoxication effects on speech suprasegmental before and after consuming an amount of alcohol that would result in a blood alcohol concentration (BAC) of 0.10%. They examined speech that extends beyond individual phonemes or segments, such as stress, intonation, and rhythm. As per the result analysis, The participants showed a significant decrease in the F0 range, which indicates a reduction in the ability to modulate pitch; intonation patterns were altered; speech rate increased significantly after alcohol consumption, with longer pauses between sentences; and a decrease in the ability to accentuate syllables. Ma et al. (2021) [13] presented a comprehensive existing literature review on voice features for the smoking estimation objective. The review provides an overview of the various voice characteristics. In this review, they studied their potential to identify smoking status, including fundamental frequency, shimmer, jitter, harmonics-to-noise ratio formants, and others. As per their findings, the HNR value rises when smoking is stopped.

Additionally, jitter and shimmer are significantly reduced. The F0 value increases while abstaining from smoking and decreases once smoking is resumed. Schiel et al. (2012) [14] described the development of the first public corpus of alcoholized German speech and its analysis. The corpus was created by recording speech samples from individuals intoxicated to the point of impairment and then transcribing and annotating the recordings for analysis. The corpus consists of 120 speech samples from 40 participants, who were asked to perform a series of speech tasks while intoxicated to the point of impairment. The authors analyze the corpus to investigate how alcohol affects speech production and explore alcoholized speech's linguistic and non-linguistic features. The analysis focuses on various speech production characteristics, including phonetics, prosody, syntax, and discourse, as well as non-linguistic features such as social and emotional cues. The researchers are investigating various features that could be used to distinguish intoxicated speech from sober speech, such as fundamental frequency (F0) in different contexts, rhythm parameters, and disfluencies. Landman (2018) [15] investigates the impact of alcohol on the vocal range using qualitative analysis. The analysis focused on three main areas: pitch range, volume, and tone quality. The results showed that after consuming alcohol, participants' pitch range decreased, with a significant decrease in the upper range. Additionally, the participants' volume increased, particularly in the mid-range, and tone quality became rougher and less controlled. The vocal range after drinking was significantly higher than their actual performance, suggesting that alcohol may impair a person's ability to assess their vocal range accurately.

Alcohol consumption can have a range of impacts on both animals and humans. However, the specific effects can vary depending on the species and other factors, such as the amount of alcohol consumed. Some of the researchers have experimented with alcohol consumption with animals and humans. Namazi et al. (2021) [16] analyzed how brain activity changes relate to a human's rhythmic pattern. Shannon and sample entropy were used to study voice and EEG signals, and concepts related to complexity and information were used. They exposed ten subjects—five male and five female—to four distinct smells of varying complexity to influence brain activity. The authors evaluated the resulting changes in the subjects' voices and calculated the Shannon and sample entropy of the EEG and voice signals. According to the findings, changes in the complexity and information content of voice and EEG signals are strongly correlated, with r values of 0.8659 and 0.9423, respectively. This study suggests a strong correlation between the alterations in brain activity and the rhythmic pattern of voice, which can be evaluated using complexity and information concepts. Tisljár-Szabó et al. (2014) [17] examined alcohol effects and the ability to produce fluent and accurate speech. For this examination, they used 15 (8 male and seven female) students with a mean age of 20.73+-1.79. As per the results and findings, alcohol consumption had a significant effect on speech production, and the participants significantly produced many speech faults (word substitutions and mispronunciations) when they consumed alcohol compared to when they had the placebo drink.

Moreover, the participants' speech rate was also significantly slower when they consumed alcohol. The study also found that the consequences of alcoholic beverages on delivered production varied depending on the type of speech task. In particular, the participants showed a more significant increase in speech errors and slower speech rates when performing more complex speech tasks (such as a tongue-twister task) than more straightforward speech tasks (such as reading a list of words).

Olson et al. (2014) [18] researched the effects of alcohol on the learned songs of zebra finches, a species of bird known for their ability to learn and produce complex vocalizations. The researchers exposed male zebra finches to either water or an alcohol solution containing a concentration of 0.5 g of ethanol per kg of body weight and recorded their vocalizations. The main finding in this research is that alcohol exposure affects the ability of birds to produce complex songs; alcohol exposure impairs the birds' capability to develop diverse vocalizations, and alcohol exposure affects the ability of birds to learn and integrate new information. As per findings, alcohol may affect the ability of animals, including humans, to learn and produce complex vocalizations, which could affect communication and social interactions. Wang et al. (2019) [19] studied detecting alcohol intoxication through a ResNet-based model for the task of speech. The ResNet network was trained for 50 epochs, and the mean number of times it was introduced was three. They discovered that the UAR (unweighted average recall) was only 0.633 when speaker normalization was not performed on the network.

On the other hand, the UAR increased to 0.677 when they carried out batch z-normalization using the actual speaker label. The performance also improved when they used the speaker label from clustering, and the UAR was 0.671. They extracted the i-vector and used the predicted label from spectral clustering as the speaker label. They found that neither the baseline model nor the proposed model significantly degraded. Kang et al. (2018) [20] studied the association between voice hygiene habits and the Korean (K-VRQOL) Voice-Related Quality of Life among classical singers. For this research, they chose 128 male 35 (27.3%) and female 93 (72.7%)) singers in South Korea who completed a questionnaire on their hygienic voice habits and the K-VRQOL scale, which evaluates the effect of voice problems on an individual's life quality. The data collected was analyzed using correlation analysis, descriptive statistics, and multiple regression analysis.

The study found a substantial negative correlation between K-VRQOL scores and vocal problems, indicating that individuals with more severe vocal problems had lower life quality related to their voices. Multiple regression analysis showed that voice hygiene habits, particularly avoiding alcohol and smoking, were significant predictors of K-VRQOL scores. This study suggests that voice hygiene habits, particularly avoiding alcohol and smoking, are crucial for maintaining good vocal health and quality of life among classical singers. Terband et al. (2018) [21] investigated the relationship between fetal alcohol spectrum disorders (FASD) and speech impairments in boys. FASD is a spectrum of disorders induced by prenatal alcohol influence and can lead to a physical range, problems of cognition, and behavior issues. Twenty-six children (twelve girls and fourteen boys) with typical development (ages 4.1–8.7) and ten boys with FASD participated in the study. The researchers conducted a comprehensive speech assessment, including speech production, speech perception, and phonological awareness measures. The results showed that boys with FASD had lower scores on significant speech production, speech perception, and phonological awareness than the control group. The findings have important implications for clinical practice, highlighting the need for early identification and intervention for speech impairments in children with FASD.

One of the most noticeable effects of alcohol on speech production is slurred speech. Alcohol can also lead to the impairment of cognitive functions such as attention, memory, and concentration, which are essential for fluent and accurate speech production. Schuller et al. (2014) [22] provides a comprehensive review of research on speaker states, particularly sleepiness, and intoxication, and the challenges associated with detecting them in speech. The authors begin by discussing the importance of seeing the speaker's states, such as sleepiness and intoxication, in various settings, including driving and workplace safety. They then provide some research on detecting these states in a speech, highlighting the various acoustic features that effectively see sleepiness and intoxication, like changes in pitch, speech rate, speech intensity, etc. In this review, they note that continued research in this area is essential for improving safety in various settings and for better understanding the effects of intoxication and sleepiness on speech over the long term. Singer et al. (2007) [23] investigated the psychosocial factors contributing to successful voice rehabilitation after laryngectomy surgery. Laryngectomy is a surgical procedure involving removing the larynx (voice box), which can result in the loss of voice. The authors used several measures to assess the participants' psychosocial status, including the Hospital Anxiety and Depression Scale, the Life Orientation Test, the Social Network Index, and the Questionnaire of Social Support. Van et al. (2018) [24] reviewed the literature on voice stress psychoanalysis and described four fundamental components of the framework: (1) physiological processes, (2) cognitive processes, (3) voice features, and (4) performance outcomes. They argued that understanding the interplay between these factors could provide new insights into human performance. Based on the idea of voice stress analysis, the review article proposed a new framework for comprehending the connection between effort and voice in human performance. Liu et al. (2019) [25] investigated how listeners deal with atypical pronunciations of words during speech perception. Specifically, the study examines whether listeners attribute unusual pronunciations to individual talker characteristics or speech errors. The authors analyzed the data and found that the effect of atypical pronunciation on visual target identification was more significant for words that were phonologically similar to their typical pronunciation. This suggests that listeners are more sensitive to deviations from expected pronunciations when the pronunciation deviates slightly from the norm.

Alcohol consumption can have a significant impact on a person's ability to speak languages, mainly if they are not fluent in the language they are speaking. Alcohol is a depressant, which means it can slow down the central nervous system and impair cognitive function. Hendricks et al. (2019) [26] aimed to summarize and synthesize the findings from longitudinal studies on the effects of prenatal alcohol exposure on language, speech, and communication outcomes in children. The review included studies that followed children from birth to adolescence or young adulthood and assessed their language, speech, and communication abilities. Liu et al. (2018) [27] proposed a computational model that captures how listeners integrate acoustic cues and contextual information to infer the causes of speech sounds. For this, they conducted two experiments. In the first experiment, contestants were presented with a sequence of speech sounds and were asked to indicate which of two possible causes they thought produced the sounds. The causes were defined by different combinations of speaker identity, speaking rate, and vowel context. Participants listened to speech sounds in the second experiment with manipulated speaking rates and vowel contexts. The results showed that contestants could use acoustic cues to infer the speech sounds' speaking rate and vowel context. The findings have implications for understanding speech perception and developing computational speech processing models. Cooney et al. (1998) [28] investigated how alcohol affects speech using acoustic analysis. The study involved recording the speech of participants before and after consuming alcohol and then analyzing the recordings to assess the alcohol effects on various acoustic measures of speech. The study found that alcohol consumption resulted in changes in several acoustic measures of speech, including pitch, intensity, and spectral tilt. Specifically, the outcomes revealed that alcohol consumption led to a decrease in pitch and intensity and an increase in spectral tilt. The study also found that these results were more pronounced in female than male participants. The author discusses the possible physiological mechanisms underlying how alcohol affects speech, including alterations in the tension of the vocal cords and the coordination of the muscles involved in speech production. The physical effects on the vocal cords and alcohol can also affect a singer's overall health and well-being, leading to fatigue, decreased lung capacity, and an increased risk of illness [29][30]. These factors can all impact a singer's ability to perform at their best and may lead to long-term damage to the voice.

This survey aims to review the existing literature on voice analysis for identifying alcohol consumers, highlight the limitations of existing approaches, and propose a new approach that addresses these limitations. This literature survey aims to analyze existing methods for identifying alcohol consumers using voice analysis. It offers an intelligent, novel approach that has the potential to improve accuracy and adaptability. The survey includes studies that investigate the relationship between alcohol consumption and various voice parameters, such as pitch, frequency, and ferment patterns, and studies that examine the accuracy and reliability of voice analysis for identifying alcohol consumption.

## 3. MATERIALS AND MODELS

The proposed ML models for identifying alcohol consumers using a vowelized voice dataset can be implemented using a pre-processed vowelized voice dataset. In the preprocessing, we remove noise and normalize the audio levels. The pre-processed data is split into training and testing sets. The training set is used to train the model, while the testing set is used to evaluate the model's performance. For this, we use 10-fold cross-validation. We chose some suitable machine-learning algorithms for this task. The chosen algorithms are k-NN, C4.5, SVM, and Random Forest algorithms, as well as NNs. The selected ML algorithm is trained on the training set using the extracted voice features. The model's performance is evaluated using accuracy, precision, recall, and F1-score metrics. A collection of voice recordings of people pronouncing different vowels. The dataset should include both alcohol consumers and non-alcohol consumers.

### A. *Proposal Model*

Figure 1 shows the proposed model and describes the drinkers' identification system with a voice data set utilizing the back propagation ANN model's incremental hidden layer (HL) neurons. In this, we collected the voice records and personal information from drinkers and non-drinkers. All voice records have been vowelized (a, e, i, o, u). Before storing *.wav files in the storage, it removes the unwanted voice data from the original voice information. For the experiment, extract the vocalizations' hidden values with voice parameters like pitch, pulses, voicing, jitter, shimmer, and harmonica. Club the voice parameters of hidden values with relative persons' data with cleaning and normalization, then create the *.csv data file and store it in the secondary storage. The statistical analysis is conducted using the *.csv file, and the intelligent incremental hidden layer neurons ANN classifier is applied for the classification of drinkers and non-drinkers. After getting the optimal ANN model, examine the alcohol drinker's predictions with unknown voice parameter values. This model helps identify alcohol consumers using mobile applications.

### B. *Dataset Description*

Table 3 depicts all attributes of the data, as well as their data types and range values, which are also described. The whole set of data contains digital values. The vowels are described as a—1, e—2, and so on, u—5. The pitch parameters are 5, which are mean values, median, STD, minimum, and maximum values. The target classes are two, which are 0 and 1. '0' represents non-drinkers, and '1' describes drinkers relatively. The selective information about the dataset is presented in Table 3, which contains information about data attributes and their descriptions.

Figure 2 (A) shows the unwanted voice data or noise marked with red circles. We have removed this noise data without affecting the original voice records. After removing the noise, the voice records are shown in Figure 2(B), with green circles as indicators. Slurred speech is a common symptom of alcohol consumption. Alcohol affects the central nervous system, which can lead to impaired motor function and coordination, including the muscles involved in speech production [31]. Specifically, alcohol can affect the muscles in the face, tongue, and throat, making it difficult to articulate words and form coherent sentences. Alcohol consumption can affect voice pulses, the slight variations in frequency that occur in the human voice during speech [32].

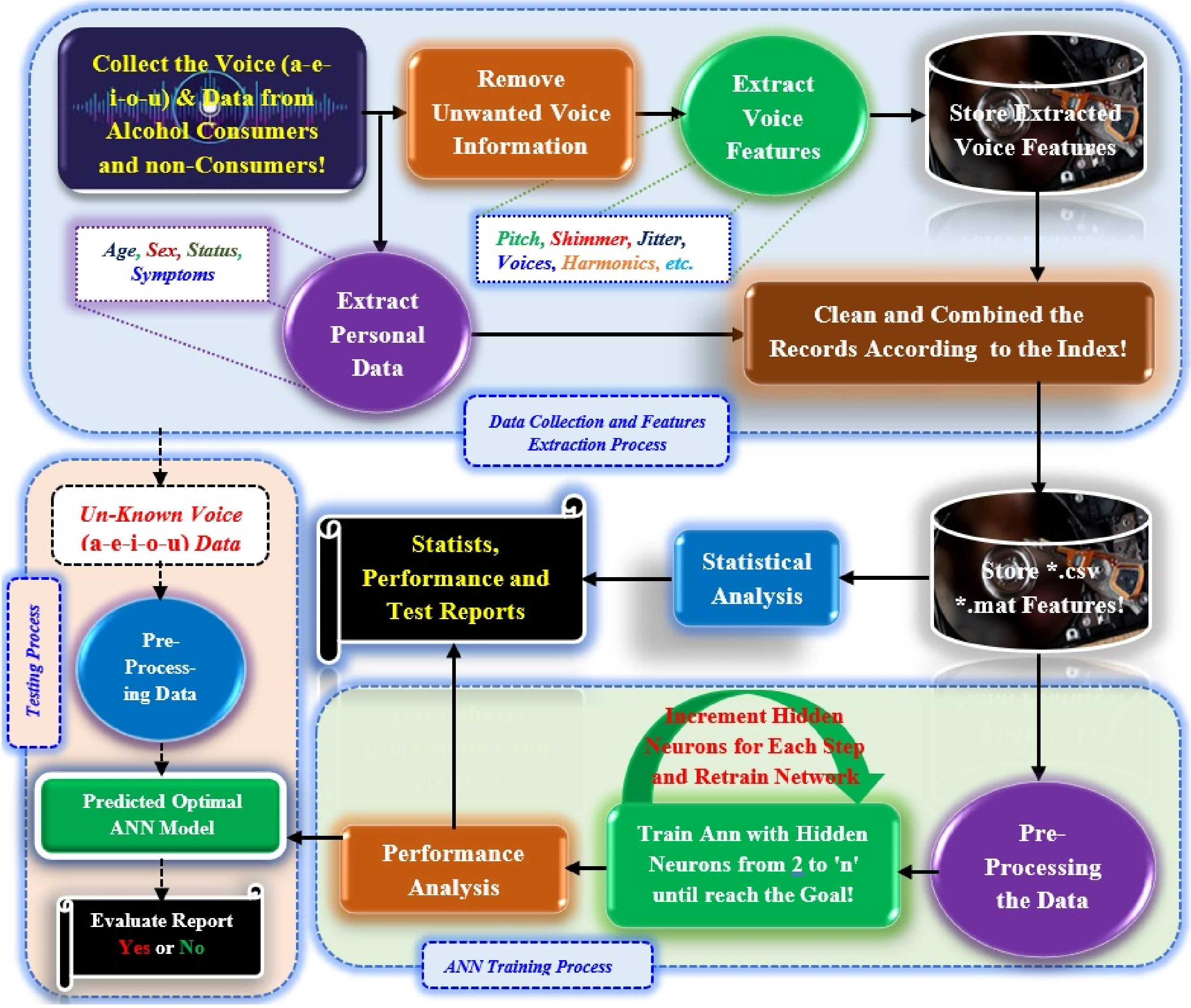
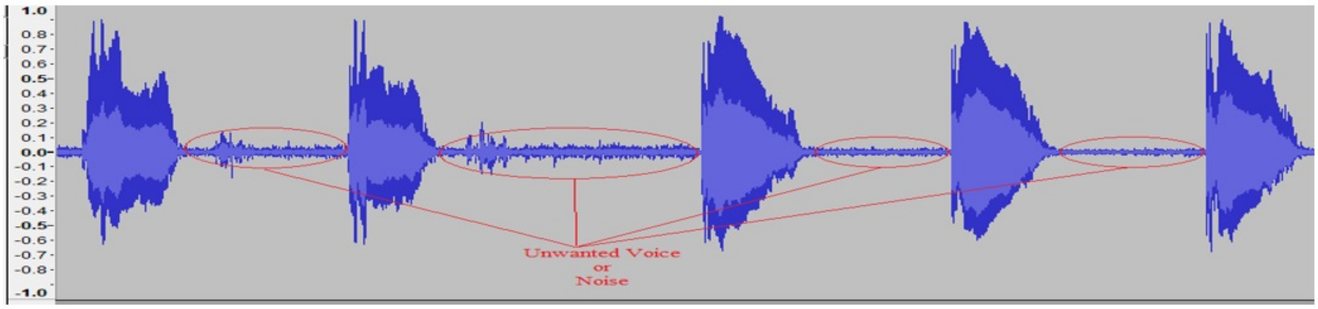



Figure 1. Drinkers prediction proposal model using Incremental Hidden Layer neurons ANN

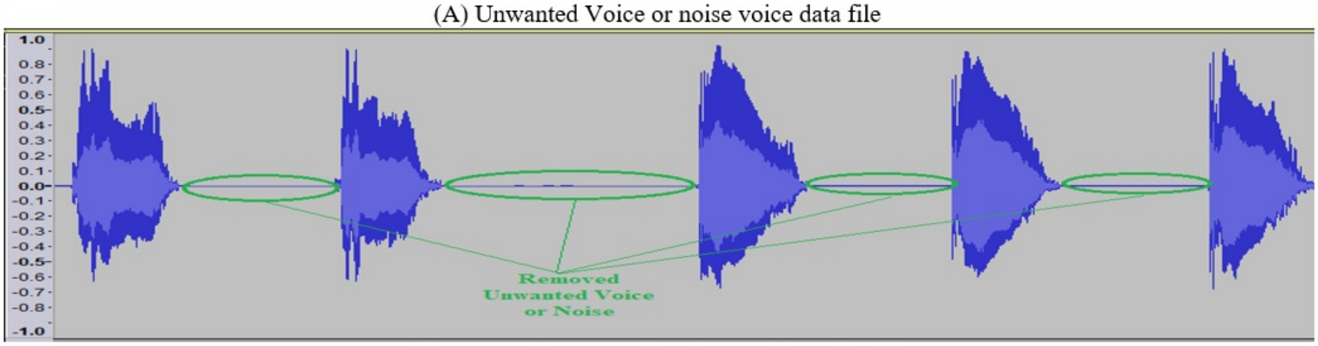
Figure 3(A) shows the pulses associated with the pronunciation of the vowel sound '/a' by alcohol drinkers. The sound'/a' pulses appear to be lagging and narrow with flickers. Figure 3(B) shows the non-drinkers' voice pulses for the vowel sound '/a', revealing a clear difference in voice pulses compared to the drinkers.

Sound is produced by the vocal system, which includes the tongue, pharynx, larynx, and lips. The vocal cords are housed in the larynx, the voice box. The larynx vibrates as air passes through it, resulting in sound waves [33]. The tension and thickness of the vocal cords and the amount of air passing through them determine the sound's pitch and volume. Figure 4 shows the model of producing sound in humans: Muscles force the lungs and make a sound from the nose and mouth. The air passes through the trachea to the vocal folds. The speech signal travels through vocal tract filters produced by human vocal cords. Human speech mainly contains two different classes: vowels and consonants in periodic sources and episodic or noisy sources. The vocal cords' fundamental frequency or vibration is the pitch, defined as F0 [34].

Pitch is the most noticeable acoustic attribute of the voice. According to this property, we can differentiate between genders in humans, as the pitch value of men's voices is lower than that of women's [35]. It affects humans in terms of characteristics, leadership qualities, and more. Pitch is an acoustic property of the voice and affects patterns of features associated with human processes, such as management abilities and dominance. Voice pitch also affects gender differences because a woman's pitch rate is

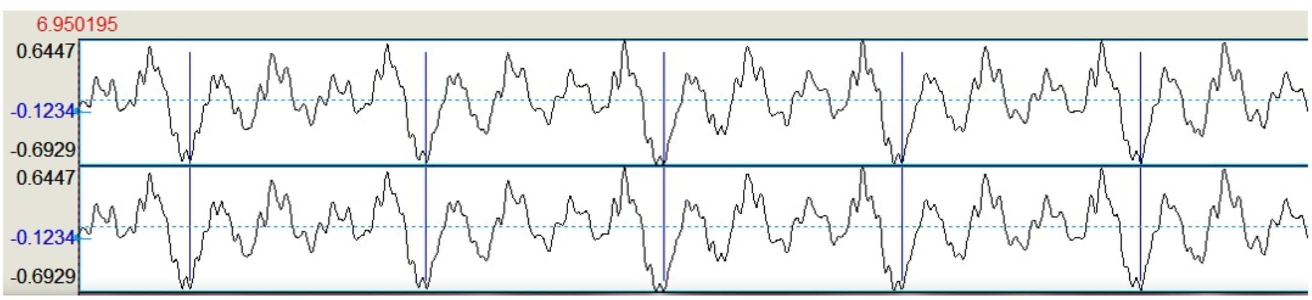



(A) Unwanted Voice or noise voice data file

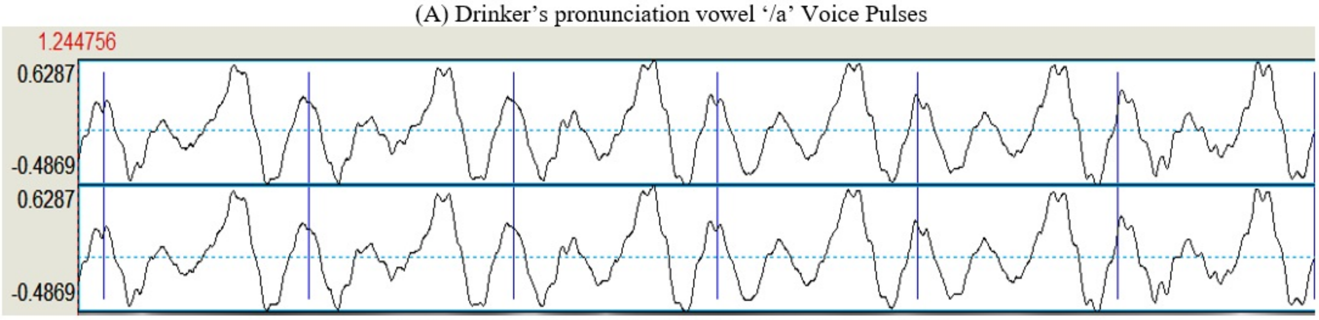



(B) Removed Unwanted Voice or noise voice file

Figure 2. Unwanted Voice Removal Visualizations




(A) Drinker's pronunciation vowel '/a' Voice Pulses




(B)Non-Drinker's pronunciation vowel '/a' Voice Pulses

Figure 3. Difference between Alcohol Consumers and non-consumers voice pulses of pronunciation of '/a'

TABLE I. ALCOHOL DRINKERS AND NON-DRINKERS VOICE DATASET DESCRIPTION

| Sl. No | Attributes | Data Type | Description |
|---|---|---|---|
| 1 | Vowel | Discrete (Integer) | Vowel Sounds 1- a(105), 2-e(105), 3-i(104), 4-o(101), 5-u(104) |
| 2 | Age | Continuous (Integer) | Age of Drinkers and non-Drinkers range is 22 to 34 years |
| 3 | Median pitch | Continuous (Real) | Median pitch in Hz (Hertz) range is 102.252 to 271.529 Hz |
| 4 | Mean pitch | Continuous (Real) | Mean pitch in Hz (Hertz) 101.016 to 289.79 Hz |
| 5 | Std. Div. | Continuous (Real) | Standard deviation Pitch in Hz between 0.886 to 141.719 Hz |
| 6 | Minimum pitch | Continuous (Real) | Minimum pitch between 66.592 to 253.562Hz |
| 7 | Maximum pitch | Continuous (Real) | Maximum pitch between 106.964 to 527.64 Hz |
| 8 | No. of pulses | Continuous (Integer) | Number of pulses between 13 to118 |
| 9 | No. of periods | Continuous (Integer) | Number of periods between 12 to 117 |
| 10 | Mean period | Continuous (Real) | Mean period in seconds range is 0.00344 to 0.00991 seconds |
| 11 | Std. Div. of period | Continuous (Real) | Standard deviation of period, range is 0.000029 to0.00306 seconds |
| 12 | Fraction of UVFL | Continuous (Real) | Fraction of locally unvoiced frames range is 0 to 33.333 |
| 13 | No of Unvoiced | Continuous (Integer) | Number of unvoiced frames range is 0 to 19 |
| 14 | Total frames | Continuous (Integer) | Total number of frames range is 10 to 81 |
| 15 | Number of VBs | Continuous (Integer) | Number of voice breaks between 0 to 2 |
| 16 | Degree of VBs | Continuous (Real) | Degree of voice breaks range is 0.0 to 29.96(seconds/seconds) |
| 17 | Jitter (loc.) | Continuous (Real) | Jitter (local) in % range is 0.26% to 4.93% |
| 18 | Jitter (loc., abs) | Continuous (Real) | Jitter (local, absolute) in seconds range is 0.000014 to 0.000389 |
| 19 | Jitter (rap) | Continuous (Real) | Jitter (rap) in % range is 0.08% to 2.98% |
| 20 | Jitter (ppq5) | Continuous (Real) | Jitter (ppq5) in % range is 0.1 to 3.51 |
| 21 | Jitter (ddp) | Continuous (Real) | Jitter (ddp) in % range is 0.24 to 8.92 |
| 22 | Shimmer (loc.) | Continuous (Real) | Shimmer (local) in % range is 1.81 to 23.35 |
| 23 | Shimmer (loc.,abs) | Continuous (Real) | Shimmer (local, dB) in decibel range is 0.157 dB to 1.893dB |
| 24 | Shimmer (apq3) | Continuous (Real) | Shimmer (apq3) in % range is 0.85% to 11.85% |
| 25 | Shimmer (apq5) | Continuous (Real) | Shimmer (apq5) in % range is 1% to 15.16% |
| 26 | Shimmer (apq11) | Continuous (Real) | Shimmer (apq11) in % range is 0.77% to 28.33% |
| 27 | Shimmer (dda) | Continuous (Real) | Shimmer (dda) in % range is 2.55% to 35.55% |
| 28 | Mean AC | Continuous (Real) | Mean autocorrelation: 0.658867 to 0.994578 |
| 29 | Mean NHR | Continuous (Real) | Mean noise-to-harmonics ratio: 0.005474 to 0.61314 |
| 30 | Mean HNR | Continuous (Real) | Mean harmonics-to-noise ratio: 3.197 to 26.516 dB |
| 31 | Target Class (0 or 1) | Discrete (Integer) | 0-Non-Drinker (219) 1-Drinker (290) |

higher than a man's. Figure 5 shows the demographic results of the spectrum of voice signals and its voice parameters.

amsmath

$$jitter = \frac{1}{N_p - 1} \sum_{l=1}^{N_p} |T_l - T_{l-1}| \quad (1)$$

$$Jitter_{Relative} = \frac{\frac{1}{N_p-1} \sum_{l=1}^{N_p} |T_l - T_{l-1}|}{\frac{1}{N_p} \sum_{l=1}^{N_p} T_l} \times 100\% \quad (2)$$

$$Jitter_{Relative} = \frac{\frac{1}{N_p-1} \sum_{l=1}^{N_p} |T_l - T_{l-1}|}{\frac{1}{N_p} \sum_{l=1}^{N_p} T_l} \times 100\% \quad (3)$$

$$Jitter_{ppq5} = \frac{\frac{1}{N_p-1} \sum_{l=2}^{N_p-2} |T_l - (\frac{1}{5} \sum_{m=l-2}^{l+2} T_m)|}{\frac{1}{N_p} \sum_{l=1}^{N_p} T_l} \times 100\% \quad (4)$$

$$Shimmer_{dB} = \frac{1}{N_p - 1} \sum_{l=1}^{N_p-1} |20 \log_{10} \left( \frac{A_{l-1}}{A_l} \right)| \quad (5)$$

$$Shimmer_{Relative} = \frac{\frac{1}{N_p-1} \sum_{l=1}^{N_p-1} |A_l - A_{l+1}|}{\frac{1}{N_p} \sum_{l=1}^{N_p} A_l} \times 100\% \quad (6)$$

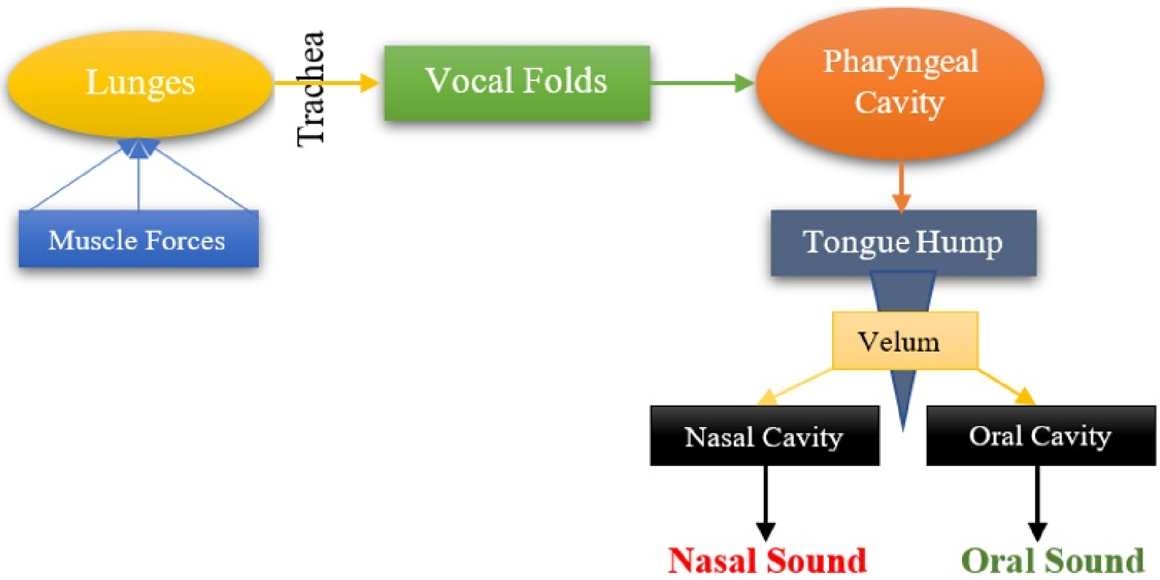



Figure 4. Model for Sound Produced by the Human

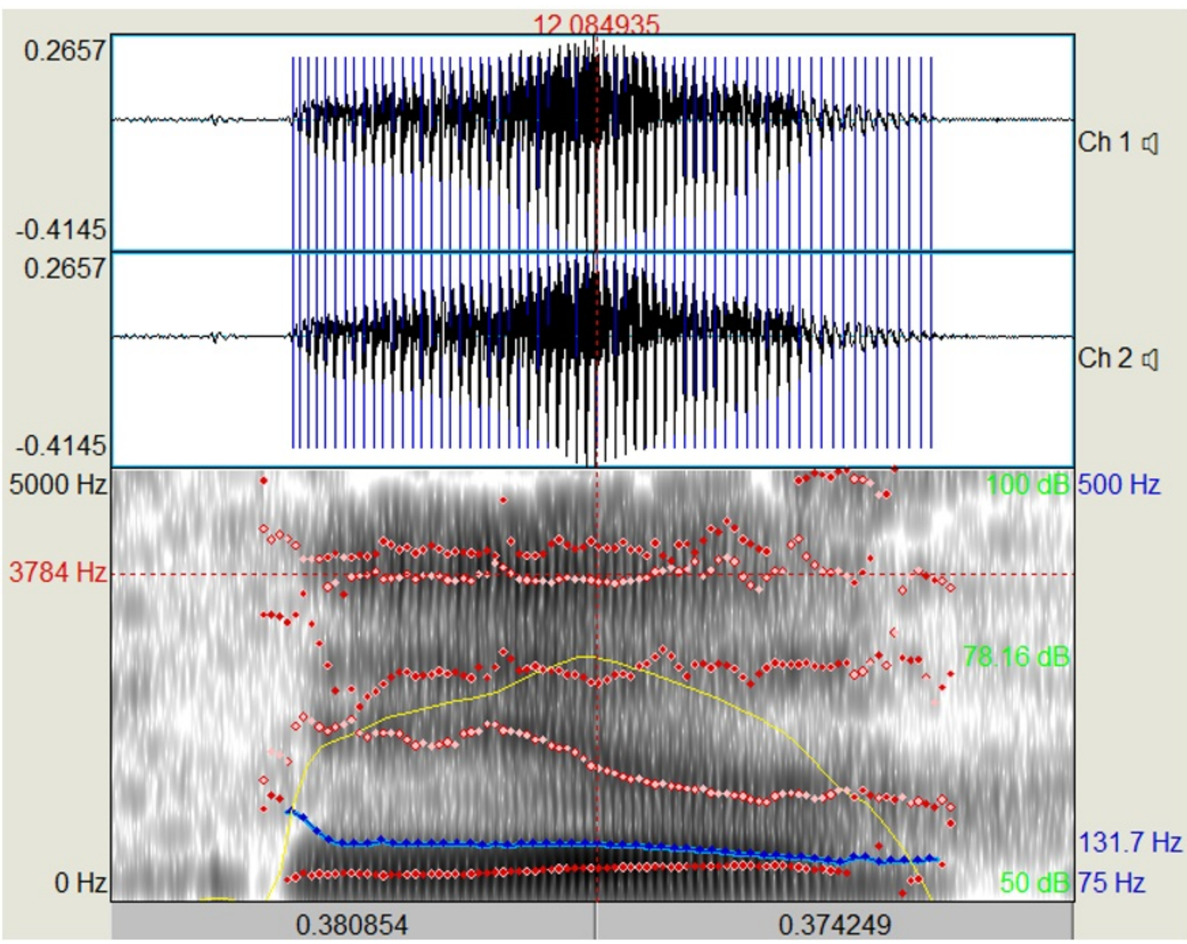



Figure 5. Spectrum of Voice '/a' Sound Audio Signal

$$"Shimmer"_{rapq3} = \frac{\frac{1}{N_p-1}\sum_{l=1}^{N_p-1}|A_l - (\frac{1}{3}\sum_{m=l-1}^{l+1}A_m)|}{\frac{1}{N_p}\sum_{l=1}^{N_p}A_l} \times 100\% \quad (7)$$

$$Shimmer_{ppq5} = \frac{\frac{1}{N_p-1}\sum_{l=2}^{N_p-2}|A_l - (\frac{1}{5}\sum_{m=l-2}^{l+2}A_m)|}{\frac{1}{N_p}\sum_{l=1}^{N_p}A_l} \times 100\% \quad (8)$$

$$Harmonic\_NoiseRatio = 10\log_{10}\left(\frac{V_{AC}(T)}{V_{AC}(0) - V_{AC}(T)}\right) \quad (9)$$

*C. ANN Model*

Artificial Neural Networks (ANNs) are ML model types that act as the human brain to solve complex problems. ANNs are composed of interconnected nodes, or "neurons," which work together to process and interpret input data [36]. These neurons are organized into layers, each performing a specific task in the overall process. At the most basic level, an ANN model consists of three key components: an input layer, one or more hidden layers, and an output layer. The input layer receives data from the outside world, and the output layer produces the final result of the network's computation. The hidden layer(s) in between is where the actual processing occurs. The training process of an ANN model involves feeding input data into the network and adjusting the weights and biases of each neuron to minimize the error between the network's output and the actual output [37] [38]. This work is done through a process called back-propagation, where the error signal is propagated backward through the network to adjust the weights and biases of each neuron.

The artificial neural network (ANN) model for alcohol consumer detection analysis with a back-propagation algorithm is described in Figure 6. The input layer takes the input features or attributes X1, X2, ..., X30 for identification with an objective or target class of either yes or no (drinker or not). The ANN is made up of three layers: the input (IL) layer, the hidden (HL) layer, and the output (OL) layer. The training of the NN is demonstrated with input- and output-based matches utilizing feature attribute values. Specifically, NNs perform this function by working with an input transformation set. According to our analysis, the number of HL neurons is optimized by two to five, and this can continue until the maximum accuracy is reached. In this process, the values of feature attributes are transformed through the HL, and then the predicted result appears at the OL. All these changes are based on the bias (B) and weight (W) values. During the training, the NN learns and adjusts the weight values to minimize the loss (L) between target and actual output values. These weights are adjusted using the gradient descent (GD) optimization process at each epoch. Compute the activation in the forward direction and allocate weights to the hidden layer neurons. The computation measures the loss function using the outputs and the original target values. Back-propagate to the output layer and update the weights of the relative neurons. Figure 7 shows the results of the forward-direction computations.

$$\mathbf{W}^{(n+1)} = \mathbf{W}^{(n)} + \varepsilon\frac{\partial L}{\partial \mathbf{W}} \quad (10)$$

W represents the value of the weight, n denotes the nth value of the weight, the learning rate is denoted by , and L is the loss value; then, L/W is the gradient weight value to loss (Equation (10)). This value is changed in the gradient descent cycle. The overall configuration of one feature neuron is determined using Equation (11).

$$\mathbf{A} = \mathbf{XW} + \mathbf{B} \quad (11)$$

The network uses weight update rules and activation functions to adjust internal connections based on training data, learning complex relationships between input features and output targets (Eq. (11) and Eq. (12)).

$$\mathbf{A_i} = \sum_{j=1}^{n} X_{ij}W_j + B_i \quad (12)$$

The output, likely from a linear combination of weights and inputs, is fed to the activation function  (sigma), applied independently to each element (i = 1, 2..., m) (Equation (13)).

$$\sigma(x) = \frac{1}{1+\exp(-x)} \quad (13)$$

Compose the sequence OL neurons output Y or Yi that is represented in eq. (14)

$$Y = \sigma(\mathbf{A}) = \sigma(\mathbf{XW} + \mathbf{B})$$

OR

$$\mathbf{Y_i} = \sigma(A_i) = \sigma(\sum_{j=1}^{n} X_{ij}W_j + B_i) \quad (14)$$

As per Eq. (14), we will compose the general equations that are Eq. (15) and (16) that represent H and Y values.

$$\mathbf{H} = \sigma(\mathbf{XW}^1 + \mathbf{B}_0) \quad (15)$$

$$\mathbf{Y} = \sigma(\mathbf{XW}^2 + \mathbf{B}_1) \quad (16)$$

The equation calculates the activation of a neuron in the first layer of a neural network by multiplying the input vector (X) by the weight matrix (W1), adding the bias

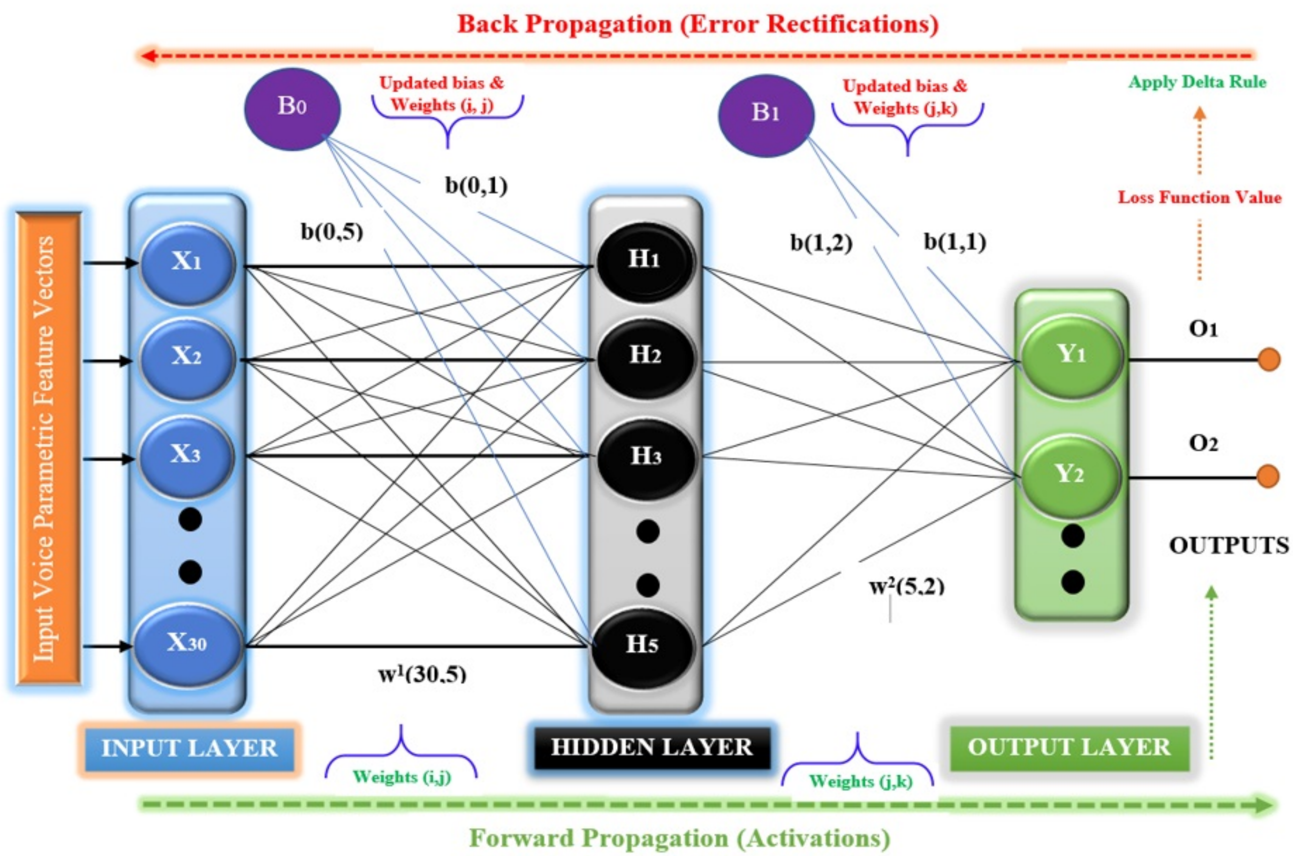



Figure 6. ANN Back Propagation model for Alcohol Consumer Detection

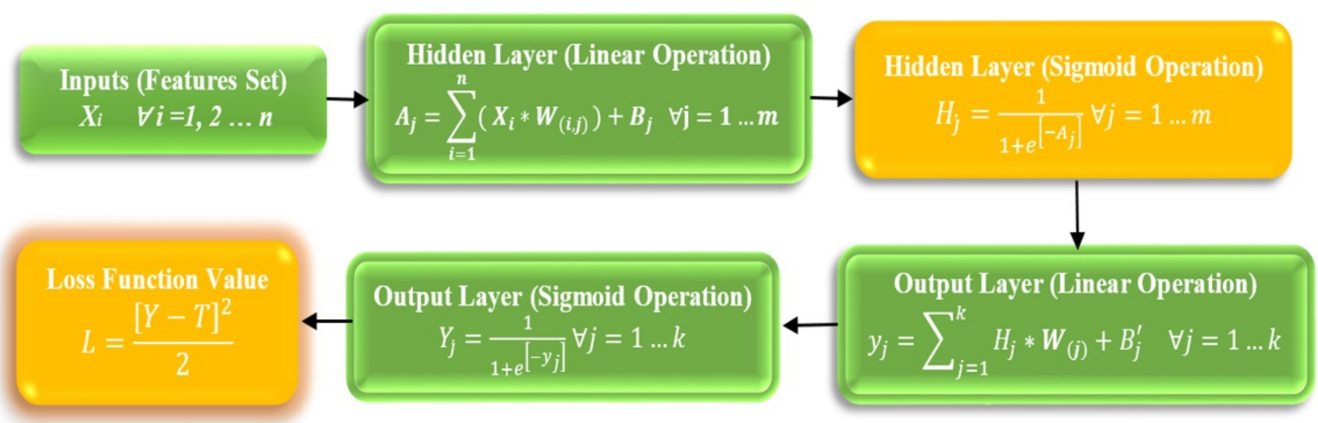



Figure 7. ANN Back Propagation Mathematical model

vector (B0), and applying the activation function (sigma) element-wise, introducing non-linearity and allowing the network to learn complex relationships between input and output. Equation(17) depicts the error or loss L value in mean squared actual out value Y and target value T.

$$\mathbf{L} = \frac{1}{2}\|\mathbf{Y} - \mathbf{T}\|^2 \tag{17}$$

$$\frac{\partial L}{\partial \mathbf{W}^2} = \frac{1}{2}\frac{\partial\, \|\mathbf{Y} - \mathbf{T}\|^2}{\partial \mathbf{W}^2} = \frac{1}{2}\left[\frac{\partial \mathbf{Y}^2}{\partial \mathbf{W}^2} - 2\mathbf{T} \odot \frac{\partial \mathbf{Y}}{\partial \mathbf{W}^2}\right] \tag{18}$$

Equation (18) calculates the gradient of the weight matrix (W²) to minimize the squared error between the network's output and desired targets, considering both weight changes and differences. As per Eq. (17) in Eq. (18) then get the solution in Eq. (19)

$$\frac{\partial L}{\partial \mathbf{W}^2} = \mathbf{H}\left[\frac{\exp(-\mathbf{H}\mathbf{W}^2 - \mathbf{B}_1)}{((1 + \exp(-\mathbf{H}\mathbf{W}^2 - \mathbf{B}_1))^2}\right]\left[\frac{1}{(1 + \exp(-\mathbf{H}\mathbf{W}^2 - \mathbf{B}_1))} - \mathbf{T}\right] \tag{19}$$

*D. Gradient Decent - BP of ANN Analysis*

Figure 8 shows the general BP-ANN architecture for calculating loss functions. In this process, an input linear matrix X is given to the network, and the activation function A1 is applied with weights and bias values. The output of A1 is the input of the hidden layer neurons. Activation A2 is applied with bias B1 and weights W2, and then the output Y is obtained [39][40]. The loss of L value is calculated using the measures of Y and T with Eq. (17).

Figure 9 shows a detailed description of the weight-updating process using the back-propagation algorithm. In this process, the given model produces an output (O/P), which is compared to the target value, and the mean squared error value is calculated. If the error value is small or close to the goal, the process is stopped and an optimal NN model is produced. Otherwise, the weights of the neurons are updated, and the NN model is processed again. This procedure is repeated until the goal is reached [41].

Equation (20) derives the loss L with chain derivations with weights W2.

$$\frac{\partial L}{\partial W^2} = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial W^2}\right) = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial A_2}\right)\left(\frac{\partial A_2}{\partial W^2}\right) \tag{20}$$

Equations (21) and (22) define the loss L value at HL along chain derivation.

$$\frac{\partial L}{\partial W^2} = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial W^1}\right) = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial A_2}\right)\left(\frac{\partial A_2}{\partial W^1}\right) \tag{21}$$

$$\doteq \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial A_2}\right)\left(\frac{\partial A_2}{\partial H}\right)\left(\frac{\partial H}{\partial W^1}\right) = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial A_2}\right)\left(\frac{\partial A_2}{\partial H}\right)\left(\frac{\partial H}{\partial A_1}\right)\left(\frac{\partial A_1}{\partial W^1}\right) \tag{22}$$

Equations (23) and (24) derivation in back-propagation error rectified values concerning B1 and B0.

$$\frac{\partial L}{\partial B_1} = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial B_1}\right) = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial A_2}\right)\left(\frac{\partial A_2}{\partial B_1}\right) \tag{23}$$

$$\frac{\partial L}{\partial B_0} = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial B_0}\right) = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial A_2}\right)\left(\frac{\partial A_2}{\partial B_0}\right) \tag{24}$$

$$\doteq \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial A_2}\right)\left(\frac{\partial A_2}{\partial H}\right)\left(\frac{\partial H}{\partial B_0}\right) = \left(\frac{\partial L}{\partial Y}\right)\left(\frac{\partial Y}{\partial A_2}\right)\left(\frac{\partial A_2}{\partial H}\right)\left(\frac{\partial H}{\partial A_1}\right)\left(\frac{\partial A_1}{\partial B_0}\right) \tag{25}$$

$$L(\mathbf{Y}, \mathbf{T}) = \frac{1}{n}\left(\sum_{i=1}^{n}\left(-T^{(i)}\log(Y^{(i)}) - (1 - T^{(i)})\log(1 - Y^{(i)})\right)\right) \tag{26}$$

*E. Confusion Matrix*

In this, we present the importance of the confusion matrix (CM) of the machine learning (ML) model related to the analysis of alcohol consumers. The CM is a measurement of the performance of an ML model with two or more target classes' classification issues [69]. The measurements are actual and predicted values in four blocks: true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). Using the confusion matrix (shown in Figure 10), we can compute some significant parameters, such as accuracy, the F1-value, specificity, precision, recall, etc. We can also measure the area under the receiver operating characteristic (AUC-ROC) curves. Table 4 shows the confusion matrix.

We have computed the parameters for performance accuracy values such as True Positive Rate (TPR), False Negative Rate (FNR), Positive Predictive Value (PPV), Sensitivity, Recall, Miss Rate, Specificity (SPC), True Negative Rate (TNR), Precision, False Omission Rate (FOR), Negative Predictive Value (NPV), Likelihood Ratio (LR), Accuracy (ACC), False Discovery Rate (FDR), Diagnostic Odds Ratio (DOR), and F1-Score. These performance parameters are specified in the equations below (Eq(s). 26–33).

$$Accuracy(ACC) = \frac{\sum TruePositive + \sum TrueNegative}{\sum TotalPopulation} \tag{27}$$

$$TPR = \frac{\sum TruePositive}{\sum ConditionPositive} \tag{28}$$

$$FNR = \frac{\sum FalseNegative}{\sum ConditionPositive} \tag{29}$$

$$FPR = \frac{\sum FalsePositive}{\sum ConditionNegative} \tag{30}$$

$$F_1 score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \tag{31}$$

$$SPCorTNR = \frac{\sum TrueNegative}{\sum ConditionNegative} \tag{32}$$

$$Prevalence = \frac{\sum ConditionPositive}{\sum TotalPopulation} \tag{33}$$

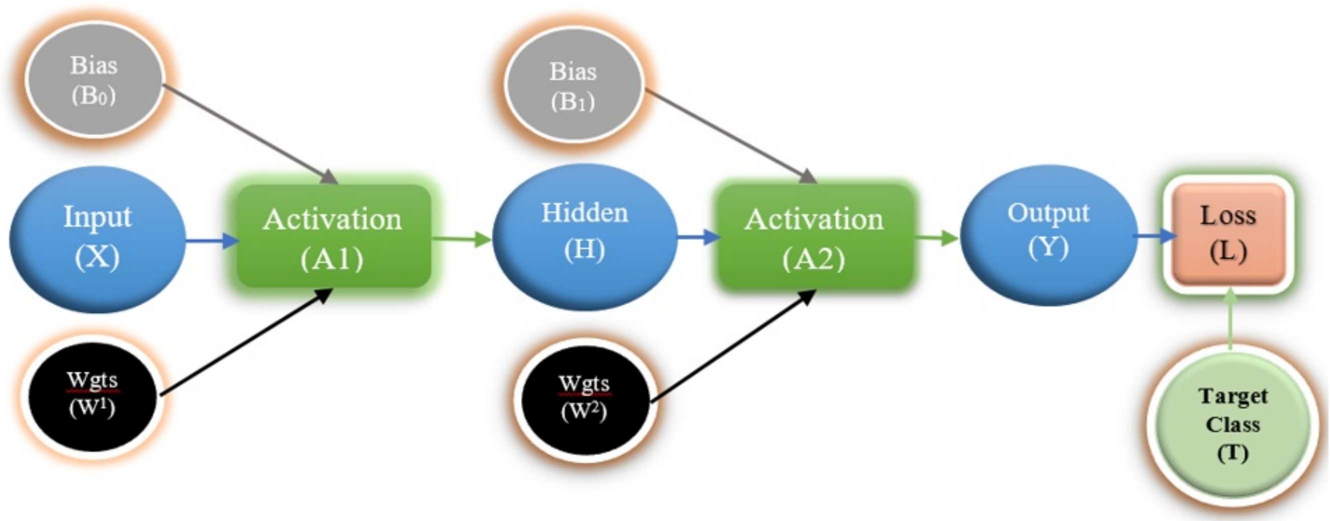



Figure 8. General BP-ANN Architecture for Loss Function Computations

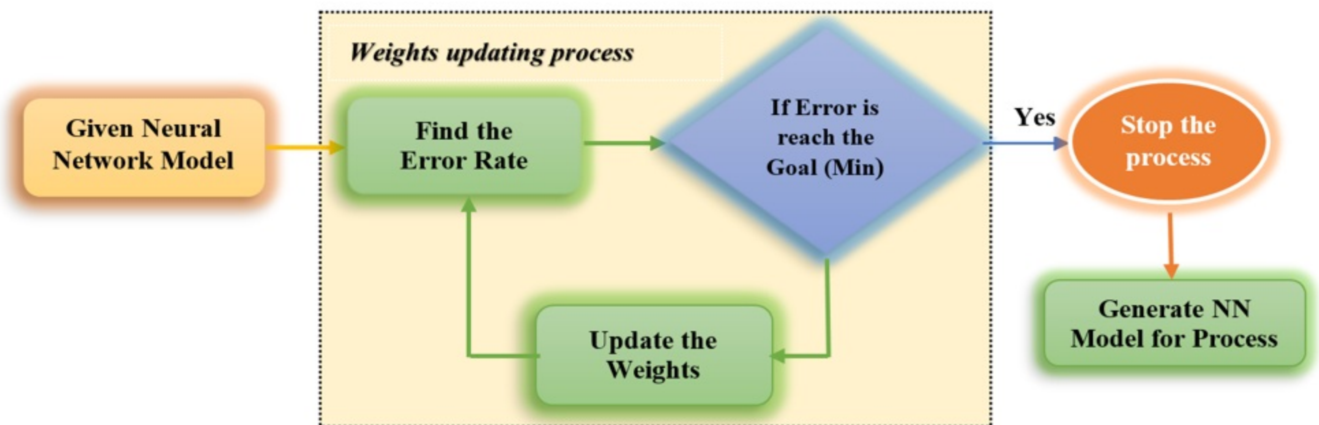



Figure 9. ANN Weights Updating BP-Process

| | | Predicted values | |
|---|---|---|---|
| Actual Values | Classes | Drinker (1) | Non-Drinker (2) |
| | Drinker (1) | (1,1) TP | (1,2) FP |
| | Non-Drinker (2) | (2,1) FN | (2,2) TN |

Figure 10. Alcohol Consumers Confusion Matrix

$$PPVorPRC = \frac{\sum TruePositive}{\sum PredictedConditionPositive} \quad (34)$$

## 4. MATERIALS AND MODELS

In this section, we conduct experiments on drinkers' datasets using statistical and classification models. Firstly, we analyze and describe the dataset in detail using statistical analysis, such as the mean, median, minimum, and maximum values of each attribute for class 2 (non-drinkers), class 1 (drinkers), and the total dataset. After that, we analyze and discuss classification algorithms, including an ANN-based predictive model.

### *A. Authors and Affiliations*

In this section, we conduct experiments on drinker data sets using statistical and classification models. Firstly, we analyze and describe the data set in detail using statistical analysis, such as the mean, median, minimum, and maximum values of each attribute for Class 2 (non-drinkers), Class 1 (drinkers), and the total. We gather personal and voice information from all individuals (drinkers and non-drinkers) with 509 vowels (/'a'/' e'/' i'/' o'/' u') voice records in *.wav format. After that, we analyze and discuss classification algorithms, including an ANN-based proposal model. We divide the data set into two class sets, which are Class 1 set and Class 2. We perform the statistical analysis on each class set, and after combining both classes, we again conduct the statistical analysis and get the total dataset's statistical results. The collected information contains personal data for males aged 22 to 34 and their hidden voice record values. Class 1 describes the drinkers, and Class 2 specifies the non-drinker specifications. As per statistical results, all mean and median values of pitch (mean, median, SD, min, and max) in the Class 1 (drinkers) set are higher than those in Class 2 (non-drinkers), as is the case for the entire set. The number of unvoiced mean values in Class 1 is higher than in Class 2. All the mean and median values of the jitter and shimmer voice parameters in Class 1 are higher than those in the Class 2 set. Table II specifies all voice parameters' minimum and maximum values concerning Classes 1 and 2 and the total data set. The 22- to 34-year-olds are involved in all Class 2, Class 1, and complete dataset categories. It describes a detailed analysis of every voice parameter's minimum and maximum values, such as mean pitch, median pitch, jitter, shimmer, harmonic ratio, etc.

Table III specifies the minimum and maximum values of all the voice parameters concerning Class 1, Class 2, and the total dataset. The age group of 22 to 34 years is involved in all Class 2, Class 1, and total dataset categories. It describes the detailed analysis of minimum and maximum values of every voice parameter, such as mean pitch, median pitch, jitter, shimmer, harmonic ratio, and so on.

Figure 11 shows the class attribute drinkers and non-drinkers bar graph. The blue bar represents the Drinkers' total voice records of 290 and the orange bar represents

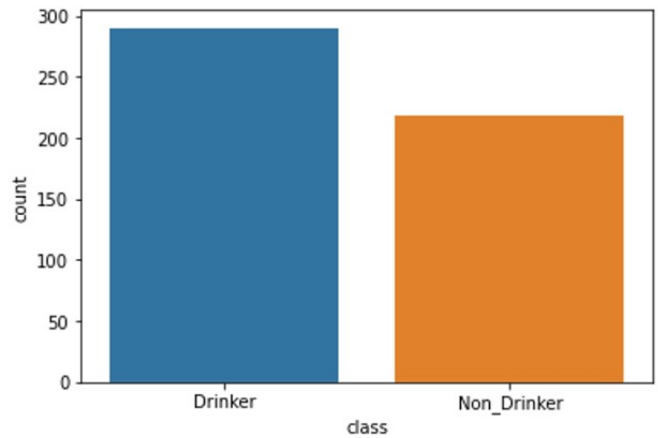



Figure 11. Class attribute Drinker and Non-Drinker Statistics

the Non-drinkers with a total count of 219 records. Figure 12 shows the correlation matrix for the Drinker and Non-Drinker Data attribute relations. The correlation values are between (-1, +1) values. The red color indicates the positive value, and the blue shows the negative value. As per analysis, the mean pitch and median pitch, the attributes pulses and periods, jitter attributes, and shimmer attributes are highly correlated and are nearer to value one. Some other qualities are negatively correlated with different characteristics, such as mean periods, mean auto-correlation, and mean harmonic-to-noise ratio values.

The correlation between the two variables has been measured. Correlation is derived from two things: negative and positive. A positive correlation occurs when two factors or variables change similarly or in the same direction; if one variable increases, the other value also increases relatively. A negative correlation occurs when two factors move in the opposite or inverse direction, meaning that if one increases, the further decreases. The matrix structure is utilized if there are numerous factors, and the objective is to find the correlation between them and store them using the proper data structure. Such a network is known as a correlation matrix. A correlation matrix is a table that shows the correlation coefficients between all factors in the dataset. The correlation matrix finds closely related pairs of factors or feature variables. Using this matrix, analysts can analyze the relationships between multiple variables. The Pearson correlation coefficient can be calculated using this formula. If X and Y are two variables, X and Y are the means, and Xi and Yi are the individual values of X and Y, then the correlation calculation is computed as Equation (35).

$$CorrelationValue = \frac{\sum (X_i - \bar{X})(X_i - \bar{X})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \quad (35)$$

### *B. Machine Leaning Models and Analysis*

In this, we have been analyzing the 5 ML models like k-NN, C4.5, SVM, and Random Forest algorithms using

TABLE II. STATISTICAL MEASURES MEAN AND MEDIAN VALUES OF ALL ATTRIBUTES THROUGH CLASS 2, CLASS 1 AND TOTAL SET

| Attributes | Mean | | | Median | | |
|---|---|---|---|---|---|---|
| | Class-1 | Class-2 | Total set | Class-1 | Class-2 | Total set |
| Age | 29.10513 | 26.41379 | 27.57171 | 29 | 27 | 27 |
| Median pitch | 183.09015 | 138.3734 | 157.613 | 166.409 | 129.749 | 140.479 |
| Mean pitch | 191.26150 | 139.0441 | 161.5109 | 172.955 | 129.087 | 142.677 |
| Std. Div. | 31.55574 | 11.26888 | 19.99741 | 19.173 | 9.703 | 11.351 |
| Minimum pitch | 153.78539 | 120.9325 | 135.0676 | 134.831 | 116.1295 | 123.697 |
| Maximum pitch | 270.79840 | 160.098 | 207.7275 | 273.539 | 145.7995 | 167.905 |
| No. of pulses | 59.757991 | 49.1069 | 53.68959 | 59 | 48 | 52 |
| No. of periods | 58.109589 | 48.06207 | 52.38507 | 57 | 47 | 50 |
| Mean period | 0.00562 | 0.007398 | 0.006633 | 0.00588 | 0.007735 | 0.00702 |
| Std. Div. of period | 0.000757 | 0.000582 | 0.000657 | 0.00057 | 0.000579 | 0.000579 |
| Fraction of UVFL | 5.656735 | 0.948935 | 2.974491 | 3.571 | 0 | 0 |
| No of Unvoiced | 2.045662 | 0.472414 | 1.149312 | 1 | 0 | 0 |
| Total frames | 34.721461 | 36.95517 | 35.99411 | 34 | 36 | 35 |
| Number of VBs | 0.191781 | 0.02069 | 0.094303 | 0 | 0 | 0 |
| Degree of VBs | 1.965881 | 0.208159 | 0.964428 | 0 | 0 | 0 |
| Jitter (loc.) | 1.850183 | 1.181759 | 1.469352 | 1.77 | 1.02 | 1.3 |
| Jitter (loc., abs) | 0.000103 | 0.000089 | 0.000095 | 0.000093 | 0.000074 | 0.000082 |
| Jitter (rap) | 0.853607 | 0.436276 | 0.615835 | 0.76 | 0.315 | 0.49 |
| Jitter (ppq5) | 0.881187 | 0.443793 | 0.631984 | 0.8 | 0.34 | 0.52 |
| Jitter (ddp) | 2.559406 | 1.306862 | 1.845776 | 2.27 | 0.945 | 1.48 |
| Shimmer (loc.) | 9.234795 | 5.993103 | 7.387859 | 8.58 | 5.68 | 6.54 |
| Shimmer (loc.,abs) | 0.903361 | 0.601462 | 0.731356 | 0.87 | 0.5855 | 0.692 |
| Shimmer (apq3) | 4.420228 | 2.471517 | 3.309961 | 4.19 | 2.085 | 2.69 |
| Shimmer (apq5) | 5.483836 | 3.129034 | 4.1422 | 4.81 | 2.665 | 3.34 |
| Shimmer (apq11) | 7.533836 | 4.777724 | 5.963556 | 6.52 | 4.215 | 5.09 |
| Shimmer (dda) | 13.259178 | 7.413207 | 9.928468 | 12.57 | 6.255 | 8.07 |
| Mean AC | 0.897638 | 0.94832 | 0.926514 | 0.900901 | 0.957036 | 0.94016 |
| Mean NHR | 0.147506 | 0.067739 | 0.102059 | 0.138518 | 0.051095 | 0.07467 |
| Mean HNR | 12.955301 | 16.27091 | 14.84435 | 12.342 | 16.279 | 14.969 |

performance parameters like classification accuracy, AUC, precision, recall, and F1 value using confusion matrices. The input dataset is trained using 10-fold cross-validation.

### C. Confusion Matrix Analysis of Each Experimental ML Models

Figure 13 shows the confusion matrices of all experimental ML models. We describe this in detail as follows. The k-NN is constructed with k = 5, and the metric distance is measured with Euclidean uniform weights. Figure 13 (A) shows the confusion matrix of the k-NN: in class 2 (non-DRK), 187 cases are classified correctly, 32 cases are wrongly classified as instances in the DRK class, and 268 instances in class 1 are classified correctly, while 22 instances are classified incorrectly. The C4.5 computational ML model considers numeric and categorical attributes. For each categorical one, the C4.5 computes the information gained and chooses the most esteemed value in the selection process. It then uses the attribute to produce numerous results that have different values for attributes. In this, C4.5 has been configured with a binary tree with a minimum of 2 instances of leaves; not 5 lesser subsets split, and a maximum depth of the tree of 20. Figure 13 (B) shows the confusion matrix of the C4.5: 206 cases of class 2 (N-DRK) are classified correctly, but thirteen sample cases are classified wrongly as instances in the DRK class; 277 sample instances of class 1 are classified correctly, but 13 examples are classified incorrectly. Figure 13 (C) shows the confusion matrix of the SVM, wherein in class 2 (N-DRK) 200 case samples are classified correctly, and 19 case instances are classified incorrectly. 214 sample instances in class 1 are classified correctly, and 14 case instances are classified incorrectly. Figure 13 (D) shows the confusion matrix of the RFs, wherein class 2 (N-DRK), 207 cases are classified correctly, and 12 samples are classified incorrectly, with these instances in the DRK class, and 278 sample instances in class 1 are classified correctly, but 12 case instances are classified incorrectly. Figure 13 (E) shows that the confusion matrix of the NB correctly classified 119 cases of class 2 (N-DRK) and 250 sample cases of class 1. 27 examples of class 2 and 40 sample examples of class 1 were incorrectly classified.

TABLE III. STATISTICAL MEASURES MAXIMUM AND MINIMUM VALUES OF ALL ATTRIBUTES THROUGH CLASS 2, CLASS 1, AND TOTAL SET

| Attributes | Minimum | | | Maximum | | |
|---|---|---|---|---|---|---|
| | Class-1 | Class-2 | Total set | Class-1 | Class-2 | Total set |
| Age | 24 | 22 | 22 | 34 | 34 | 34 |
| Median pitch | 120.32 | 102.252 | 102.252 | 271.529 | 229.056 | 271.529 |
| Mean pitch | 122.994 | 101.016 | 101.016 | 289.79 | 227.555 | 289.79 |
| Std. Div. | 2.115 | 0.886 | 0.886 | 141.719 | 131.084 | 141.719 |
| Minimum pitch | 76.595 | 66.592 | 66.592 | 253.562 | 214.191 | 253.562 |
| Maximum pitch | 126.606 | 106.964 | 106.964 | 527.64 | 492.45 | 527.64 |
| No. of pulses | 25 | 13 | 13 | 118 | 107 | 118 |
| No. of periods | 24 | 12 | 12 | 117 | 106 | 117 |
| Mean period | 0.00344 | 0.00439 | 0.00344 | 0.00811 | 0.00991 | 0.00991 |
| Std. Div. of period | 0.000061 | 0.000029 | 0.000029 | 0.00292 | 0.00306 | 0.00306 |
| Fraction of UVFL | 0 | 0 | 0 | 33.333 | 30.159 | 33.333 |
| No of Unvoiced | 0 | 0 | 0 | 19 | 19 | 19 |
| Total frames | 20 | 10 | 10 | 69 | 81 | 81 |
| Number of VBs | 0 | 0 | 0 | 2 | 1 | 2 |
| Degree of VBs | 0 | 0 | 0 | 29.96 | 25.71 | 29.96 |
| Jitter (loc.) | 0.43 | 0.26 | 0.26 | 4.93 | 4.58 | 4.93 |
| Jitter (loc., abs) | 0.000022 | 0.000014 | 0.000014 | 0.000384 | 0.000389 | 0.000389 |
| Jitter (rap) | 0.15 | 0.08 | 0.08 | 2.98 | 2.56 | 2.98 |
| Jitter (ppq5) | 0.18 | 0.1 | 0.1 | 3.51 | 1.62 | 3.51 |
| Jitter (ddp) | 0.44 | 0.24 | 0.24 | 8.92 | 7.68 | 8.92 |
| Shimmer (loc.) | 3.19 | 1.81 | 1.81 | 23.35 | 17.63 | 23.35 |
| Shimmer (loc.,abs) | 0.302 | 0.157 | 0.157 | 1.893 | 1.543 | 1.893 |
| Shimmer (apq3) | 1.25 | 0.85 | 0.85 | 11.85 | 9.93 | 11.85 |
| Shimmer (apq5) | 1.79 | 1 | 1 | 15.16 | 11.18 | 15.16 |
| Shimmer (apq11) | 2.01 | 0.77 | 0.77 | 28.33 | 16.86 | 28.33 |
| Shimmer (dda) | 3.75 | 2.55 | 2.55 | 35.55 | 29.79 | 35.55 |
| Mean AC | 0.658867 | 0.736003 | 0.658867 | 0.993707 | 0.994578 | 0.994578 |
| Mean NHR | 0.006643 | 0.005474 | 0.005474 | 0.61314 | 0.420146 | 0.61314 |
| Mean HNR | 3.197 | 4.964 | 3.197 | 26.516 | 24.729 | 26.516 |

### *D. Performance Parameters Analysis of MLs*

We need to measure performance parameters such as accuracy, precision, recall, and F1 score.

**K-NN Model:** Table IV shows the performance parameters of the k-NN that depicts class 2 (NDRKs), class 1(DRKs) parameter values, and average weight values also. The AUC and CA are equal performance values for both classes (N-DRK and DRK) and the values are 0.967936 and 0.89391 respectively. F1 and Recall values are somewhat higher in class1 than in class 2. The precision value of class 2 is 0.894737, superior to the class 1 value.

**C4.5 Model:** Table V shows the performance parameters of the C4.5 that describe the class 2 (NDRKs), class 1(DRKs) parameter values, and average weight values. The AUC and CA are equal performance values for both classes (N-DRK and DRK), and the values are 0.953392 and 0.948919respectively. F1, Precision, and Recall values are somewhat superior in class1 than class 2 that the values are 0.955172, 0.955172, and 0.955172 respectively. Figure 14 shows the visualization of the C4.5 tree. According to the analysis, the root node is the maximum pitch value. It will elaborate the tree according to conditions. If the Max pitch is less than or equal to 247.17, it checks the fraction of locally unvoiced frames; otherwise, it assures the total number of frames. If the full frames are less than or equal to 46, someone other than the algorithm may suspect a drinker; the next possible attribute to check is shimmer (local). If the shimmer (local) is less than or equal to 2.273, then we assume that it is possible to be a drinker (1.90) or a non-drinker (0.5). According to the C4.5 tree analysis, we can identify drinkers and non-drinkers with some crucial attributes such as mean pitch, mean period, mean noise, and so on. The detailed analysis is in Figure 14.

**SVM:** In this, the SVM configured with the cost (C) is 1.00, and the regression loss value is 0.10 where the kernel is RBF (exp(-g—x-y—$\hat{2}$ where g = 0.001). The numerical tolerance is 0.001, and the iteration limit is 100. Table VI shows the performance parameters of the SVM that describe the merits and values of class 2 (NDRKs) and class 1(DRKs) parameters' merits and values. The average weight values are also calculated. The AUC and CA are equal

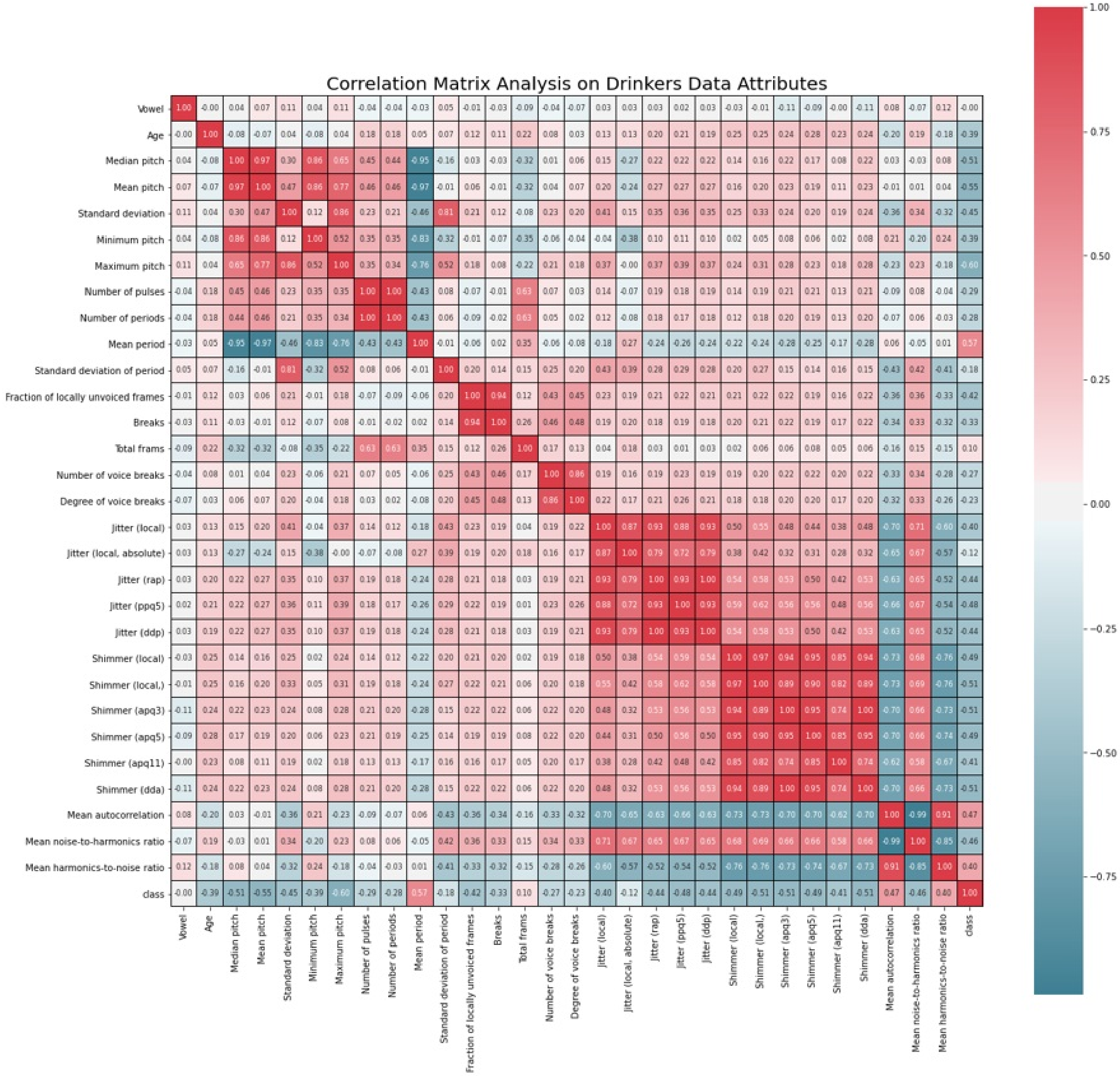



Figure 12. Correlation Matrix for Drinker's Dataset Attributes

performance values for both classes (N-DRK and DRK), and the values are 0.984565 and 0.935167, respectively. F1, Precision, and Recall values are somewhat superior in class1 than class 2, which are 0.94359, 0.935593, and 0.951724, respectively.

**Random Forest (RFs):** In this, the RF is configured that the number of trees is 10, several attributes considered at each spilled 5, replicable training mode discourses slipping subsets not lesser than 5. Table VII shows the performance parameters of the RFs that describe the class 2 (NDRKs) and class 1 (DRKs) parameter values, and average weight values also. The CA values of N-DRK and DRK are equal and the value is 0.952849. The CA, F1, Precision, and Recall values are somewhat superior in class 1 than in class 2 are 0.952849, 0.958621, 0.958621, and 0.958621 respectively.

**Naive Bayes (NB):** In this, the NB is configured using the Bayes probability theorem. Table VIII shows

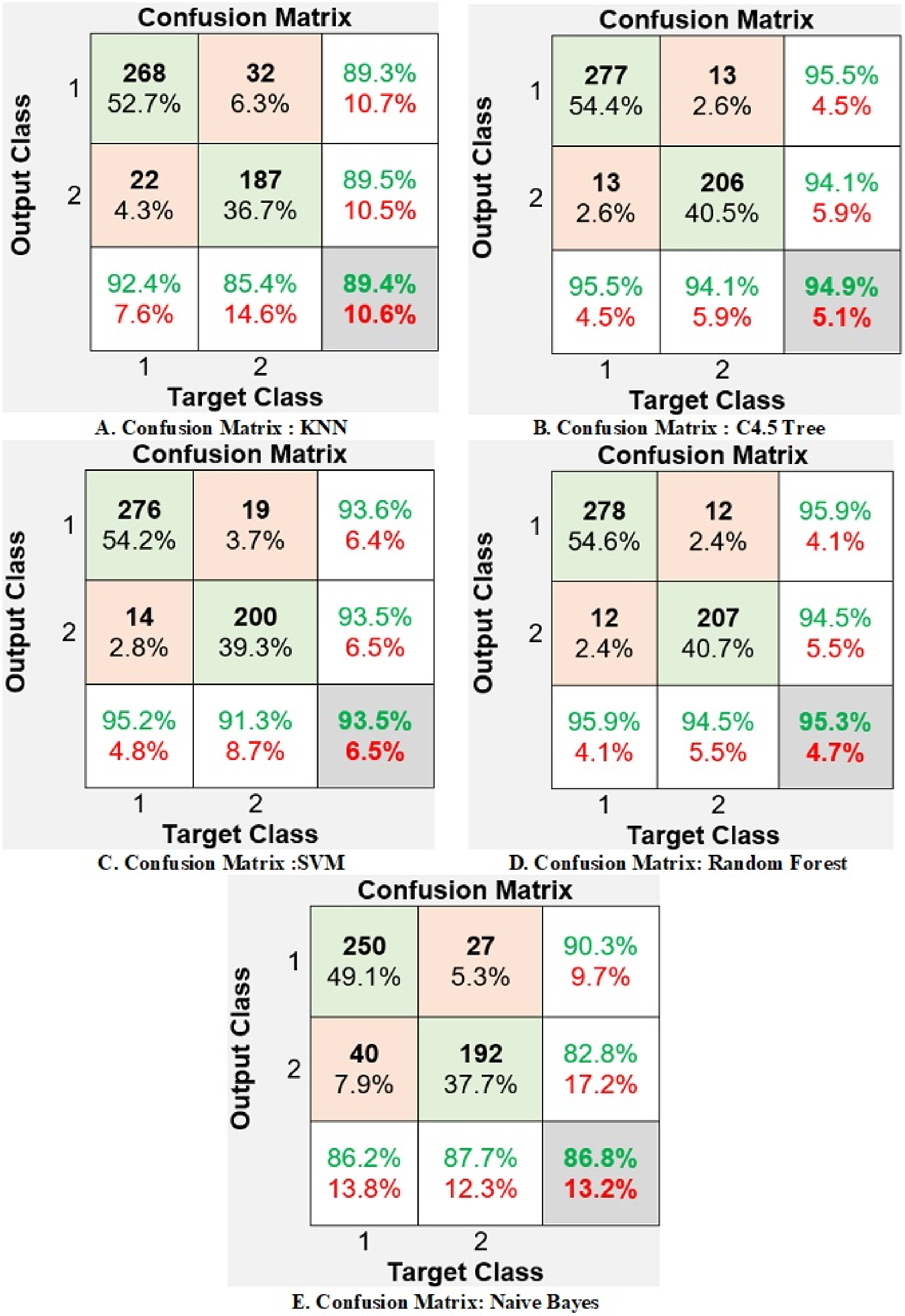



Figure 13. Confusion Matrix for all Experimental ML models

TABLE IV. PERFORMANCE PARAMETER VALUES OF K-NN

| k-NN | AUC | CA | F1 | Precision | Recall |
|---|---|---|---|---|---|
| Class 2 | 0.967936 | 0.89391 | 0.873832 | 0.894737 | 0.853881 |
| Class 1 | 0.967936 | 0.89391 | 0.908475 | 0.893333 | 0.924138 |
| Avg. Wait | 0.968887 | 0.89391 | 0.893569 | 0.893937 | 0.89391 |

TABLE V. PERFORMANCE PARAMETER VALUES OF K-NN C4.5

| C4.5 | AUC | CA | F1 | Precision | Recall |
|---|---|---|---|---|---|
| Class 2 | 0.953392 | 0.948919 | 0.940639 | 0.940639 | 0.940639 |
| Class 1 | 0.953392 | 0.948919 | 0.955172 | 0.955172 | 0.955172 |
| Avg. Wait | 0.958092 | 0.948919 | 0.948919 | 0.948919 | 0.948919 |

TABLE VI. PERFORMANCE PARAMETER VALUES OF SVM

| SVM | AUC | CA | F1 | Precision | Recall |
|---|---|---|---|---|---|
| Class 2 | 0.984565 | 0.935167 | 0.923788 | 0.934579 | 0.913242 |
| Class 1 | 0.984565 | 0.935167 | 0.94359 | 0.935593 | 0.951724 |
| Avg. Wait | 0.984565 | 0.935167 | 0.93507 | 0.935157 | 0.935167 |

TABLE VII. PERFORMANCE PARAMETER VALUES OF RF'S

| RFs | AUC | CA | F1 | Precision | Recall |
|---|---|---|---|---|---|
| Class 2 | 0.988805 | 0.952849 | 0.952849 | 0.952849 | 0.952849 |
| Class 1 | 0.988805 | 0.952849 | 0.958621 | 0.958621 | 0.958621 |
| Avg. Wait | 0.988805 | 0.952849 | 0.952849 | 0.952849 | 0.952849 |

the performance parameters of the NB that describe the class 2(NDRKs) and class 1(DRKs) parameter values, and average waited-for values also. The AUC and CA are equal performance values for both classes (N-DRK and DRK) and the values are 0.934371 and 0.868369 respectively. F1 and Precision values are somewhat superior in class 1 than in class 2 which are 0.881834 and 0.902527 respectively. The Recall value 0.876712 is in class 2 greater than class 1.

Figure 15 shows the analysis of ROC all ML methodologies. We analyze the target class 1 ROC curves. Each model ROC curve specified with each color that the SVM is specified with Aquamarine; NB is specified with orange, the k-NN is specified with purple, the RF model is specified with pink, and C4.5 Tree model specified with the dark green. The RF model performs well with 0.989 value of AUC; secondly, the SVM model presents 0.983 value of AUC. The AUC value of the NB model is 0.934 that performs least in comparison to other experimental models.

### E. Experimental MLs Comparative Study Results

In this analysis, we compare all the experimental machine learning algorithms concerning the average performance (average weight of class 2 and class 1) parameters like F1, CA, AUC, recall, and precision values. As per the study, the RF model shows the best performance to compare other experimental models with a classification accuracy of 0.953 (95.3%), a value of precision is 0.953, and a recall value is 0.953, F1 value is 0.952, measurement of AUC is 0.989. The superior values are depicted in bold and marked with + in Table IX. The second highest-performed model is the C4.5 model specified with a * mark in Table IX. The second highest AUC value is allotted to the SVM model. The detailed performance parameter values are shown in Table IX.

### F. Artificial Neural Network (Incrementing neurons of HL) Models:

In this part, we have compared and discussed incremental hidden layer neurons of BP-ANN models using each performance parameter like CA, AUC, recall, precision, F1, MSE, regression (R) and gradient values, and mue values. As per analysis, the five neurons HL BP-ANN give the best solutions for the identification of alcoholics.

### G. Confusion Matrices BP-ANN (2-5 HL Neurons) Models and Comparative Study on Performance Parameters

Figure 16 shows the analysis of confusion matrices of 2 to 5 neurons of HL BP-ANN models. Figure 16 (A) represents that two neurons of HL BP-ANN classified 288 instances of drinkers (class 1) correctly and two sample cases were classified wrongly. On other hand, there are 205 instances of class 2 (non-drinkers) were classified correctly and fourteen instances wrongly. The total accuracy of the two HL neurons BP-ANN was nearly 96.9%. Figure 16 (B) shows that three neurons of HL BP-ANN correctly classified 281 instances of drinkers (class 1) and nine instances were classified wrongly, and 216 instances of

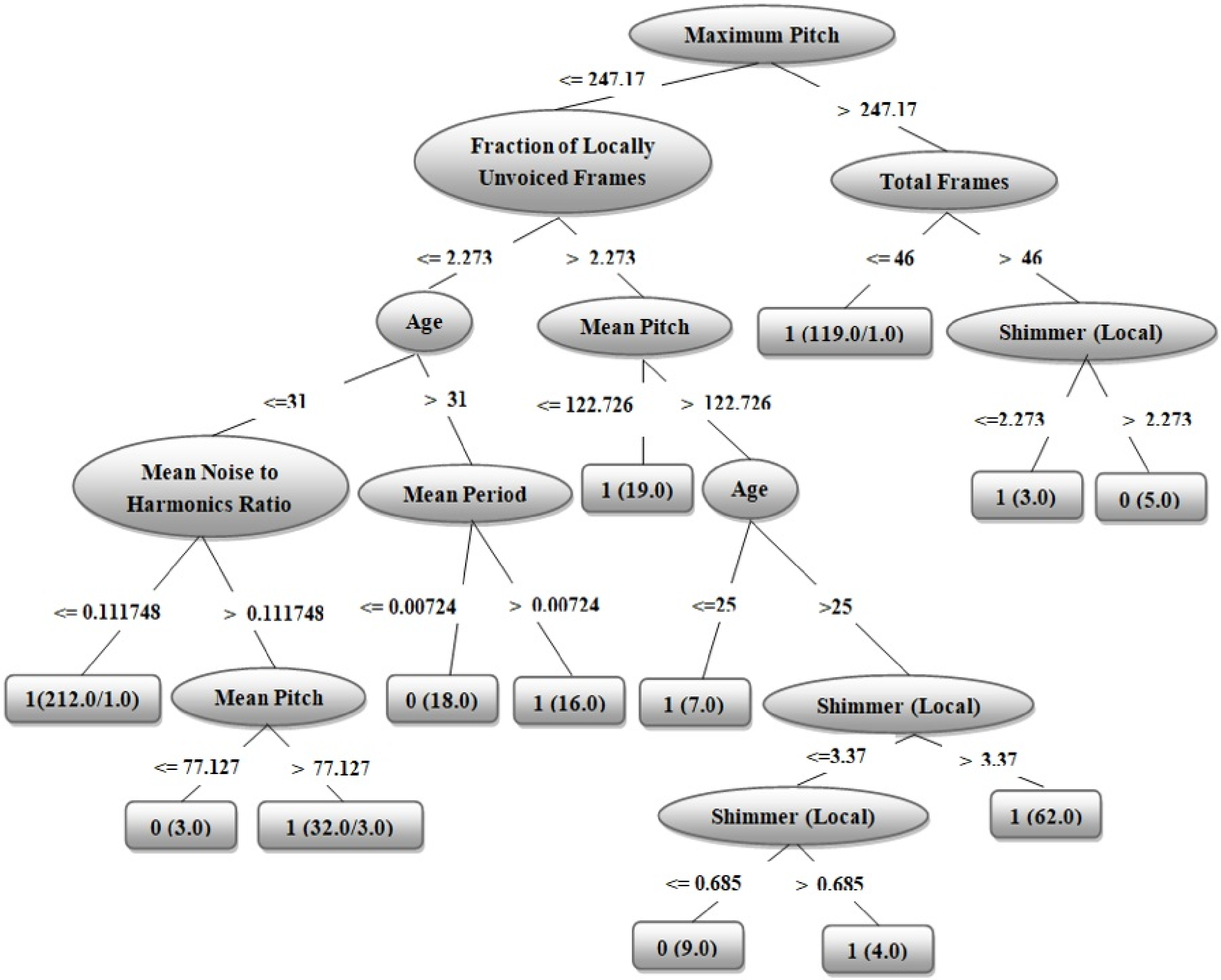



Figure 14. C4.5 Visualization Tree

TABLE VIII. PERFORMANCE PARAMETER VALUES OF NB'S

| NBs | AUC | CA | F1 | Precision | Recall |
|---|---|---|---|---|---|
| Class 2 | 0.934371 | 0.868369 | 0.851441 | 0.827586 | 0.876712 |
| Class 1 | 0.934371 | 0.868369 | 0.881834 | 0.902527 | 0.862069 |
| Avg. Wait | 0.934371 | 0.868369 | 0.868757 | 0.870283 | 0.868369 |

class 2 (non-drinkers) were classified correctly versus three sample instances wrongly. The total accuracy of the three HL neurons BP-ANN was 97.6%, which was better than two neurons HL BP-ANN. Figure 16 (C) recognizes four neurons of HL BP-ANN. It performed with the accuracy of class 1 was 98.6%, as well as 98.6% of the accuracy of class 2. Therefore, the total accuracy of the four HL neurons BP-ANN was nearly 98.6%. Figure 16 (D) depicts those five neurons of HL BP-ANN correctly classified 289 instances of drinkers (class 1), and one instance was classified wrongly. And there are 217 instances of class 2 (non-drinkers) that were classified correctly and two instances wrongly. The total accuracy of Five HL neurons BP-ANN was 99.4

Table X shows a competitive analysis of 2 to 5-neuron HL BP-ANN models with average performance parameters. We compared all the experimental 2 to 5-neuron HL BP-ANN models in terms of average performance (average weight of class 2 and class 1) of parameters such as CA, AUC, recall, precision, and F1-values. According to the study, the 5-neuron HL BP-ANN model shows the best per-

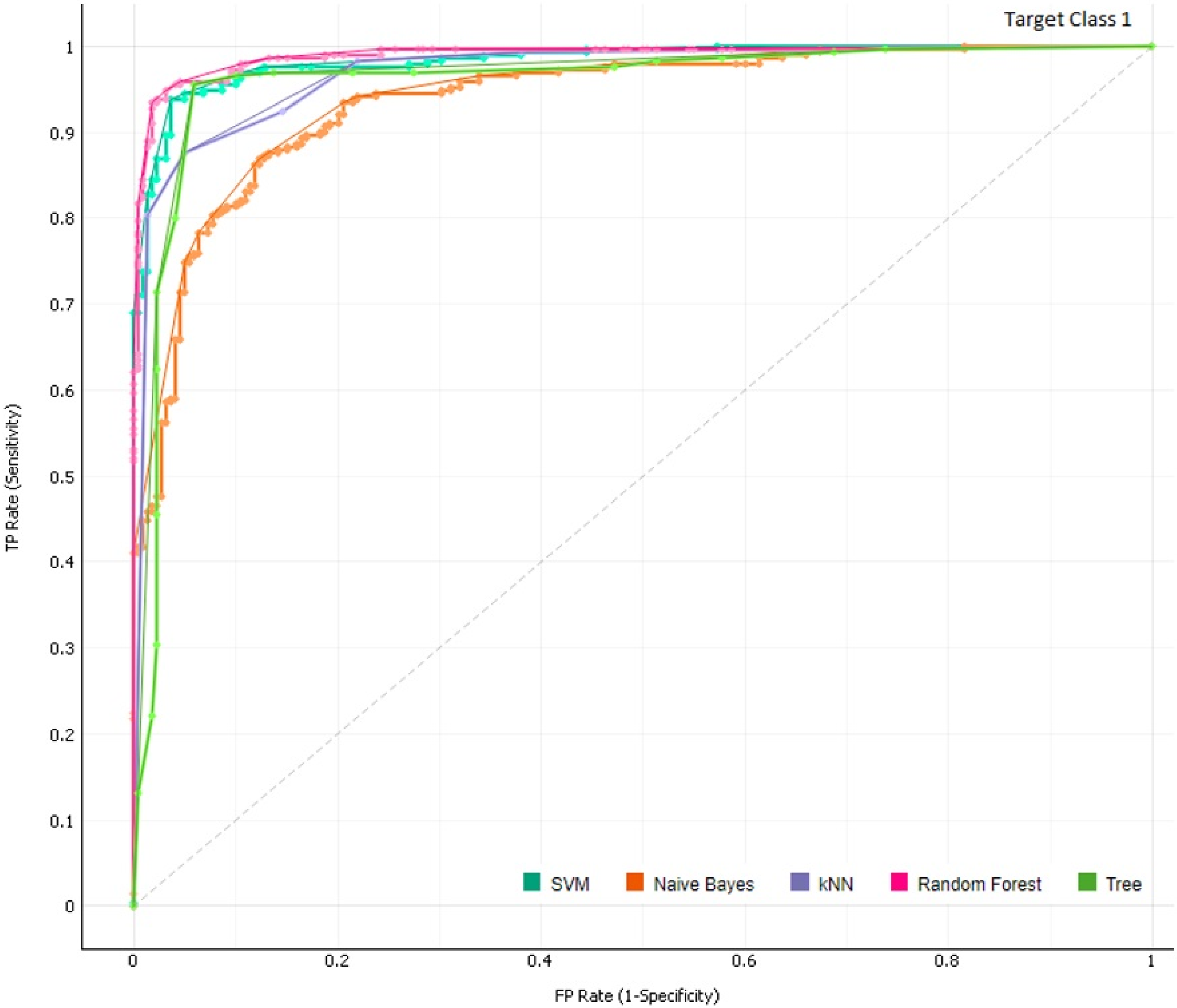



Figure 15. ROC curves of Experimental Machine Learning (ML) Algorithms

formance when compared to other experimental BP-ANN models, with a classification accuracy of 0.9941 (99.4%), precision, recall, and F1-values of 1. The superior values are represented with bold and * marks in Table X. The second highest performing model is the 4-neuron BP-ANN model, marked with the symbol +, with a classification accuracy of 0.99804 and AUC and precision values of 1. Therefore, we can conclude that the 5-neuron HL BP-ANN is the best model to predict drinkers with 100% accuracy. The test results also show 99.4% accuracy for this model.

**ROC Curves of 2 to 5 HL neurons BP-ANN Models and AUC Analysis:** Figure 17 shows the analysis of ROC curves for class 1 (drinkers) and class 2 (non-drinkers) neurons in the hidden layer (HL) of back-propagation artificial neural networks (BP-ANNs). The ROC is constructed between the values of specificity (X-axis) and sensitivity (Y-axis). The class curves are represented as blue (class 1) and green (class 2) colors in Figure 17.

**Best Training Performances of 2 to 5 HL Neurons BP-ANN Models and Comparative Analysis** Figure 18 depicts the study of the best training performances of 2-5 neurons (HL BP-ANNs). On the X-axis, the number of epochs is specified, and on the Y-axis, the MSE (mean squared error) values are specified. The line with dots indicates the best performance line indicator, and the red and blue color lines indicate the training and testing performances. Figure 18 (A) describes the two neurons' best HL BP-ANN ROC performance analysis. For this, the model uses 1000 epochs, and the performance value is 0.0069262.

TABLE IX. COMPETITIVE ANALYSIS WITH AVERAGE PERFORMANCE PARAMETER OF ML VALUES

| ML Model | CA | AUC | Recall | Precision | F1 |
|---|---|---|---|---|---|
| k-NN | 0.894 | 0.9689 | 0.894 | 0.894 | 0.893 |
| C4.5 | 0.949* | 0.958 | 0.949* | 0.949* | 0.948* |
| SVM | 0.936 | 0.983* | 0.935 | 0.936 | 0.935 |
| RFs | 0.953+ | 0.989+ | 0.953+ | 0.953+ | 0.952+ |
| NB | 0.869 | 0.934 | 0.869 | 0.870 | 0.868 |

TABLE X. COMPETITIVE ANALYSIS WITH AVERAGE PERFORMANCE PARAMETER OF 2 TO 5 HL BP-ANNS MODEL

| Model | CA | AUC | Recall | Precision | F1 |
|---|---|---|---|---|---|
| Two Neurons HL | 0.9686 | 0.9799 | 0.9536 | 0.9931 | 0.9730 |
| Three Neurons HL | 0.9764 | 0.9898 | 0.9894 | 0.9690 | 0.9791 |
| Four Neurons HL | 0.9862 | 0.9911 | 0.9862 | 0.9896 | 0.9879 |
| Five Neurons HL | 0.9941 | 1.0000 | 0.9931 | 0.9966 | 0.9948 |

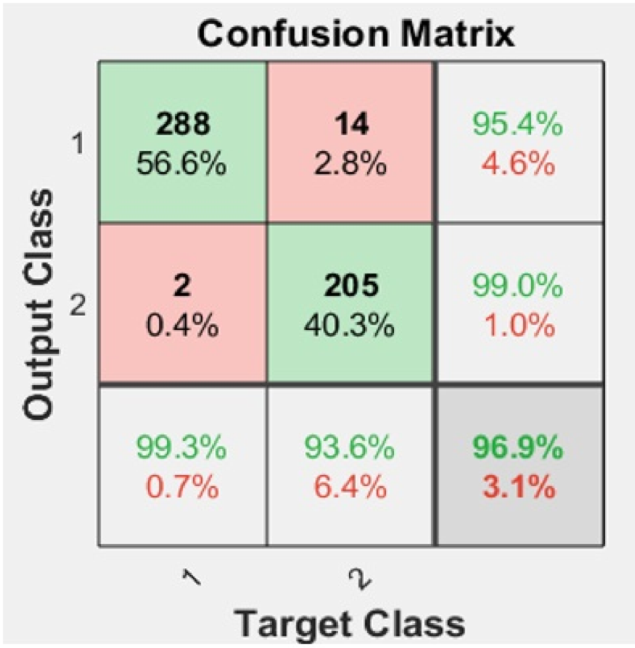
Figure 18 (B) specifies the three neurons' best HL BP-ANN ROC performance analysis. For this, the model uses 723 epochs, and the performance value is 0.0072902. Figure 18 (C) indicates the four HL BP-ANN ROC best performance analyses. In this, the model uses 1000 epochs, and the performance value is 0.0092679. Figure 18 (D) represents the five-HL BP-ANN ROC's best performance analyses. In this, the model uses 150 epochs, and the performance value is 0.0023297.

**Error Histograms Analysis of 2 To 5 Neurons HL BP-ANN Models:** Figure 19 shows the analysis of error histograms for 2–5 HL BP-ANN models. On the X-axis, the error values are specified as targets subtracted by the actual outputs, with targets represented as zero for non-drinkers and one for drinkers. The Y-axis specifies the number of instances. The blue and red bars indicate the number of training and testing instances with error values. The orange color bar indicates the zero-error line. The red color portions indicate the testing instances' error rate, and the blue color portions in the strip specify the training instances. Figure 19 (A) describes two neurons' HL BP-ANN error histogram with 20 bins. Most have error values of zero, and very few have error values of -0.05001 to +0.05001. The fewer data points' error values are very high (outside the boundary); some are -0.9502, and some are +0.9451. Figure 19 (B) specifies three neurons' HL BP-ANN error histogram with 20 bins. Most instances have error values of zero, and very few have error values of -0.04989 to +0.04989. The fewer data points' error values are very high; some are -0.9479, and some are +0.9479. Figure 19 (C) indicates four HL BP-ANN Error histograms with 20 bins. Most instances are in error values is zero, and very few are in the range of -0.04968 to +0.04968. Very few training data points are available for the boundary. Figure 19 (D) represents a five-HL BP-ANN error histogram with 20 bins. Almost all instances in the boundary error values are nearer to zero.

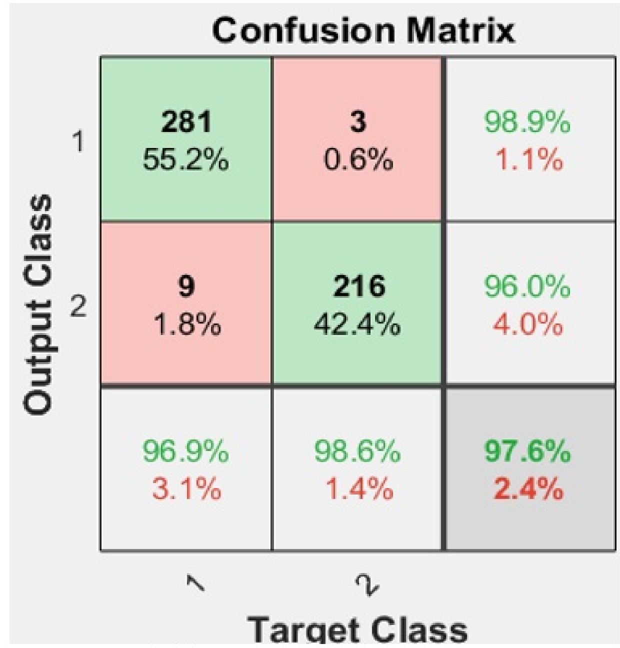
## 5. DISCUSSIONS AND COMPARATIVE STUDY

Alcohol consumption has been shown to have various effects on the human body, including changes to the voice [42]. Alcohol is a known irritant to the vocal cords, and excessive consumption can lead to inflammation and swelling. This can result in changes to the voice, including hoarseness and reduced vocal range[43].

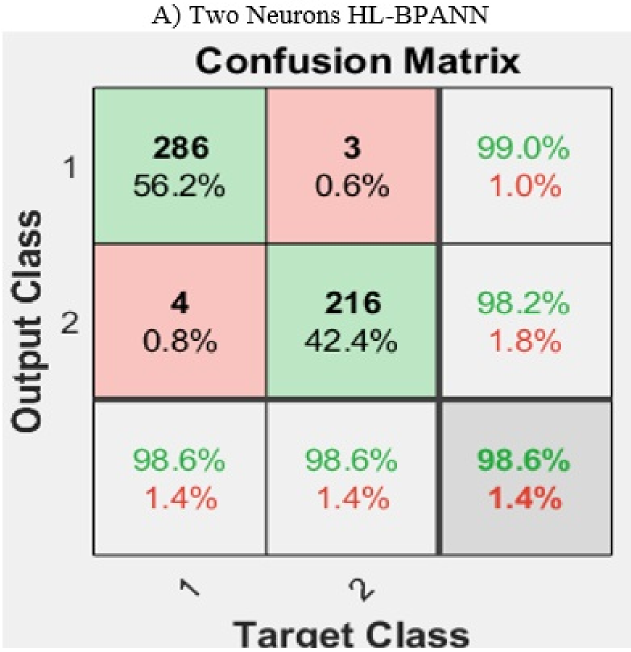
Chronic alcohol consumption has been linked to various voice disorders, including vocal nodules, polyps, and laryngeal cancer. These conditions can have long-lasting effects on the voice and may require medical treatment. Individuals who are dependent on alcohol and undergo detoxification may experience a range of withdrawal symptoms, including changes to their voice. This can include vocal tremors, hoarseness, and difficulty speaking. Research has shown that women may be more susceptible to the effects of alcohol on the voice, and older individuals may be at a higher risk for vocal changes due to alcohol consumption [44] [45]. Alcohol consumption affects the human body's motor system and leads to AUD. It directly affects the brain and leads to neurological diseases such as Alzheimer's, Parkinson's [46], and so on. This bad habit has an impact on all parts of the human body, particularly the liver, heart, and kidneys [47]. It is closely related to the acoustic (voice) system, resulting in errors in speaking words, lagging, wrong wording, and changes in voice parameters such as voicing, pulses, and fundamental frequency (F0). Some research has shown this to be proper concerning changes. Many researchers focused on differentiating differences in voice parameter values in conditions of sobering and intoxication. Klinghol et al. (1988) [48] researched the relationship between intoxication of alcohol in low-level and speech signals. For this research, they used 11 male people's voices of text-reading words in alcohol-intoxicated and sober situations. They determined pitch (F0), SNR (signal-to-noise ratio), 1st (F1) to 2nd (F2) format frequency ratio, LTAS, and frequency speeds and determined differences in sobriety and intoxication. As per observations, they found SNR, F0, and LTAS were discriminated against by 5% error in both (somber and intoxication) conditions. Liquor intoxication is known to influence the characteristics of the human way of behaving and consciousness. The system that




A) Two Neurons HL-BPANN




B) Three Neurons HL-BPANN



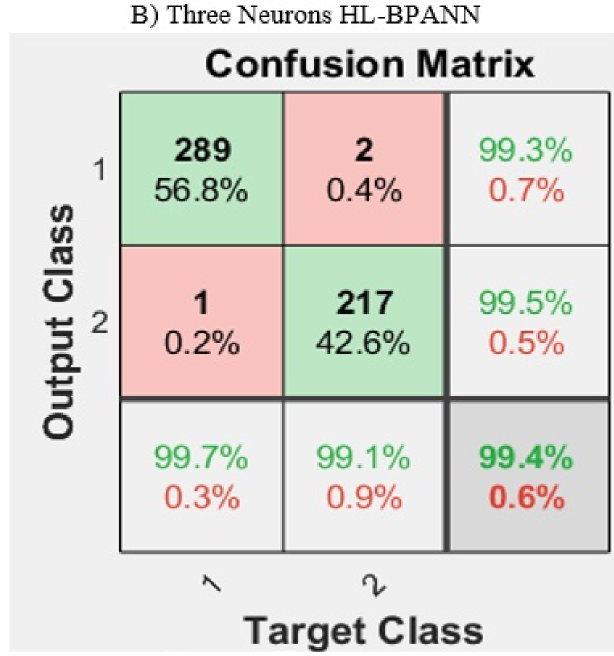




Figure 16. Confusion Matrices of 2 to 5 HL BP-ANNs Models

has the most impact is speech production after beverages. Offrede et al. (2021) [49] researched speech languages and alcohol consumption. This research mainly aims to compare the pronunciations of L1-Duch and L2-English after an alcoholic beverage. For the experiment setup, they chose 80 individuals who speak Dutch and English and are natives of Nederland, aged 20 to 64 (mean values are 31 and SD is 9.5). The speaking English rate is 3 to 9 out of 10, the mean value is 7.5, and the SD is 1.2. The BAC (blood alcohol concentration) levels ranged from 0.8 to 1.59. An additive-mixed general methodology was used for this analysis. As per the analysis of results, participants' BAC affected negative pronunciation in L1 and had no critical outcome on the L2 language pronunciations. Table XII shows some research on alcohol and acoustics, with detailed results and analysis.

Some of the research is covered by the acoustic and facial expressions of the drinkers. Some research has deeply described the culture, behavior, way of speaking, and social

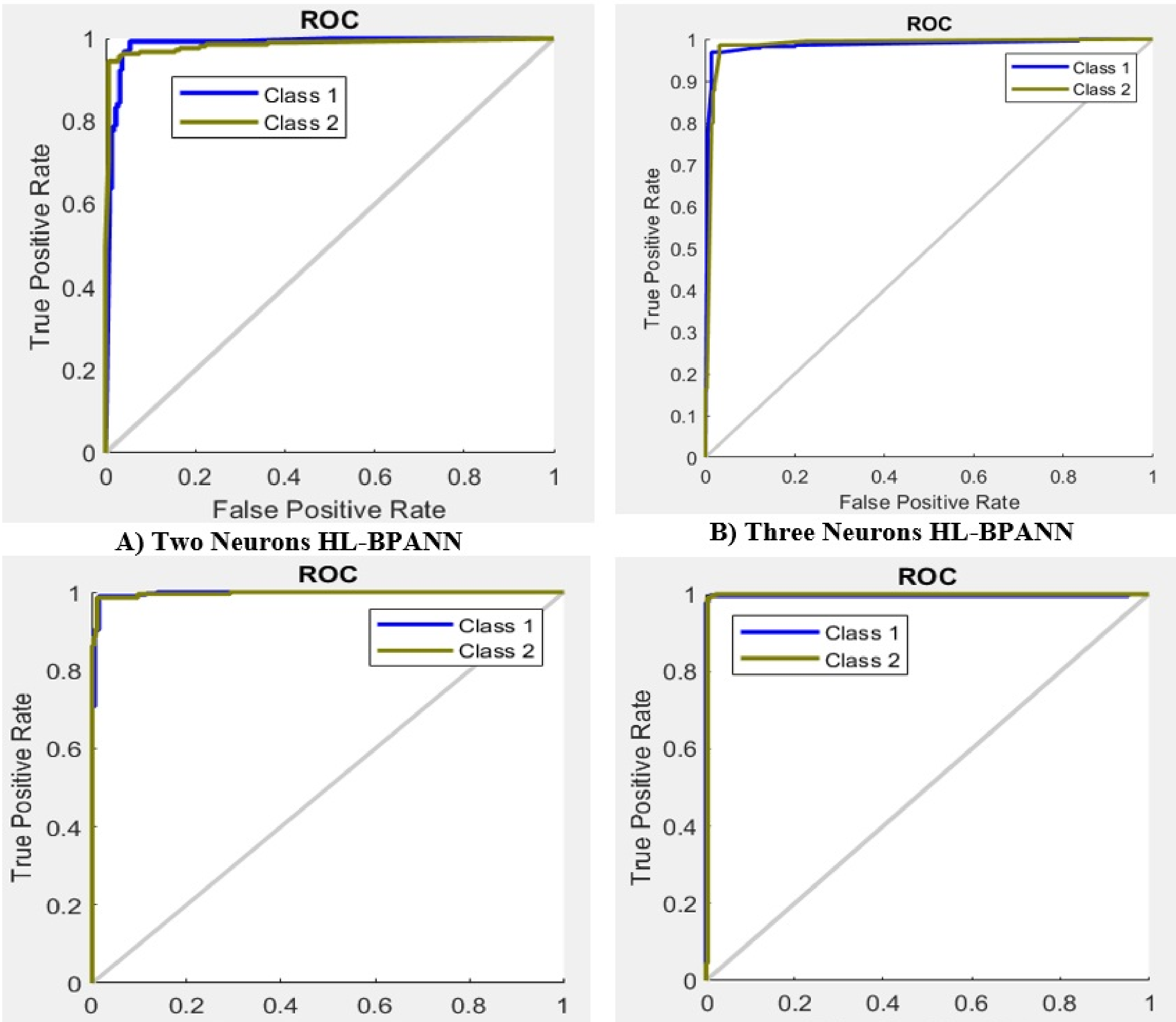



Figure 17. ROC class 2(non-drinkers) and class 1 (drinkers) curves analysis of 2 to 5 neurons HL BP-ANNs Models

issues with alcohol consumption. A healthy voice represents a healthy mind without neurological diseases like Parkinson's, Alzheimer's, and dementia[56] [57]. Sayette et al. (2012) [58] researched alcohol consumption's effects on social bonding and emotions. For this experiment, they chose 360 male and 360 female alcohol consumers and derived three groups that were unfamiliar to everyone. Participants consumed a dose of moderate alcohol for over 36 minutes. Meanwhile, the sequence recorded facial and speech behaviors. The outcomes demonstrated that alcohol works with binding during group establishment. The impact of alcohol on social reactions during gatherings and collaborative interactions is concerning. Kirchner et al. (2006) [59] researched alcohol effects on male drinkers' group formations. In this research, they chose 54 male social consumers collected into 3-man clusters of unknown individuals (strangers). All individuals from each cluster group were directed to consume either 0.82 g/kg of alcohol or a medicinal drug (placebo) for 30 minutes. Alcohol consumption enhanced and expanded individual and group coordination levels with smiling and ways of behaving with speech over the long haul, resulting in better bonding. The dysfunction of the focal nervous system bringing about mental deficiency is likely the most damaging element of FAS (fetal alcohol syndrome). Greene et al. (1990) [60] examined the effects of FAS on speech and language. The count of inconsistencies and the extent of birth weight are more delicate marks of FAS than language exploitation.

**Comparative Study**

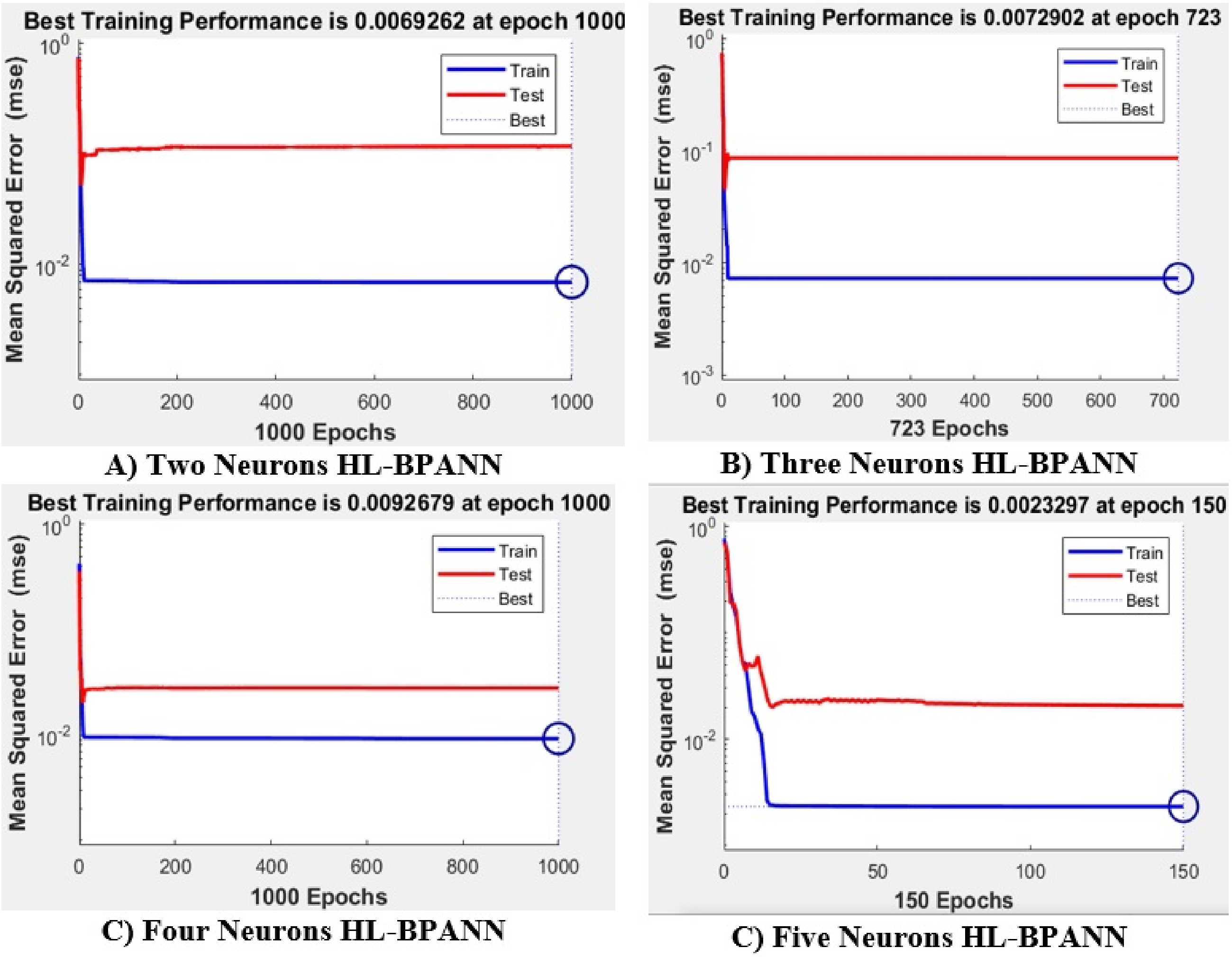
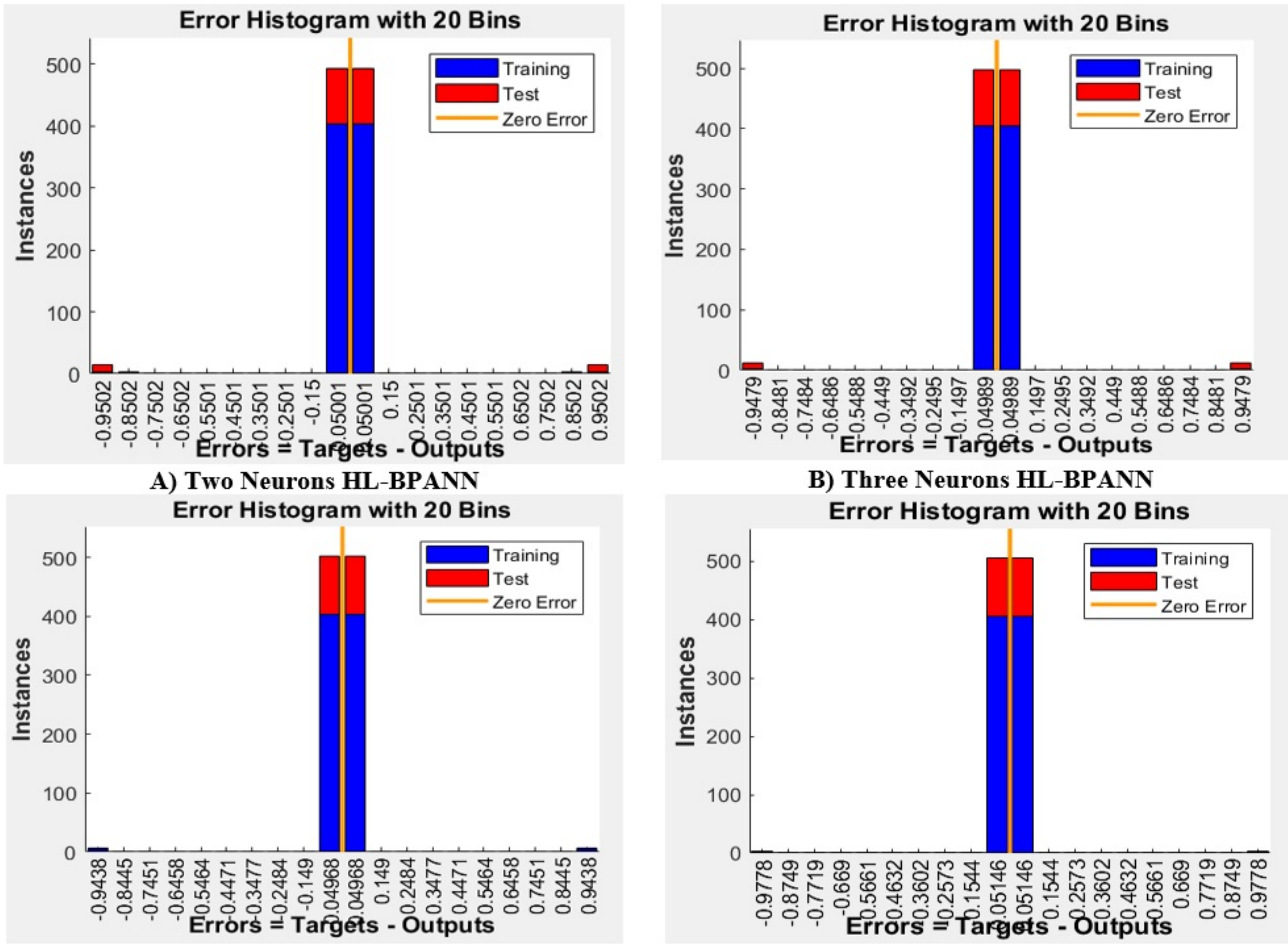



Figure 18. Best performance analysis of 2 to 5 HL BP-ANNs Models

TABLE XI. ANALYSIS OF BP-ANN MODELS TIME AND PERFORMANCES LIKE MSE, MU VALUES, R VALUE, ACCURACY VALUES

| BP-ANN Model | Epoch | Time in Sec. | MSE | Gradient | R Value | Accuracy (%) |
|---|---|---|---|---|---|---|
| Two Neurons HL | 1000 | 0.26 | 0.006926 | $4.4593 \times 10^{-6}$ | 0.94112 | 96.856581 |
| Three Neurons HL | 723 | 0.17 | 0.007290 | $9.9071 \times 10^{-8}$ | 0.95327 | 97.642436 |
| Four Neurons HL | 1000 | 0.21 | 0.009267 | $4.9862 \times 10^{-7}$ | 0.97364 | 98.624754 |
| Five Neurons HL | 150 | 0.09 | 0.002329 | $9.3724 \times 10^{-8}$ | 0.98796 | 99.410609 |

According to a comparison with the background research work, this research approach is a novel and low-cost technology for the early identification of alcoholics with smartphone apps using voice signals. The model of five neurons (HL BP-ANN) is suitable as it classifies 99.4% of drinkers and non-drinkers. Figure 25 shows how each experimental ML model compares to others using the CA and AUC performance parameter values. All ML models have been represented on the X-axis, and the Y-axis represents the performance values (0 to 1). The blue bar represents the AUC values, and the brown bar specifies the CA from 0 to 1. The visualization reports indicate that the RF model's AUC and CA values (0.989 and 0.953) are superior to other experimental ML models. The SVM model's AUC value (0.983) is the second highest, and the CA value (0.936) is in the third position in the comparison. According to the study and critical analysis of the background research work, the current research approach is a novel and low-cost technology for identifying alcoholics with smartphone apps using voice signals.




**A) Two Neurons HL-BPANN**

**B) Three Neurons HL-BPANN**

**C) Four Neurons HL-BPANN**

**D) Five Neurons HL-BPANN**

Figure 19. Error Histograms analysis of 2 to 5 HL BP-ANNs Models

The model of five neurons (HL BP-ANN) is suitable, as it classifies 99.4% of drinkers and non-drinkers. Figure 25 represents each experimental ML model's comparison to others using the CA and AUC performance parameter values. All ML models are specified on the x-axis, and the y-axis represents the performance values (0 to 1). The blue bar represents the AUC values, and the brown bar specifies the CA from 0 to 1. The visualization reports say the RF model's AUC and CA values (0.989 and 0.953) are superior to other experimental ML models. Accuracy (CA) of Experimental Machine Learning (ML) Models. The SVM model's AUC value (0.983) is the second highest, and its CA value (0.936) is in the third position in the comparison. Again, we conducted the statistical analysis in both classes and obtained the total statistical results for the dataset. The collected information contained age groups of 22 to 34 years for male persons' data and hidden voice record values. Class 1 describes the drinkers, and Class 2 specifies the non-drinkers' specifications. According to the statistical results, all mean and median values of pitch (mean, median, SD, min, and max) in Class 1 (drinkers) set are higher than those in Class 2 (non-drinkers), as is the case for the entire set. The number of unvoiced mean values in Class 1 is higher than in Class 2. All mean and median values of the jitter and shimmer voice parameters in Class 1 are higher than those of the Class 2 set. Table XIII specifies the minimum and maximum of all voice parameter values concerning Class 1, Class 2, and the total dataset. The age group of 22 to 34-year-olds is involved in all Class 2, Class 1, and total dataset categories. It describes the detailed analysis of the minimum and maximum of every voice parameter value, such as mean pitch, median pitch, jitter, shimmer, harmonic ratio, etc. According to the critical analysis, the present research work is novel and suitable for the classification and predictions of alcohol drinkers. It is more sophisticated for early detection technology than other systems and tools. As per ML model analysis, the RF model is more efficient at 95.3% than other experimental ML models. In the general external background survey and internal comparative study, the proposal model 5-neurons

TABLE XII. EXTERNAL RESEARCH WORKS ON ALCOHOL EFFECTS ON ACOUSTIC ANALYSIS

| Ref.No. | Author Ref. | Research Models and Description | Data set and Results |
|---|---|---|---|
| [50] | Leung et al. (2019) | The main aim is a health risk in Alcohol use of the youth and sex contrasts in liquor in LMIC. This examination exposed the sex divergences in the consequences and preponderance of liquor use among young people living in middle and low-income nations. | 271,156 students participated, age of 13 to 17 years Alcohol Use (Females = 56.15% and males=59.74%) from GSHS reports from 68 countries. As per reports, the male person had higher chances of alcohol or liquor use (OR=2.38 [1.91-2.96]), a history of intoxication (OR=2.64 [2.11-3.31]), and liquor-related issues (OR=1.72 [1.41-2.10]) than females. |
| [51] | Spindler et al. (2021) | Researched AUD- (Alcohol Use Disorder) is related to GM-grey Matter volume. | 27 studies on AUD patients - 1045 and healthy controls -1054. GM decrease in AUD could interrupt the neuron network correspondence responsible for the neurocognitive damages concerned with high-pitched chronic alcohol utilization. |
| [52] | Vogel et al. (2021) | The main aim of the research is speech production effects linked with tobacco and alcohol versus CANNabis use. For this, they calculate acoustic parameters. | Gathering data from 40 control individuals using alcohol and tobacco, and 31 groups using cannabis. Alcohol (drinks) 90% Controls (900 ± 2460), cannabis 100% (2678 ± 3765), p-value¡0.001, effect size- 0.71. Tobacco (cigarettes) 28% Controls (22 ± 54), cannabis 77% (860 ± 1832), p-value¡0.001, effect size-0.61. |
| [53] | Davidson et al. (2001) | The theme of this research was evaluating the stimulant impacts of alcohol on humans and their behaviors. | 19 subjects were analyzed. BAES level: t-tests (t = 2.66, P¿.005) found consumed alcohol compared to soda, the values are 5.0 ± 2.3 S.E.M. (alcohol) vs 3.8 ± 2.3 S.E.M. (soda). |
| [54] | Braun et al. (2003) | Studied the effects of alcohol on speech, prosodic changes, fundamental frequency (F0), and breath alcohol concentration (BRAC). | Data from 33 male drinkers, aged 19-24. Syllables produced: sober 161 ± 46, intoxicated 192 ± 54. F0 at max BRAC ¡0.08%: sober 126.8, intoxicated 133.7. |
| [55] | Andrews et al. (1977) | Effects of alcohol on the speaker's speech of sober and intoxicated. | Voice data from 27 students. Parameters in sober and inebriated conditions showed significant differences, with t-test values indicating statistical significance for various characteristics. |

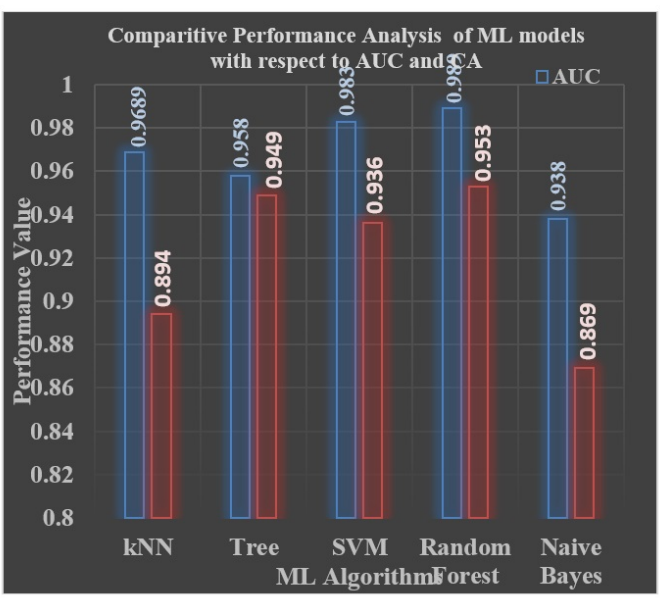



Figure 20. Comparative Analysis of AUC and CA of ML Models

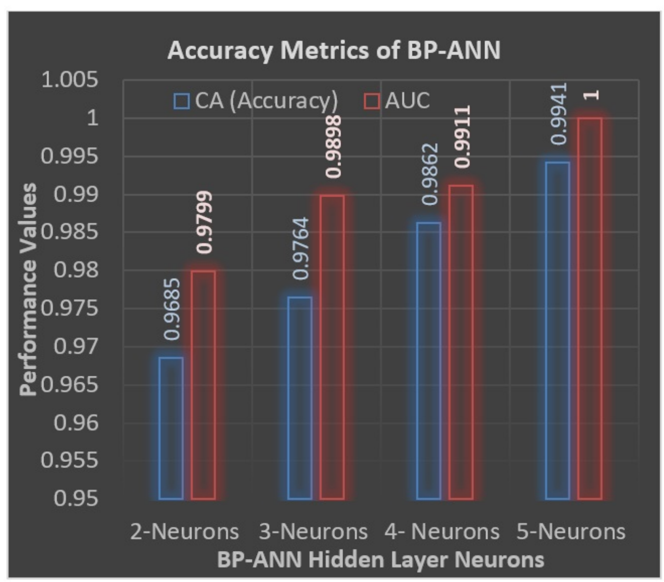



Figure 21. Competitive analysis with AUC and CA of 2 to 5 neurons HL BP-ANNs Models

HL BP-BP-ANN is the best solution for acoustic alcoholics' problems, with a classification accuracy of 99.4%.

## 6. CONCLUSION

Heavy alcohol consumption is one of the evils in society because it impacts the socioeconomic system and social and family lives. It also affects the human vocal cords, changing their voice at the time of alcohol consumption for various reasons. The Intelligent Novel Approach for Identification of Alcohol Consumers Using an ANN-Based Model on Vowelized Voice Dataset is a method for using artificial neural networks (ANNs) to identify individuals who consume alcohol based on their voice patterns. The approach involves collecting a dataset of vowelized voice samples from both alcohol consumers and non-consumers. The voice samples are pre-processed to extract relevant features to distinguish between the two groups. These features include pitch, tone, and other acoustic characteristics of the voice. The pre-processed voice samples are then used to train an ANN model. The model is designed to learn the patterns and relationships between the extracted features and the corresponding labels of alcohol consumers or non-consumers. Once trained, the model can predict the label of new, previously unseen voice samples. This novel study described the auto-detection of alcohol consumers using vowelized (/a /e /i /o /u) voice data with machine learning and neural network models. In this, five eminent existing machine learning models were used, such as Naïve Bayes (NB), Random Forest (RF), k-NN, SVM, and C4.5 (Tree). The RF algorithm was performed in existing machine learning models, with a classification accuracy value of 0.953. On the other hand, we used BP-ANN models in which we incremented the neurons in Hidden Layers 2 to 5 and trained again for each increment step. The 5-neuron HL BP-ANN model performed better than all other experimental models, with 99.4% classification accuracy. Therefore, the 5-HL neuron BP-ANN is the best solution for this dataset. In this research, the exemplars used were only male individuals. So, further, we aim to include female drinker individuals, measure alcohol consumption quantity and quality for the detection training process with better models (PNN, 1-D CNN, and stacked auto-encoders), and construct equipment tools named the Voice Intention Transcript Alcohol Level Identification Tool.

**ACKNOWLEDGEMENT** The author gratefully acknowledges the Aditya Institute of Technology and Management, Tekkali, for providing the resources essential to this research. Thanks to Director Prof. VV Nageswara Rao, Principal Dr A. S. Srinivasa Rao, and the FIST Lab central coordinator of AITAM for their unwavering support and assistance.

## References

[1] S. Gutterman, "Alcohol and mortality: An actuarial issue," *North American Actuarial Journal*, vol. 26, no. 2, pp. 184–204, 2022.

[2] W. H. Organization, "Alcohol," 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/alcohol

[3] J. M. Ford, B. J. Roach, K. W. Jorgensen, J. A. Turner, G. G. Brown, R. Notestine, A. Bischoff-Grethe, D. Greve, C. Wible, J. Lauriello *et al.*, "Tuning into the voices: a multisite fmri study of auditory hallucinations," *Schizophrenia bulletin*, vol. 35, no. 1, pp. 58–66, 2009.

[4] V. Balasankar, S. S. V. Penumatsa, and T. P. R. Vital, "Empirical statistical analysis and cluster studies on socio-economic status (ses) dataset," in *IOP Conference Series: Materials Science and Engineering*, vol. 1085, no. 1. IOP Publishing, 2021, p. 012030.

[5] V. A. Eashwar, R. Umadevi, and S. Gopalakrishnan, "Alcohol consumption in india–an epidemiological review," *Journal of family medicine and primary care*, vol. 9, no. 1, p. 49, 2020.

[6] P. Schultz, "Vocal fold cancer," *European annals of otorhinolaryngology, head and neck diseases*, vol. 128, no. 6, pp. 301–308, 2011.

[7] T. V. Wang and P. C. Song, "Neurological voice disorders: a review," *International Journal of Head and Neck Surgery*, vol. 13, no. 1, pp. 32–40, 2022.

[8] K. Cunningham and S. Nandamudi, "Interprofessional approaches to acoustic voice analysis," *Voice and Speech Review*, vol. 16, no. 3, pp. 292–303, 2022.

[9] P. B. Dasgupta, "Detection and analysis of human emotions through voice and speech pattern processing," *arXiv preprint arXiv:1710.10198*, 2017.

[10] K. Johnson, D. B. Pisoni, and R. H. Bernacki, "Do voice recordings reveal whether a person is intoxicated? a case study," *Phonetica*, vol. 47, no. 3-4, pp. 215–237, 1990.

[11] G. W. Wakista, S. J. Abhayarathne, G. T. Mendis, S. M. Arunatileka, K. D. Sandaruwan, S. A. Gunawardena, and R. Fernando, "Effect of alcohol on suprasegmental features of voice," in *2014 4th World Congress on Information and Communication Technologies (WICT 2014)*. IEEE, 2014, pp. 269–274.

[12] H. Hollien, G. DeJong, C. A. Martin, R. Schwartz, and K. Liljegren, "Effects of ethanol intoxication on speech suprasegmentals," *The Journal of the Acoustical Society of America*, vol. 110, no. 6, pp. 3198–3206, 2001.

[13] Z. Ma, C. Bullen, J. T. W. Chu, R. Wang, Y. Wang, and S. Singh, "Towards the objective speech assessment of smoking status based on voice features: a review of the literature," *Journal of Voice*, vol. 37, no. 2, pp. 300–e11, 2023.

[14] F. Schiel, C. Heinrich, and S. Barfüsser, "Alcohol language corpus: the first public corpus of alcoholized german speech," *Language resources and evaluation*, vol. 46, pp. 503–521, 2012.

[15] S. M. Landman, "Qualitative analysis of alcohol's acute effect on vocal range," *Alpenglow: Binghamton University Undergraduate Journal of Research and Creative Activity*, vol. 4, no. 1, p. 8, 2018.

[16] H. Namazi, M. R. Ashfaq Ahamed, M. H. Babini, and O. Krejcar, "Analysis of the correlation between the human voice and brain activity," *Waves in Random and Complex Media*, pp. 1–13, 2021.

[17] E. Tisljár-Szabó, R. Rossu, V. Varga, and C. Pléh, "The effect of

TABLE XIII. PROPOSAL MODEL VERSUS ALL OTHER EXPERIMENTAL MODELS

| Models | CA (Accuracy) | AUC | Recall | Precision | F1 |
|---|---|---|---|---|---|
| k-NN | 0.894 | 0.9689 | 0.894 | 0.894 | 0.893 |
| C4.5 | 0.949 | 0.958 | 0.949 | 0.949 | 0.948 |
| SVM | 0.936 | 0.983 | 0.935 | 0.936 | 0.935 |
| RFS | 0.953 | 0.989 | 0.953 | 0.953 | 0.952 |
| Naive Bayes | 0.869 | 0.934 | 0.869 | 0.870 | 0.868 |
| 2-HL ANN | 0.9685 | 0.9799 | 0.9536 | 0.9931 | 0.9729 |
| 3-HL ANN | 0.9764 | 0.9898 | 0.9894 | 0.9689 | 0.9790 |
| 4-HL ANN | 0.9862 | 0.9911 | 0.9862 | 0.9896 | 0.9879 |
| 5-HL ANN | 0.9941 | 1 | 0.9931 | 0.9966 | 0.9948 |

alcohol on speech production," *Journal of psycholinguistic research*, vol. 43, pp. 737–748, 2014.

[18] C. R. Olson, D. C. Owen, A. E. Ryabinin, and C. V. Mello, "Drinking songs: alcohol effects on learned song of zebra finches," *PLoS One*, vol. 9, no. 12, p. e115427, 2014.

[19] W. Wang, H. Wu, and M. Li, "Deep neural networks with batch speaker normalization for intoxicated speech detection," in *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE, 2019, pp. 1323–1327.

[20] H. Kang, S. Kim, and J. Yoo, "Voice hygiene habits and the characteristics of korean voice-related quality of life (k-vrqol) among classical singers," *Phonetics and Speech Sciences*, vol. 10, no. 3, pp. 49–59, 2018.

[21] H. Terband, M. Spruit, and B. Maassen, "Speech impairment in boys with fetal alcohol spectrum disorders," *American journal of speech-language pathology*, vol. 27, no. 4, pp. 1405–1425, 2018.

[22] B. Schuller, S. Steidl, A. Batliner, F. Schiel, J. Krajewski, F. Weninger, and F. Eyben, "Medium-term speaker states—a review on intoxication, sleepiness and the first challenge," *Computer Speech & Language*, vol. 28, no. 2, pp. 346–374, 2014.

[23] S. Singer, M. Merbach, A. Dietz, and R. Schwarz, "Psychosocial determinants of successful voice rehabilitation after laryngectomy," *Journal of the Chinese Medical Association*, vol. 70, no. 10, pp. 407–423, 2007.

[24] M. Van Puyvelde, X. Neyt, F. McGlone, and N. Pattyn, "Voice stress analysis: a new framework for voice and effort in human performance," *Frontiers in psychology*, vol. 9, p. 1994, 2018.

[25] L. Liu and T. F. Jaeger, "Talker-specific pronunciation or speech error? discounting (or not) atypical pronunciations during speech perception." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 45, no. 12, p. 1562, 2019.

[26] G. Hendricks, S. Malcolm-Smith, C. Adnams, D. J. Stein, and K. A. M. Donald, "Effects of prenatal alcohol exposure on language, speech and communication outcomes: A review longitudinal studies," *Acta neuropsychiatrica*, vol. 31, no. 2, pp. 74–83, 2019.

[27] L. Liu and T. F. Jaeger, "Inferring causes during speech perception," *Cognition*, vol. 174, pp. 55–70, 2018.

[28] O. Cooney, "Acoustic analysis of the effects of alcohol on the human voice," Ph.D. dissertation, Dublin City University, 1998.

[29] M. Gateway, "Losing Voice After Drinking — musicgateway.com," https://www.musicgateway.com/blog/how-to/alcohol-singers-how-does-it-affect-your-voice, [Accessed 02-03-2024].

[30] S. B. Chin and D. B. Pisoni, *Alcohol and speech*. Academic Press New York, 1997.

[31] Y.-C. Jung and K. Namkoong, "Alcohol: intoxication and poisoning–diagnosis and treatment," *Handbook of clinical neurology*, vol. 125, pp. 115–121, 2014.

[32] D. Abur, M. K. MacPherson, A. C. Shembel, and C. E. Stepp, "Acoustic measures of voice and physiologic measures of autonomic arousal during speech as a function of cognitive load in older adults," *Journal of Voice*, vol. 37, no. 2, pp. 194–202, 2023.

[33] S. S. Shaabdurakhimova, "Voice formation mechanism and causes of voice impairment," *ACADEMICIA: An International Multidisciplinary Research Journal*, vol. 11, no. 3, pp. 2259–2265, 2021.

[34] H. Cai, "Acoustic analysis of resonance characteristics of head voice and chest voice," in *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 2019, pp. 1–6.

[35] "Alcohol Use Disorder (AUD) Treatment: MedlinePlus — medlineplus.gov," https://medlineplus.gov/alcoholusedisorderaudtreatment.html, [Accessed 22-08-2022].

[36] "Alcohol Use Disorder: What It Is, Risks & Treatment — my.clevelandclinic.org," https://my.clevelandclinic.org/health/diseases/3909-alcoholism, [Accessed 26-08-2022].

[37] P. V. Terlapu, R. P. R. Sadi, R. K. Pondreti, and C. R. Tippana, "Intelligent identification of liver diseases based on incremental hidden layer neurons ann model," *International Journal of Computing and Digital System*, 2021.

[38] P. V. Terlapu, S. B. Gedela, V. K. Gangu, and R. Pemula, "Intelligent diagnosis system of hepatitis c virus: A probabilistic neural network based approach," *International Journal of Imaging Systems and Technology*, vol. 32, no. 6, pp. 2107–2136, 2022.

[39] P. V. Terlapu, M. Yugandhar, B. Ramesh, B. V. Kumar, and R. Pemula, "Student cognitive learning capability (sclc) prediction system using pca-ann based model," in *2022 International Conference on Computing, Communication and Power Technology (IC3P)*. IEEE, 2022, pp. 11–18.

[40] V. Balasankar, S. Penumatsa, and P. Terlapu, "Intelligent socio-

economic status prediction system using machine learning models on rajahmundry ap, ses dataset," *Indian Journal of Science and Technology*, vol. 13, no. 37, pp. 3820–3842, 2020.

[41] T. P. Vital, "Intellectual gestational diabetes diagnosis system using mlp-whale optimization algorithm including statistical analysis," *International Journal of Computing and Digital Systems*, vol. 14, no. 1, pp. 1–1, 2023.

[42] P. R. V. Terlapu, D. Jayaram, S. Rakesh, M. V. Gopalachari, B. Ramana, N. Tangudu, and K. R. Kalidindi, "Optimizing chronic kidney disease diagnosis in uddanam: A smart fusion of ga-mlp hybrid and pca dimensionality reduction," *Procedia Computer Science*, vol. 230, pp. 522–531, 2023.

[43] A. S. Resuli, C. Cingi, and E. De Corso, "Vocal hoarseness (trachyphonia) causes and treatment," in *Airway diseases*. Springer, 2023, pp. 1–9.

[44] L. Greaves, N. Poole, and A. C. Brabete, "Sex, gender, and alcohol use: implications for women and low-risk drinking guidelines," *International journal of environmental research and public health*, vol. 19, no. 8, p. 4523, 2022.

[45] M. P. Messina, A. D'Angelo, G. Battagliese, G. Coriale, L. Tarani, S. Pichini, D. Rasio, G. Parlapiano, M. Fiore, C. Petrella *et al.*, "Fetal alcohol spectrum disorders awareness in health professionals: Implications for psychiatry," *Rivista di Psichiatria*, vol. 55, no. 2, pp. 79–89, 2020.

[46] T. P. R. Vital, J. Nayak, B. Naik, and D. Jayaram, "Probabilistic neural network-based model for identification of parkinson's disease by using voice profile and personal data," *Arabian Journal for Science and Engineering*, vol. 46, no. 4, pp. 3383–3407, 2021.

[47] T. P. Vital, "Empirical study on uddanam chronic kidney diseases (uckd) with statistical and machine learning analysis including probabilistic neural networks," in *Handbook of Computational Intelligence in Biomedical Engineering and Healthcare*. Elsevier, 2021, pp. 283–314.

[48] F. Klingholz, R. Penning, and E. Liebhardt, "Recognition of low-level alcohol intoxication from speech signal," *The Journal of the Acoustical Society of America*, vol. 84, no. 3, pp. 929–935, 1988.

[49] T. F. Offrede, J. Jacobi, T. Rebernik, L. De Jong, S. Keulen, P. Veenstra, A. Noiray, and M. Wieling, "The impact of alcohol on l1 versus l2," *Language and speech*, vol. 64, no. 3, pp. 681–692, 2021.

[50] J. Leung, V. Chiu, J. P. Connor, A. Peacock, A. B. Kelly, W. Hall, and G. C. Chan, "Alcohol consumption and consequences in adolescents in 68 low and middle-income countries–a multi-country comparison of risks by sex," *Drug and alcohol dependence*, vol. 205, p. 107520, 2019.

[51] C. Spindler, S. Trautmann, N. Alexander, S. Bröning, S. Bartscher, M. Stuppe, and M. Muehlhan, "Meta-analysis of grey matter changes and their behavioral characterization in patients with alcohol use disorder," *Scientific reports*, vol. 11, no. 1, p. 5238, 2021.

[52] A. P. Vogel, V. Pearson-Dennett, M. Magee, R. A. Wilcox, A. Esterman, D. Thewlis, J. M. White, and G. Todd, "Adults with a history of recreational cannabis use have altered speech production," *Drug and Alcohol Dependence*, vol. 227, p. 108963, 2021.

[53] D. Davidson, K. Hutchison, C. Dagon, and R. Swift, "Assessing the stimulant effects of alcohol in humans," *Pharmacology Biochemistry and Behavior*, vol. 72, no. 1-2, pp. 151–156, 2002.

[54] A. Braun, H. J. Künzel, D. Recasens, J. Romero, and M. Solé, "The effect of alcohol on speech prosody," in *Proceedings of the international congress of phonetic sciences, barcelona*, vol. 2645, 2003, p. 2648.

[55] M. L. Andrews, W. M. Cox, and R. G. Smith, "Effects of alcohol on the speech of non alcoholics," 1977.

[56] H. M. Fletcher, M. J. Drinnan, and P. N. Carding, "Voice care knowledge among clinicians and people with healthy voices or dysphonia," *Journal of voice*, vol. 21, no. 1, pp. 80–91, 2007.

[57] T. d. C. Moreira, C. Gadenz, L. R. Figueiró, D. M. Capobianco, K. Cunha, M. Ferigolo, H. M. Barros, and M. Cassol, "Substance use, voice changes and quality of life in licit and illicit drug users," *Revista CEFAC*, vol. 17, pp. 374–384, 2015.

[58] M. A. Sayette, K. G. Creswell, J. D. Dimoff, C. E. Fairbairn, J. F. Cohn, B. W. Heckman, T. R. Kirchner, J. M. Levine, and R. L. Moreland, "Alcohol and group formation: A multimodal investigation of the effects of alcohol on emotion and social bonding," *Psychological science*, vol. 23, no. 8, pp. 869–878, 2012.

[59] T. R. Kirchner, M. A. Sayette, J. F. Cohn, R. L. Moreland, and J. M. Levine, "Effects of alcohol on group formation among male social drinkers," *Journal of Studies on Alcohol*, vol. 67, no. 5, pp. 785–793, 2006.

[60] T. Greene, C. B. Emhart, S. Martier, R. Sokol, and J. Ager, "Prenatal alcohol exposure and language development," *Alcoholism: Clinical and Experimental Research*, vol. 14, no. 6, pp. 937–945, 1990.

**Dr.PanduRanga Vital Terlapu** , is a Professor at AITAM, Srikakulam in India, with 18 years of research experience. He holds a Bachelor of Science in Computer Science from Andhra University, a Master of Computer Application from Acharya Nagarjuna University, and an M.Tech in Computer Science and Engineering from GITAM University. He completed his PhD in Computer Science and Engineering in 1999. Dr. Terlapu has published over 59 research papers and filed three patents, including one granted by South Africa. His research interests include Machine Learning, Image Processing, Deep Learning, Data Mining, Big Data Analytics, IoT, Computational Intelligence, Voice Analysis, and Software Engineering.

**Dr. Ram Prasad Reddy Sadi** , is a seasoned and dedicated Engineering Professor at ANITS, Visakhapatnam. He holds a Ph.D. in Computer Science and Engineering, specializing in Machine Learning, and has published three books, 21 papers, 7 conference proceedings, and 4 patent grants. His research interests include Artificial Intelligence and Machine Learning. He is a life member of IEI and a senior member of ACM.