International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Review Article Open Access

Machine Learning-Based Acoustic Signal Processing for Bowl Sound Analysis

Ratul Ali[1], A.H.M. Saifullah Sadi[2*], Aktarul Islam[3], Md. Shohel Rana[4], Saila Nasrin[5] and Sohel Afzal Shajol[6]

[1]Department of Computer Science and Engineering, Uttara University (UU), Dhaka, Bangladesh. E-mail: 2161081002@uttarauniversity.edu.bd

[2]Professor, Department of Computer Science and Engineering, Uttara University (UU), Dhaka, Bangladesh. E-mail: saifullah.cse@uttarauniversity.edu.bd

[3]Department of Computer Science and Engineering, University of Rajshahi (RU), Rajshahi, Bangladesh. E-mail: aktarul857@gmail.com

[4]Department of Computer Science and Engineering, University of Rajshahi (RU), Rajshahi, Bangladesh. E-mail: msr.cse.ru@gmail.com

[5]Department of Computer Science and Engineering, Daffodil International University (DIU), Dhaka, Bangladesh. E-mail: sailanasrin92@gmail.com

[6]Department of Computer Science and Engineering, University of Development Alternative (UODA), Dhaka, Bangladesh. E-mail: sohelafzalshajol@gmail.com

Article Info

Volume 4, Issue 2, July 2024
Received : 09 January 2024
Accepted : 03 June 22024
Published : 05 July 2024
doi: 10.51483/IJAIML.4.2.2024.9-22

Abstract

Acoustic data plays a pivotal role in scientific and engineering research across various fields, including biology, communications, and Earth science. This study investigates recent advancements in acoustics, specifically focusing on machine learning (ML) and deep learning. ML, with its statistical techniques, autonomously identifies patterns in data. Unlike traditional acoustics, ML uncovers complex relationships among features and labels using extensive training data. Applying ML to acoustic phenomena like human speech and reverberation shows promising results. Additionally, this paper reviews acoustic signal processing for bowel sound analysis, emphasizing noise reduction, segmentation, feature extraction, and ML techniques. The integration of advanced signal processing and ML holds significant potential.

Keywords: *Acoustic data, Machine Learning, Signal processing, Bowel sound analysis, Artificial Intelligence*

1. Introduction

Acoustic data play a pivotal role in various scientific domains, including the interpretation of human speech and animal vocalizations, ocean source localization, and imaging geophysical structures in the ocean. Despite

* *Corresponding author: A.H.M. Saifullah Sadi, Professor, Department of Computer Science and Engineering Uttara University (UU), Dhaka, Bangladesh. E-mail: saifullah.cse@uttarauniversity.edu.bd*

the broad applications, challenges such as data corruption, missing measurements, reverberation, and large data volumes complicate the analysis. Machine learning (ML) techniques have emerged as a powerful solution to address these challenges, offering automated data processing and pattern recognition capabilities. ML in acoustics is a rapidly evolving field, with significant potential to overcome intricate acoustics challenges (Abaeikoupaei and Osman, 2023; Abrams *et al.*, 2008; Ackermann *et al.*, 2023; Akhtar *et al.*, 2023; Allen *et al.*, 1977; Allen and Berkley, 1979; Almeida *et al.*, 2019; Anagnostopoulos *et al.*, 2015; Anguera *et al.*, 2007; Bianco *et al.*, 2019).

ML, a family of techniques for detecting and utilizing patterns in data, proves beneficial in predicting future data or making decisions from uncertain measurements. It can be categorized into supervised and unsupervised learning, each serving distinct purposes. The historical focus in acoustics on high-level physical models is juxtaposed with the success of data-driven approaches facilitated by ML, indicating a shift towards hybrid models combining advanced acoustic models with ML (Breining *et al.*, 1999; Burgess and Granato, 2007; Carter and Bidelman, 2023; Caspary *et al.*, 1995; Chen *et al.*, 2014; Chibelushi *et al.*, 2002; Corcoran *et al.*, 2023).

In this dynamic landscape, ML in acoustics has witnessed remarkable progress, offering superior performance compared to traditional signal processing methods. However, challenges, such as the need for large datasets and the interpretability of ML models, persist. Despite these challenges, ML holds considerable potential in advancing acoustics research, as demonstrated. Including references, e.g. (Davis and Johnsrude, 2003; Denby *et al.*, 2010; Didier *et al.*, 2023; Dietzen *et al.*, 2023; Ding and Simon, 2013; Elliott and Theunissen, 2009; Ermilov *et al.*, 2009; Fang *et al.*, 2023).

The historical context of stethoscopes in medical practice, particularly in listening to the heart, lungs, and bowel sounds. Scientific analysis of bowel sounds dates back to the early 1900s, with observations and recordings dating even further. The sounds produced by the gastrointestinal tract offer valuable insights into the anatomy and physiology of the human gut, potentially revealing activities of the microbiome (Gabler *et al.*, 2023; Gajecki *et al.*, 2023; Gandour *et al.*, 2004; Gfeller *et al.*, 2007; Ghitza, 1994; Gillis *et al.*, 2023; Goli and Par, 2023).

The study further discusses the intersection of big data analytics and artificial intelligence in diverse applications, including bowel sound analysis. Artificial intelligence models, driven by advancements in

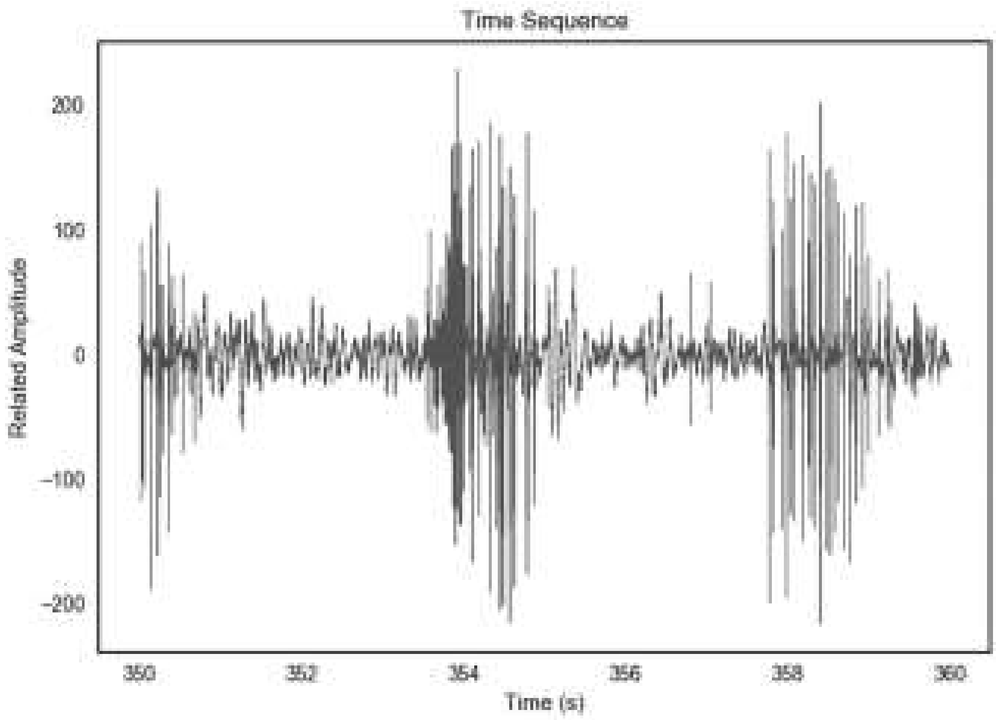

Figure 1: Time Domain Acoustic Signal Recorded from the Gut

computer processing power, have found utility in areas such as disease diagnosis and civil engineering. The application of these technologies to identify and analyze bowel sounds represents a notable advancement, offering a deeper understanding of gut functions and potential applications in healthcare (Hamsa *et al.*, 2023; Hansen, 1996; Hansen and Hasan, 2015; Hickok and Poeppel, 2007; Hollfelder *et al.*, 2023; Huang *et al.*, 2023).

The discussion concludes by highlighting improvements in acoustic signal processing methods, particularly in noise reduction and signal enhancement. Pioneering work in the 1970s utilized computers to analyze bowel sounds, marking the beginning of a journey that incorporated advanced signal processing techniques like Fourier transformation and short time Fourier transformation. These advancements culminated in the automatic detection of bowel sounds, showcasing the evolution of acoustic signal processing techniques in bowel sound applications (Figure 1).

2. Literature Review

2.1. Acoustic Signal Processing and Machine Learning Fundamentals

Machine Learning (ML) operates on a data-driven paradigm, capable of uncovering intricate relationships between features that conventional methods may overlook. While classic signal processing techniques rely on provable performance guarantees and simplifying assumptions, ML, particularly Deep Learning (DL), has demonstrated enhanced performance in various tasks. However, the increased flexibility of ML models introduces complexities, impacting both performance guarantees and model interpretability. ML models often necessitate substantial training data, though the requirement for 'vast' quantities is not mandatory to leverage ML techniques. Despite challenges, ML's benefits may outweigh the issues, especially when high performance is essential for a specific task (Johnson *et al.*, 2005; Jung *et al.*, 2020; Khoria *et al.*, 2023; Kong *et al.*, 2023; Krause and Braida, 2004).

Inputs and Outputs: In ML, the goal is often to train a model to produce a desired output (y) given inputs (x) (Figures 2 and 3). The supervised learning framework, represented by the equation

$$y = f(x) + \varepsilon$$

involves predicting outputs based on labeled input and output pairs. Here, x represents N features, y represents P desired outputs, $f(x)$ is the predicted output, and ε is the error. Training an ML model requires numerous examples, with X representing the inputs and Y representing the corresponding outputs. Supervised learning

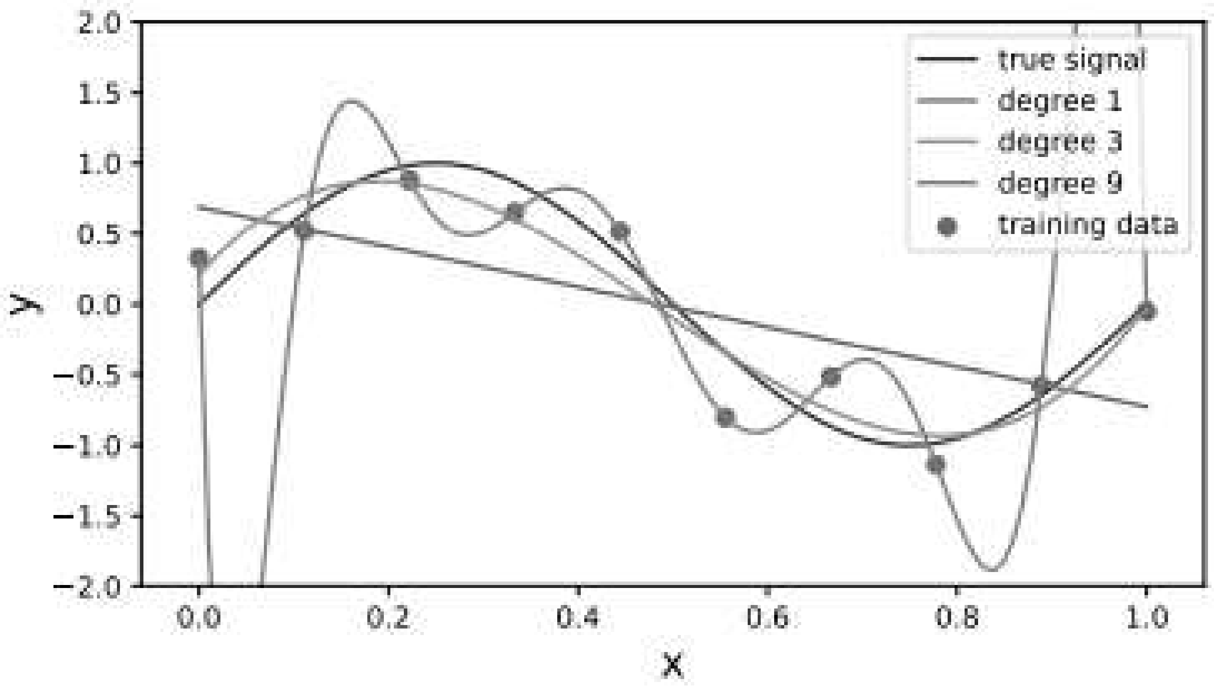

Figure 2: Model Generalization (Color Online)

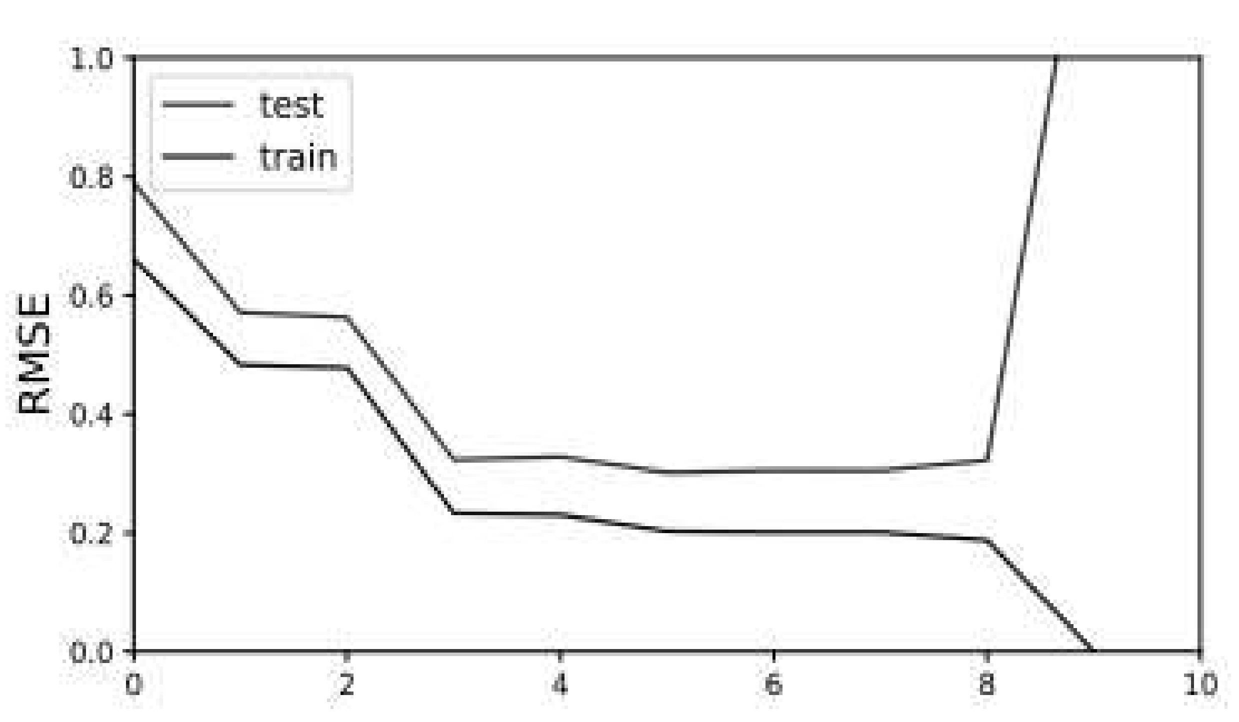

Figure 3: Model Generalization (Color Online)

focuses on predicting specific outputs, while unsupervised learning aims to discover patterns in data without explicit output specifications. Unsupervised learning often involves learning a model that approximates the features themselves (Krishna and Semple, 2000; Langner, 1992; Lee and Narayanan, 2005; Lenk *et al.*, 2023; Little *et al.*, 2007).

2.2. Signal Identification and Enhancement

Sounds result from mechanical deformation, generating energy waves detected by the ear or transducer. Acoustic signal processing and ML techniques contribute to understanding these phenomena (Liu and Vicario, 2023; Luthra, 2023; Magnuson and Nusbaum, 2007; Markovich *et al.*, 2009; Martin and Boothroyd, 1999).

1) **Time Domain Signal:** The raw data, a time domain signal, is crucial for acoustic analysis. Features like SNR, duration, and event count are extracted, aiding in signal quality assessment. Filtering methods, including adaptive filtering, enhance signals by removing unwanted components.

2) **Frequency Domain Signal:** Transforming signals into the frequency domain through Fourier analysis reveals information unobservable in the time domain. The FFT technique provides features like centroid frequency and spectral bandwidth, but may lose some time domain information.

3) **Time-Frequency Domain Signal:** Simultaneous time and frequency information is obtained using Short-Time Fourier Transform (STFT) or Wavelet Transform (WT). Spectrograms from STFT enable speech recognition and noise suppression. WT, known for noise suppression, offers varied time and frequency domain information.

3. Advanced Signal Processing

3.1. Supervised Learning and Linear Regression in the Context of Acoustic Signal Processing

Supervised learning, a fundamental aspect of machine learning (ML), aims to establish a mapping from a set of inputs to desired outputs through labeled input-output pairs. In this discussion, we focus on real-valued features and labels, where the N features in x can be real, complex, or categorical. The corresponding supervised learning tasks are divided into two subcategories: regression and classification. Regression addresses scenarios where y is real or complex valued, while classification pertains to cases with categorical y.

The central focus in ML methods lies in finding the function f, particularly using probability tools when practical. The supervised ML task can be articulated as maximizing the conditional distribution $p(y \mid x)$, with the Maximum A Posteriori (MAP) estimator providing a point estimate for y, denoted as $y^b = f(x)$.

Linear regression serves as an illustrative example of supervised ML. In the context of Direction of Arrival (DOA) estimation in beamforming for seismic and acoustic applications, we represent the relationship between the observed Fourier-transformed measurements x and the DOA azimuth angle y using a linear measurement model. The optimization problem seeks values of weights w that minimize the difference between the observed and predicted measurements, effectively solving the linear regression problem.

The ensuing Bayesian treatment involves formulating the posterior of the model using Bayes' rule, leading to a MAP estimate for the weights. Depending on the choice of the probability density function for the weights, solutions may vary. A popular choice, the Gaussian distribution, results in the classic L^2-regularized least squares estimate, incorporating a regularization parameter for stability.

This detailed exposition highlights the foundational principles of supervised learning and its application in linear regression within the specific domain of acoustics, illustrating the seamless integration of theoretical ML concepts with practical signal processing challenges (Merchant *et al.*, 2015; Mesgarani *et al.*, 2014; Meyer, 2018; Minelli *et al.*, 2023).

1) **Advanced Signal Processing in Bowel Sound Analysis:** Acoustic signal processing in the context of bowel sound analysis involves a multi-step sequence encompassing data acquisition, preprocessing, and subsequent analysis. The reviewed literature reveals a diverse array of approaches and methodologies, with certain commonalities in the overall processing flow.

2) **Data Acquisition:** To record abdominal sounds, specialized transducers, such as electret condenser microphones or piezoelectric transducers, are designed to convert acoustic energy into electrical signals. Electronic stethoscopes, including designs like the JABES digital stethoscope and 3M Littmann 3200, demonstrate the versatility of these transducers. Additionally, innovative approaches, such as 3D-printed stethoscope heads with built-in electronics, reflect evolving design paradigms.

3) **Preprocessing and Analysis:** The preprocessing stage involves denoising, filtering, and segmentation of acoustic signals, often employing techniques like adaptive filtering and enveloping. The choice of window functions, such as rectangular, Hamming, and Hann, plays a crucial role in the slicing of acoustic recordings into small samples.

4) **Bowel Sound Analysis:** From the early 2000s, wavelet transforms (WTs) have enabled advanced feature extraction, coinciding with the integration of machine learning methods. Researchers, exemplified by groups led by Hadjileontiadis et al., have made substantial progress in noise reduction and signal enhancement for bowel sounds. Various machine learning methods, including decision trees, dimension reduction, and artificial neural networks, have been applied to characterize bowel sounds (Nagarajan *et al.*, 2023; Peelle and Wingfield, 2016; Poeppel, 2001).

In acoustics, the Fourier Transform is often used to analyze the frequency components of a signal. The Fourier Transform of a function $f(t)$ is defined as:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad ...(1)$$

where $F(\omega)$ is the Fourier Transform of $f(t)$, and ω is the angular frequency.

Let's strudy a sound signal $f(t)$ given by:

$$f(t) = A\sin(2\pi f_0 t) \quad ...(2)$$

where A is the amplitude and f_0 is the frequency of the sound.

The Fourier Transform of $f(t)$ is then calculated as:

$$F(\omega) = \int_{-\infty}^{\infty} A\sin(2\pi f_0 t) e^{-i\omega t} dt \quad ...(3)$$

This integral can be solved to find the expression for $F(\omega)$.

The literature review underscores the dynamic landscape of acoustic signal processing in bowel sound analysis, with researchers adopting diverse approaches across the processing stages. From innovative data acquisition methods to sophisticated preprocessing techniques and the application of machine learning, the field demonstrates a blend of traditional signal processing principles and contemporary methodologies. The convergence of theoretical insights and practical implementations serves as a foundation for continued advancements in acoustic signal processing for bowel sound analysis (Poeppel, 2001; Poluboina *et al.*, 2023; Randall, 2017; Ravanelli *et al.*, 2018).

3.2. Parallelization of All-Pairs Algorithm (OpenMP)

The provided algorithm outlines an approach to acoustic signal processing with parallelization using OpenMP (Algorithm 1).

1) Main Function: acousticSignalProcessing():

- This function serves as the entry point for the acoustic signal processing algorithm.
- It is marked for parallelization using the #pragma omp parallel for directive, which instructs the compiler to parallelize the loop that iterates over the model collection. For each model in the collection, the function calls processModel(i, signal).

2) Processing Each Model: processModel(i: model, signal):

- This function is also marked for parallelization using the #pragma omp parallel for reduction (+ : result[i].amplitude) directive.
- It contains a nested loop that iterates over the signal collection for each model. For each pair of models (i, j), where j is not equal to i, it calculates the similarity between the models using the calculateSimilarity(i, j) function.
- The amplitude of the result for the current model (result[i].amplitude) is adjusted based on the calculated similarity using the adjustAmplitude(i, j, similarity) function.

```
Function acousticSignalProcessing() is
    #pragma omp parallel for
    foreach i: model do
        processModel(i, signal)

Function processModel(i: model, signal) is
    #pragma omp parallel for reduction (+ :
      result[i].amplitude)
    foreach j in signal do
        if j ≠ i then
            similarity = calculateSimilarity(i, j)
            result[i].amplitude += adjustAmplitude(i,
              j, similarity)

Function calculateSimilarity(i, j) is
    // Calculations of the similarities
       between models i and j
    return similarity

Function adjustAmplitude(i, j, similarity) is
    // Adjustment of the amplitude based on
       the similarities
    return adjustedAmplitude
```

Algorithm 1: Acoustic Signal Processing Algorithm (OpenMP)

3) Calculating Similarity: calculateSimilarity(i, j):
   - The specific details of how the similarity is calculated are not provided in the algorithm and should be implemented according to the requirements of the acoustic signal processing application.
   - This function is a placeholder for calculating the similarity between two models, i and j.
4) Adjusting Amplitude: adjustAmplitude(i, j, similarity):
   - This function is a placeholder for adjusting the amplitude of a model based on the calculated similarity.
   - Again, the exact method of adjusting the amplitude is not specified and needs to be implemented based on the application's requirements.

3.3. Parallelization of All-Pairs Algorithm (CUDA)

Sequential Barnes-Hut Algorithm with Acoustic Signal Processing

1) Main Function: acousticBarnesHut():
   - This function represents the entry point for the integrated algorithm, combining he Sequential Barnes-Hut structure with acoustic signal processing.
   - It orchestrates the sequential execution of three main steps: building the tree (build_tree()), computing mass distribution (compute_mass_distribution()), and calculating forces (compute_force()).
2) Building the Tree: build_tree():
   - The function initializes the tree structure, preparing it for the insertion of acoustic models.
   - It iterates over each acoustic model in the dataset and inserts it into the root node using the insert_to_node() function.
3) Inserting Models into Nodes: insert_to_node(new_model):
   - This function is responsible for placing a new acoustic model into the appropriate quadrant of the Barnes-Hut tree.
   - It checks the number of existing models in a node. If there is more than one model, it recursively traverses the tree to find the appropriate quadrant for the new model. If there's only one model, it divides the node into quadrants, placing the existing and new models accordingly.
   - If no models exist in the node, it directly assigns the new model as the existing model.
4) Computing Mass Distribution: compute_mass_distribution():
   - This function calculates the mass distribution within each quadrant of the Barnes-Hut tree.
   - If there is only one model in a quadrant, the center of mass and mass are directly assigned from that model. Otherwise, it recursively calculates the mass distribution for child quadrants, aggregating the mass and weighted center of mass.
5) Calculating Forces: calculate_force(target):
   - This function computes the acoustic forces acting on a target model.
   - If there's only one model in the quadrant, the force is calculated using the acoustic_force() function between the target and the model. If the quadrant size is below a certain threshold (lD < theta), the force is computed using the acoustic force model.
   - If the quadrant is larger, the algorithm recursively calculates forces for child nodes and aggregates them.
6) Computing Forces for all Models: compute_force():
   - This function iterates over all acoustic models in the dataset and computes the forces acting on each model using the root_node.calculate_force(model) function.

- If there's only one model in the quadrant, the force is calculated using the acoustic_force() function between the target and the model. If the quadrant size is below a certain threshold (lD < theta), the force is computed using the acoustic force model.
- If the quadrant is larger, the algorithm recursively calculates forces for child nodes and aggregates them.

3.4. Sequential Barnes-Hut Algorithm

It represents the integrated algorithm with the Sequential Barnes-Hut structure and Acoustic Signal Processing. The algorithm includes functions for building the tree, inserting models into nodes, computing mass distribution, calculating forces, and overall coordination of the acoustic signal processing with the Barnes-Hut algorithm.

The integrated algorithm merges the Sequential Barnes-Hut structure, designed for efficient gravitational force calculations, with acoustic signal processing. The Barnes-Hut tree structure optimizes the computation of forces between acoustic models, enhancing the algorithm's scalability and efficiency in handling large datasets. The acoustic signal processing steps involve building the tree, distributing mass, and calculating forces, offering a comprehensive solution for analyzing and simulating acoustic interactions within a given system.

```
Function acousticBarnesHut() is
    build_tree()
    compute_mass_distribution()
    compute_force()

Function build_tree() is
    Reset Tree
    foreach i: model do
        root_node→insert_to_node(i)

Function insert_to_node(new_model) is
    if num_models > 1 then
        quad = get_quadrant(new_model)
        if subnode(quad) does not exist then
            create subnode(quad)
        subnode(quad)→insert_to_node(new_model)
    else if num_models == 1 then
        quad = get_quadrant(new_model)
        if subnode(quad) does not exist then
            create subnode(quad)
        subnode(quad)→insert_to_node(existing_model)

        quad = get_quadrant(new_model)
        if subnode(quad) != NULL then
            create subnode(quad)
        subnode(quad)→insert_to_node(new_model)

    else
        existing_model ← new_model
    num_models++
```

Algorithm 2: Algorithm Part 1

```
Function compute_mass_distribution() is
    if num_models == 1 then
        center_of_mass = model.position
        mass = model.mass
    else
        forall child quadrants with models do
            quadrant.compute_mass_distribution
            mass += quadrant.mass
            center_of_mass += quadrant.mass ×
              quadrant.center_of_mass
        center_of_mass /= mass
```

Algorithm 3: Algorithm Part 2

```
Function calculate_force(target) is
    Initialize force ← 0
    if num_models == 1 then
        force = acoustic_force(target, model)
    else
        if l/D < θ then
            force = acoustic_force(target, model)
        else
            forall node : child nodes do
                force += node.calculate_force(node)

Function compute_force() is
    forall models do
        force = root_node.calculate_force(model)
```

Algorithm 4: Algorithm Part 3

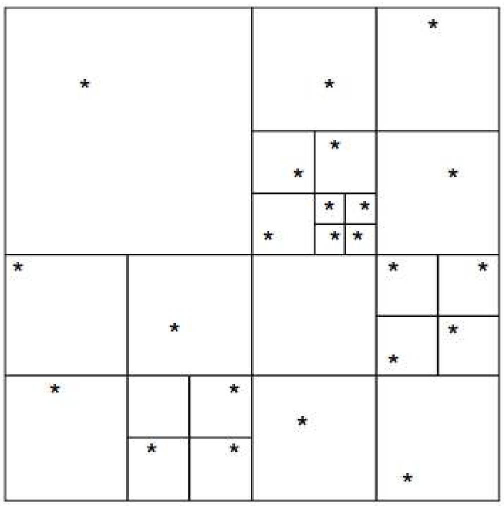

Figure 4: Barnes-Hut Tree Structure

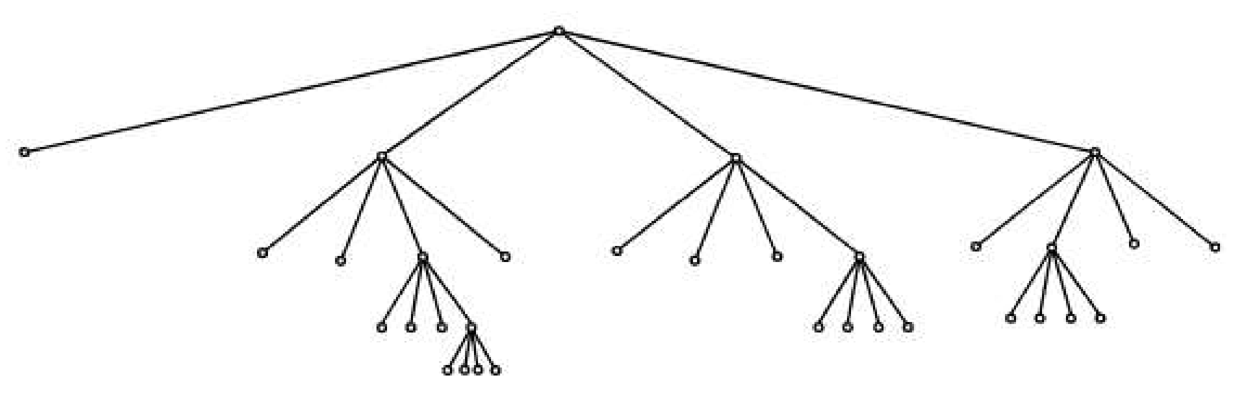

Figure 5: Barnes-Hut Domain Decomposition

4. Conclusion

In this comprehensive review, we have presented an overview of Machine Learning (ML) theory, with a particular focus on deep learning (DL), and explored its diverse applications across various acoustics research domains. While our coverage is not exhaustive, it is evident that ML has been a catalyst for numerous recent advancements in acoustics. This article aims to inspire future ML research in acoustics, emphasizing the pivotal role of large, publicly available datasets in fostering innovation across the acoustics field. The transformative potential of ML in acoustics is substantial, with its benefits amplified through open data practices (Sainath *et al.*, 2017; Schonwiesner *et al.*, 2005; Souden *et al.*, 2010; Stephen *et al.*, 2023; Stevens, 2002).

Despite the acknowledged limitations of ML-based methods, their performance surpasses that of conventional processing methods in many scenarios. However, it is crucial to recognize that ML models, being datadriven, demand substantial representative training data for optimal performance. This is viewed as a trade-off for accurately modeling complex phenomena, given the often high capacity of ML models. In contrast, standard processing methods, with lower capacity, rely on training-free statistical and mathematical models (Stowell *et al.*, 2015; Tandon and Choudhury, 1999; Telkemeyer *et al.*, 2009; Tezcan *et al.*, 2023).

This review suggests a paradigm shift in acoustic processing from hand-engineered, intuition-driven models to a data-driven ML approach. While harnessing the full potential of ML, it is essential to build upon indispensable physical intuition and theoretical developments within established sub-fields like array processing. The development of ML theory in acoustics should be undertaken while preserving the foundational physical principles that describe our environments. By blending ML advancements with established principles, transformative progress can be achieved across various acoustics fields (Ufer and Blank, 2023; Viola and Walker, 2005).

Upon on bowel sound analysis, several conclusions emerge. The choice of sensors for data acquisition, including electret condenser microphones and piezoelectric transducers, depends on research constraints. Advanced signal processing techniques, such as wavelet transforms (WTs) since the early 2000s, have enabled complex feature extraction. Machine learning methods have found application in bowel sound analysis, with varying approaches such as decision trees, dimension reduction, and clustering algorithms (Voola *et al.*, 2023; Wakita, 1973; Wu *et al.*, 2003; Xu *et al.*, 2002; Xu *et al.*, 2023; Yang *et al.*, 1992; Zmolikova *et al.*, 2023).

Conflict of Interest

The authors declare that they have no conflict of interest.

Funding

No Funding

References

Abaeikoupaei, N. and Osman, H.A. (2023). A Multi-Modal Stacked Ensemble Model for Bipolar Disorder Classification. *IEEE Transactions on Affective Computing*, 14.

Abrams, D.A., Nicol, T., Zecker, S. and Kraus, N. (2008). Right-Hemisphere Auditory Cortex is Dominant for Coding Syllable Patterns in Speech. *Journal of Neuroscience*, 28.

Ackermann, D., Domann, J., Brinkmann, F., Arend, J.M., Schneider, M., Pörschmann, C. and Weinzierl, S. (2023). Recordings of a Loudspeaker Orchestra with Multichannel Microphone Arrays for the Evaluation of Spatial Audio Methods. *AES: Journal of the Audio Engineering Society*, 71.

Akhtar, M.T., Albu, F. and Nishihara, A. (2023). Prediction Error Method (Pem)-Based Howling Cancellation in Hearing Aids: Can We Do Better?. *IEEE Access*, 11.

Allen J.B. and Berkley, D.A. (1979). Image Method for Efficiently Simulating Small-Room Acoustics. *Journal of the Acoustical Society of America*, 65.

Allen, J.B., Berkley, D.A. and Blauert, J. (1977). Multimicrophone Signal-Processing Technique to Remove Room Reverberation from Speech Signals. *Journal of the Acoustical Society of America*, 62.

Almeida, J.S., Filho, P.P.R., Carneiro, T., Wei, W., Damaševicius, R., Maskeliunas, R. and de Albuquerque, V.H.C. (2019). Detecting Parkinson's Disease with Sustained Phonation and Speech Signals Using Machine Learning Techniques. *Pattern Recognition Letters*, 125.

Anagnostopoulos, C.N., Iliou, T. and Giannoukos, I. (2015). Features and Classifiers for Emotion Recognition from Speech: A Survey from 2000 to 2011. *Artificial Intelligence Review*, 43.

Anguera, X., Wooters, C. and Hernando, J. (2007). Acoustic Beamforming for Speaker Diarization of Meetings. *IEEE Transactions on Audio, Speech and Language Processing*, 15.

Bianco, M.J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M.A., Gannot, S. and Deledalle, C.-A. (2019). Machine Learning in Acoustics: Theory and Applications. *The Journal of the Acoustical Society of America*, 146.

Breining, C., Dreiseitel, P., Hänsler, E., Mader, A., Nitsch, B., Puder, H., Schertler, T., Schmidt, G. and Tilp, J. (1999), Acoustic Echo Control, an Application of Very-High-Order Adaptive Filters. *IEEE Signal Processing Magazine*, 16.

Burgess, H.A. and Granato, M. (2007). Sensorimotor Gating in Larval Zebrafish. *Journal of Neuroscience*, 27.

Carter, J.A. and Bidelman, G.M. (2023). Perceptual Warping Exposes Categorical Representations for Speech in Human Brainstem Responses. *NeuroImage*, 269.

Caspary, D.M., Milbrandt, J.C. and Helfert, R.H. (1995). Central Auditory Aging: Gaba Changes in the Inferior Colliculus. *Experimental Gerontology*, 30.

Chen, J., Wang, Y. and Wang, D. (2014). A Feature Study for Classification-Based Speech Separation at Low Signal-to-Noise Ratios. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 22.

Chibelushi, C.C., Deravi, F. and Mason, J.S. (2002). A Review of Speech-Based Bimodal Recognition. *IEEE Transactions on Multimedia*, 4.

Corcoran, A.W., Perera, R., Koroma, M., Kouider, S., Hohwy, J. and Andrillon, T. (2023). Expectations Boost the Reconstruction of Auditory Features from Electrophysiological Responses to Noisy Speech. *Cerebral Cortex, (New York, N.Y.: 1991)*, 33.

Davis, M.H. and Johnsrude, I.S. (2003). Hierarchical Processing in Spoken Language Comprehension. *Journal of Neuroscience*, 23.

Denby, B., Schultz, T., Honda, K., Hueber, T., Gilbert, J.M. and Brumberg, J.S. (2010). Silent Speech Interfaces. *Speech Communication*, 52.

Didier, P., Waterschoot, T.V., Doclo, S. and Moonen, M. (2023). Sampling Rate Offset Estimation and Compensation for Distributed Adaptive Node-Specific Signal Estimation in Wireless Acoustic Sensor Networks. *IEEE Open Journal of Signal Processing*, 4.

Dietzen, T., Ali, R., Taseska, M. and van Waterschoot, T. (2023). Myriad: A Multi-Array Room Acoustic Database. *Eurasip Journal on Audio, Speech, and Music Processing*, 2023.

Ding, N. and Simon, J.Z. (2013). Adaptive Temporal Encoding Leads to a Background-Insensitive Cortical Representation of Speech. *Journal of Neuroscience*, 33.

Elliott, T.M. and Theunissen, F.E. (2009). The Modulation Transfer Function for Speech Intelligibility. *PLoS Computational Biology*, 5.

Ermilov, S.A., Khamapirad, T., Conjusteau, A., Leonard, M.H., Lacewell, R., Mehta, K., Miller, T. and Oraevsky, A.A. (2009). Laser Optoacoustic Imaging System for Detection of Breast Cancer. *Journal of Biomedical Optics*, 14.

Fang, J., Li, Y., Ji, P.N. and Wang, T. (2023). Drone Detection and Localization Using Enhanced Fiber-Optic Acoustic Sensor and Distributed Acoustic Sensing Technology. *Journal of Lightwave Technology*, 41.

Gabler, P., Geiger, B.C., Schuppler, B. and Kern, R. (2023). Reconsidering Read and Spontaneous Speech: Causal Perspectives on the Generation of Training Data for Automatic Speech Recognition.

Gajecki, T., Zhang, Y. and Nogueira, W. (2023). A Deep Denoising Sound Coding Strategy for Cochlear Implants. *IEEE Transactions on Biomedical Engineering*, 70.

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., Li, X. and Lowe, M. (2004). Hemispheric Roles in the Perception of Speech Prosody. *NeuroImage*, 23.

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R. and Olszewski, C. (2007). Accuracy of Cochlear Implant Recipients on Pitch Perception, Melody Recognition, and Speech Reception in Noise. *Ear and Hearing*, 28.

Ghitza, O. (1994). Auditory Models and Human Performance in Tasks Related to Speech Coding and Speech Recognition.

Gillis, M., Kries, J., Vandermosten, M. and Francart, T. (2023). Neural Tracking of Linguistic and Acoustic Speech Representations Decreases with Advancing Age. *NeuroImage*, 267.

Goli, P. and Par, S.V.D. (2023). Deep Learning-Based Speech Specific Source Localization by Using Binaural and Monaural Microphone Arrays in Hearing Aids. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 31.

Hamsa, S., Shahin, I., Iraqi, Y., Damiani, E., Nassif, A.B. and Werghi, N. (2023). Speaker Identification from Emotional and Noisy Speech Using Learned Voice Segregation and Speech VGG. *Expert Systems with Applications*, 224.

Hansen, J.H. (1996). Analysis and Compensation of Speech Under Stress and Noise for Environmental Robustness in Speech Recognition. *Speech Communication*, 20.

Hansen, J.H. and Hasan, T. (2015). Speaker Recognition by Machines and Humans: A Tutorial Review.

Hickok, G. and Poeppel, D. (2007). The Cortical Organization of Speech Processing.

Hollfelder, D., Prein, L., Jürgens, T., Leichtle, A. and Bruchhage, K.L. (2023). Influence of Directional Microphones on Listening Effort in Middle Ear Implant Users. *HNO*, 71.

Huang, Y., Ma, Y., Xiao, J., Liu, W. and Zhang, G. (2023). Identification of Depression State Based on Multi-Scale Acoustic Features in Interrogation Environment. *IET Signal Processing*, 17.

Johnson, K.L., Nicol, T.G. and Kraus, N. (2005). Brain Stem Response to Speech: A Biological Marker of Auditory Processing.

Jung, Y.H., Hong, S.K., Wang, H.S., Han, J.H., Pham, T.X., Park, H., Kim, J., Kang, S., Yoo, C.D. and Lee, K.J. (2020). Flexible Piezoelectric Acoustic Sensors and Machine Learning for Speech Processing.

Khoria, K., Patil, A.T. and Patil, H.A. (2023). On Significance of Constant-q Transform for Pop Noise Detection. *Computer Speech and Language*, 77.

Kong, F., Zhou, H., Mo, Y., Shi, M., Meng, Q. and Zheng, N. (2023). Comparable Encoding, Comparable Perceptual Pattern: Acoustic and Electric Hearing. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31.

Krause, J.C., and Braida, L.D. (2004). Acoustic Properties of Naturally Produced Clear Speech at Normal Speaking Rates. *The Journal of the Acoustical Society of America*, 115.

Krishna, B.S. and Semple, M.N. (2000). Auditory Temporal Processing: Responses to Sinusoidally Amplitude-Modulated Tones in the Inferior Colliculus. *Journal of Neurophysiology*, 84.

Langner, G. (1992). Periodicity Coding in the Auditory System.

Lee, C.M. and Narayanan, S.S. (2005). Toward Detecting Emotions in Spoken Dialogs. *IEEE Transactions on Speech and Audio Processing*, 13.

Lenk, C., Hövel, P., Ved, K., Durstewitz, S., Meurer, T., Fritsch, T., Männchen, A., Küller, J., Beer, D., Ivanov, T. and Ziegler, M. (2023). Neuro-Morphic Acoustic Sensing Using an Adaptive Microelectromechanical Cochlea with Integrated Feedback. *Nature Electronics*, 6.

Little, M.A., McSharry, P.E., Roberts, S.J., Costello, D.A. and Moroz, I.M. (2007). Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection. *Bio-Medical Engineering Online*, 6.

Liu, W. and Vicario, D.S. (2023). Dynamic Encoding of Phonetic Categories in Zebra Finch Auditory Forebrain. *Scientific Reports*, 13.

Luthra, S. (2023). Why are Listeners Hindered by Talker Variability?.

Magnuson, J.S. and Nusbaum, H.C. (2007). Acoustic Differences, Listener Expectations, and the Perceptual Accommodation of Talker Variability. *Journal of Experimental Psychology: Human Perception and Performance*, 33.

Markovich, S., Gannot, S. and Cohen, I. (2009). Multichannel Eigenspace Beamforming in a Reverberant Noisy Environment with Multiple Interfering Speech Signals. *IEEE Transactions on Audio, Speech and Language Processing*, 17.

Martin, B.A. and Boothroyd, A. (!999). Cortical, Auditory, Event-Related Potentials in Response to Periodic and Aperiodic Stimuli with the Same Spectral Envelope. *Ear and Hearing*, 20.

Merchant, N.D., Fristrup, K.M., Johnson, M.P., Tyack, P.L., Witt, M.J., Blondel, P. and Parks, S.E. (2015). Measuring Acoustic Habitats. *Methods in Ecology and Evolution*, 6.

Mesgarani, N., Cheung, C., Johnson, K. and Chang, E.F. (2014). Phonetic Feature Encoding in Human Superior Temporal Gyrus. *Science*, 343.

Meyer, L. (2018). The Neural Oscillations of Speech Processing and Language Comprehension: State of the Art and Emerging Mechanisms.

Minelli, G., Puglisi, G.E., Astolfi, A., Hauth, C. and Warzybok, A. (2023). Objective Assessment of Binaural Benefit from Acoustical Treatment in Real Primary School Classrooms. *International Journal of Environmental Research and Public Health*, 20.

Nagarajan, D., Broumi, S. and Smarandache, F. (2023). Neutrosophic Speech Recognition Algorithm for Speech Under Stress by Machine Learning. *Neutrosophic Sets and Systems*, 55.

Peelle, J.E. (2018). Listening Effort: How the Cognitive Consequences of Acoustic Challenge are Reflected in Brain and Behavior. *Ear and Hearing*, 39.

Peelle, J.E. and Wingfield, A. (2016). The Neural Consequences of Age-Related Hearing Loss.

Poeppel, D. (2001). Pure Word Deafness and the Bilateral Processing of the Speech Code. *Cognitive Science*, 25.

Poluboina, V., Pulikala, A. and Muthu, A.N.P. (2023). An Improved Noise Reduction Technique for Enhancing the Intelligibility of Sinewave Vocoded Speech: Implication in Cochlear Implants. *IEEE Access*, 11.

Randall, R.B. (2017). A History of Cepstrum Analysis and its Application to Mechanical Problems. *Mechanical Systems and Signal Processing*, 97.

Ravanelli, M., Brakel, P., Omologo, M. and Bengio, Y. (2018). Light Gated Recurrent Units for Speech Recognition. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2.

Sainath, T.N., Weiss, R.J., Wilson, K.W., Li, B., Narayanan, A., Variani, E., Bacchiani, M., Shafran, I., Senior, A., Chin, K., Misra, A. and Kim, C. (2017). Multichannel Signal Processing with Deep Neural Networks for Automatic Speech Recognition. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 25.

Schonwiesner, M., Rübsamen, R. and Cramon, D.Y.V. (2005). Hemispheric Asymmetry for Spectral and Temporal Processing in the Human Antero-Lateral Auditory Belt Cortex. *European Journal of Neuro-Science*, 22.

Souden, M., Benesty, J. and Affes, S. (2010). On Optimal Frequency-Domain Multichannel Linear Filtering for Noise Reduction. *IEEE Transactions on Audio, Speech and Language Processing*, 18.

Stephen, E.P., Li, Y., Metzger, S., Oganian, Y. and Chang, E.F. (2023). Latent Neural Dynamics Encode Temporal Context in Speech.

Stevens, K.N. (2002). Toward a Model for Lexical Access Based on Acoustic Landmarks and Distinctive Features. *The Journal of the Acoustical Society of America*, 111.

Stowell, D., Giannoulis, D., Benetos, E., Lagrange, M. and Plumbley, M.D. (2015). Detection and Classification of Acoustic Scenes and Events. *IEEE Transactions on Multimedia*, 17.

Tandon, N. and Choudhury, A. (1999). Review of Vibration and Acoustic Measurement methods for the Detection of Defects in Rolling Element Bearings. *Tribology International*, 32.

Telkemeyer, S., Rossi, S., Koch, S.P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H. and Wartenburger, I. (2009). Sensitivity of Newborn Auditory Cortex to the Temporal Structure of Sounds. *Journal of Neuroscience*, 29.

Tezcan, F., Weissbart, H. and Martin, A.E. (2023). A Tradeoff Between Acoustic and Linguistic Feature Encoding in Spoken Language Comprehension. *eLife*, 12.

Ufer, C. and Blank, H. (2023). Multivariate Analysis of Brain Activity Patterns as a Tool to Understand Predictive Processes in Speech Perception. *Language, Cognition and Neuroscience*.

Viola, F. and Walker, W.F. (2005). A Spline-Based Algorithm for Continuous Time-Delay Estimation Using Sampled Data. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 52.

Voola, M., Nguyen, A.T., Wedekind, A., Marinovic, W., Rajan, G. and Tavora-Vieira, D. (2023). A Study of Event-Related Potentials During Monaural and Bilateral Hearing in Single-Sided Deaf Cochlear Implant Users. *Ear and Hearing*, 44.

Wakita, H. (1973). Direct Estimation of the Vocal Tract Shape by Inverse Filtering of Acoustic Speech Waveforms. *IEEE Transactions on Audio and Electroacoustics*, 21.

Wu, M., Wang, D.L. and Brown, G.J. (2003). A Multipitch Tracking Algorithm for Noisy Speech. *IEEE Transactions on Speech and Audio Processing*, 11.

Xu, L., Tsai, Y. and Pfingst, B.E. (2002). Features of Stimulation Affecting Tonal-Speech Perception: Implications for Cochlear Prostheses. *The Journal of the Acoustical Society of America*, 112.

Xu, R., Sun, J., Wang, Y., Zhang, S., Zhong, W. and Wang, Z. (2023). Speech Enhancement Based on Array-Processing-Assisted Distributed Fiber Acoustic Sensing. *IEEE Sensors Journal*, 23.

Yang, X., Wang, K. and Shamma, S.A. (1992). Auditory Representations of Acoustic Signals. *IEEE Transactions on Information Theory*, 38.

Zmolikova, K., Delcroix, M., Ochiai, T., Kinoshita, K., Cernocky, J. and Yu, D. (2023). Neural Target Speech Extraction: An Overview. *IEEE Signal Processing Magazine*, 40.

Cite this article as: Ratul Ali, A.H.M. Saifullah Sadi, Aktarul Islam, Md. Shohel Rana, Saila Nasrin and Sohel Afzal Shajol (2024). Machine Learning-Based Acoustic Signal Processing for Bowl Sound Analysis. *International Journal of Artificial Intelligence and Machine Learning*, 4(2), 9-22. doi: 10.51483/IJAIML.4.2.2024.9-22.