Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

Volume 09 Issue 01
2025

Review Article
https://doi.org/10.56979/901/2025

Systematic Literature Review on Application of Naive Bayes Algorithm for Large Audio Data Classification

Rabia Tehseen[1,7*], Shazia Saqib[2,6], Uzma Omer[3,7], Anam Mustaqeem[4], Maham Mehr[4], and Shahan Yamin Siddiqui[5]

[1]Department of Computer Science, University of Central Punjab, Lahore, Pakistan.
[2]Department of Computer Science, Institute of Art and Culture, Lahore, Pakistan.
[3]Department of Computer Science, University of Education, Lahore, Pakistan.
[4]Department of Software Engineering, University of Central Punjab, Lahore, Pakistan.
[5]Department of Computing, NASTP Institute of Information Technology, Lahore, Pakistan.
[6]INTI International University, Malaysia.
[7]Department of Computer Science, University of Management & Technology, Lahore, Pakistan.
[*]Corresponding Author: Rabia Tehseen. Email: rabia.tehseen@ucp.edu.pk

Received: April 02, 2025 **Accepted:** May 18, 2025

Abstract: The increasing volume of audio data in areas like speech recognition, music genre classification, and environment sound analysis has created a need for more effective and scalable classification algorithms. This systematic literature review focuses on the use of the Naive Bayes algorithm for large-scale audio data classification and evaluates 39 peer-reviewed articles published in the last five years. The review analyses how Naive Bayes has been applied to difficulties such as feature extraction, model training, and real-time classification of audio signals considering its simplicity and computational efficiency. We assess its performance against more sophisticated machine learning techniques and its flexibility with pre-processing, ensemble models, and cross-layer control algorithms. Findings demonstrate that although Naive Bayes does not always outperform deep learning algorithms, this remains a strong option when low latency, explainability, and minimal resources are required. The review also points out gaps in existing research and discusses potential approaches to improve the algorithm's performance in audio-based tasks.

Keywords: Naïve Bayes; Audio Classification; Review; Audio-Based Applications

1. Introduction

The application of Naive Bayesian classifiers has become widespread in different fields because of their ease and effectiveness. [1] Used a Naive Bayes classifier to categorize application in the Google Play Store and achieved great results in automating app classification [1]. A similar example is provided by [2] who added a negation scope and n-grams into a sentiment analysis Naive Bayes model and improved its performance. These studies point out the efficacy of Naive Bayes classifiers for different classification tasks.

[3] Made attempts in the area of sound classification and recognition using different algorithms that would accurately attend to the tasks and recognize sound categories. Jeng et al. applied image classification techniques to audio classification by treating sound spectrograms as images with texture. [4] Studied the use of statistical methods and speech neural network models in identifying speech and music and concluded that neural networks performed better in more complicated situations. These studies demonstrate the steps taken towards better audio classification by applying modern techniques.

This study aims to develop a lightweight and scalable audio classification system capable of handling large datasets. While deep learning models offer high accuracy, they are computationally intensive and demand substantial training data. In contrast, Naive Bayes classifiers are computationally efficient but often dismissed as too simplistic for complex tasks like audio classification. This work challenges that assumption by proposing a refined Naive Bayes-based framework that incorporates tailored feature extraction and probabilistic modeling strategies, enhancing its suitability for large-scale audio classification.

This study is unique and significant because of the attempts made to utilize the efficiency and simplicity of Naive Bayes classifiers for large audio data sets. This study builds a comprehensive audio classification framework by using new methods of feature extraction and addressing the challenges posed by the diversity and complexity of audio data. The proposed approach improves the accuracy of Naive Bayes classifiers in this area and suggests a new way of accomplishing audio classification without the need for sophisticated models.

This systematic literature review has been organized into multiple sections. Section 2 presents methodology describing objectives, research questions, selection criteria and classification of articles included in this research. Results have been discussed in chapter 3. Conclusion and future directions have been provided in section 4.

2. Methodology

This systematic literature review (SLR) was developed to investigate works on Naïve Bayes and audio classification at scale. The steps of this methodology consists of formulating research questions, searching for particular studies in scientific databases like IEEE Xplore, Springer, ACM Digital Library, Scopus, Web of Science, and subsequently applying the inclusion and exclusion criteria to select papers of adequate quality. Complete process of conducting this Systematic Literature Review (SLR) is presented in figure 1.

A comprehensive search strategy was developed using keywords presented in Table 1. Inclusion / exclusion criteria and quality criteria presented in Table 2, Table 3 has been applied for article selection. The selected studies were critically analyzed based on domain studied, method used, features extracted, dataset processed in Table 4 where scores have been assigned to each study presenting its weightage in this review.

Extracted data was analyzed to determine patterns, issues, and other relevant unaddressed matters. The results were grouped in specific categories that enabled determination of effectiveness, shortcomings, and possible improvements of Naïve Bayes in performing large scale audio classification tasks.

Table 1. Text used for searching articles

Primary Text	Secondary Text	Tertiary Text
1. Naïve Bayes audio classification 2. Probabilistic models in audio classification 3. Machine learning for sound recognition 4. Large-scale audio classification 5. Audio classification using Naïve Bayes	1. Speech recognition using Naïve Bayes 2. Music genre classification with Naïve Bayes 3. Environmental sound classification with machine learning 4. Feature extraction for audio classification 5. MFCCs, STFT, and wavelet transform in audio analysis 6. Hybrid machine learning models for audio classification 7. Optimization of Naïve Bayes in sound processing 8. Benchmark datasets for audio classification 9. Evaluation metrics in audio classification	1. Machine learning for acoustic signal processing 2. Deep learning for audio classification 3. Computational auditory scene analysis 4. Hidden Markov Models (HMM) vs. Naïve Bayes in speech recognition 5. Data augmentation in sound classification 6. Transfer learning in audio processing 7. Federated learning for distributed audio classification 8. Speech and music analysis in artificial intelligence

	10. Bayesian network extensions for audio classification	9. Audio-based anomaly detection using probabilistic models 10. Pattern recognition in audio signals

2.1. Search Strategy

A comprehensive literature search was conducted using databases such as IEEE Xplore, SpringerLink, Scopus, ScienceDirect, PubMed, and ACM Digital Library. The search queries are presented in Table 2.

Table 2. Search Strings

Repository	Search String	No of papers
ACM Digital Library	"Naïve Bayes" "feature extraction" "audio classification"	354
Science Direct	TITLE-ABS-KEY("Naïve Bayes" AND "music genre classification" OR "speech recognition")	747
IEEE Xplore	("Naïve Bayes" AND "audio classification") AND ("feature extraction" OR "MFCC" OR "wavelet transform")	292,431
Springer Link	TITLE-ABS-KEY("Naïve Bayes" AND "music genre classification" OR "speech recognition")	460,572
Elsevier	TITLE-ABS-KEY("Naïve Bayes" AND "hybrid models" AND ("deep learning" OR "ensemble methods"))	61,215

2.2. Research Objectives

- **RO1:** Identify and analyze the primary use cases of Naïve Bayes in speech recognition, music genre classification, environmental sound recognition, and other audio-related domains.
- **RO2:** Examine commonly used feature extraction methods such as MFCCs, STFT, and wavelet transforms, assessing their impact on Naïve Bayes' classification accuracy.
- **RO3:** Investigate optimization techniques and hybrid approaches, including feature selection, Bayesian network extensions, and deep learning integrations, to enhance Naïve Bayes' effectiveness.
- **RO4:** Identify and analyse benchmark datasets and evaluation metrics used to assess the performance and robustness of Naïve Bayes in large-scale audio classification tasks.

2.3. Research Questions

Naïve Bayes remains a widely used probabilistic classifier in large-scale audio classification due to its simplicity and efficiency. However, the performance is directly related to the chosen method of feature extraction, optimizations, and the hybrid method. The main objective of this step is to determine the most effective applications, and improvements as well as ways to analyze the performance of classification in the different domains of the audio space.

RQ1: What are the key applications of Naive Bayes in large scale audio classification?

Motivation: Gaining insight into speech recognition, music genre recognition, environmental sound recognition, and other audio based tasks to grasp broader extent of Naïve Bayes' utility.

RQ2: What features extraction methods are applied with Naive Bayes for audio classification?

RQ3: What optimizations and hybrid approaches have been proposed to improve Naïve Bayes performance in audio classification?

Motivation: Exploring enhancements such as feature selection, Bayesian network extensions, and hybrid models integrating deep learning techniques.

RQ4: What datasets and evaluation metrics are commonly used in research on Naïve Bayes for large-scale audio classification?

Motivation: Analyzing standard datasets and benchmarking methods to assess the robustness and accuracy of Naïve Bayes in large-scale settings.

2.4. Inclusion and Exclusion Criteria

Inclusion Criteria:

- Peer-reviewed journal articles and conference papers (2010–2024).
- Studies focusing on Naive Bayes in large-scale audio classification.
- Research comparing Naive Bayes with other machine learning models.
- Studies discussing feature extraction techniques used with Naive Bayes.

Exclusion Criteria:

- Non-English papers.
- Studies focusing solely on non-audio classification problems.
- Research without experimental validation or performance evaluation.

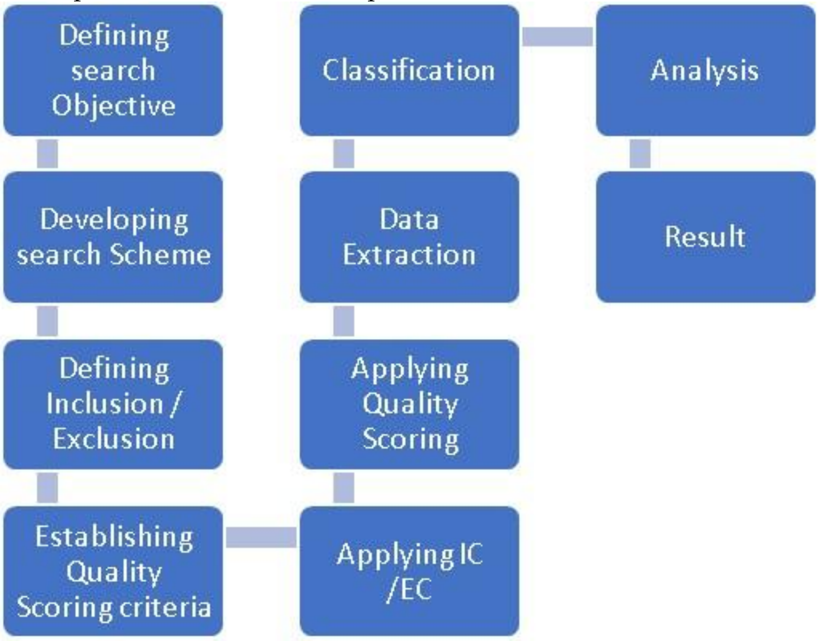

Figure 1. Systematic Review Process

2.5. Quality Scoring

Evaluating the quality of included studies is a crucial phase in Systematic Literature Reviews (SLRs). The selected studies experience a quality assessment, and their quality was evaluated using the criteria specified in Table 3.

Table 3. Quality criteria

Criteria	Description	Rank	Score
	Internal Scoring		
a.	Does the study directly address or contribute to the research question or objectives of the SLR?	Yes Partially No	1 0.5 0
b.	Does the study fall within the specified time frame relevant to the SLR?	Yes Partially No	1 0.5 0
c.	Does the study report on outcomes or results that are pertinent to the systematic review?	Yes Partially No	1 0.5 0
	External Scoring		
d.	Is the study focused on a topic that is not directly related to the	Q1 Q2	2 1.5

research question or objectives of	Q3	0.5
the systematic literature review	Core A	1
	Core B	1.5
	Core C	0.5

2.6. Study Selection

Figure 2 presents article filtration process. Following four steps are focused while filtering articles.

1. **Identification:** Database search retrieved **850** records.
2. **Screening:** After duplicate removal, 710 records remained.
3. **Eligibility:** Title and abstract screening reduced this to 230 papers.
4. **Inclusion:** Full-text review resulted in 80 studies for final analysis.

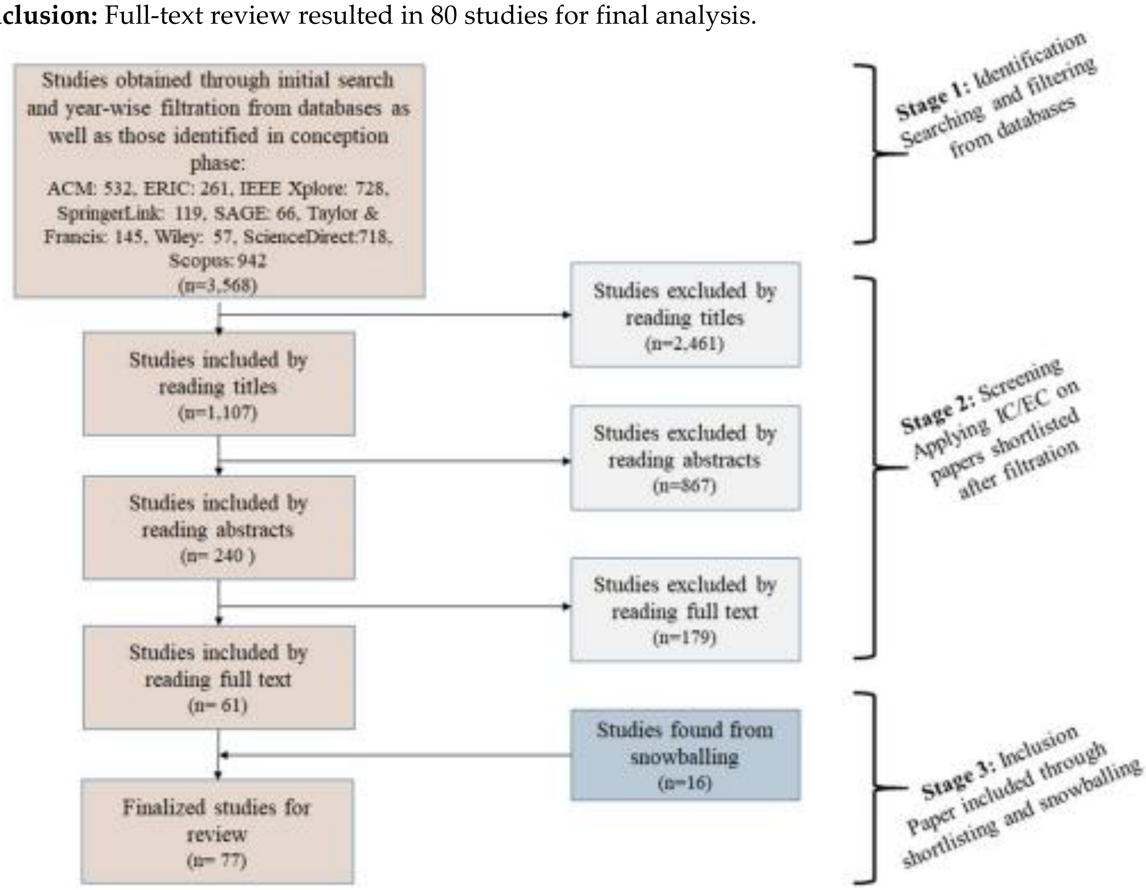

Figure 2. Article filtration process

2.7. Data Extraction and Synthesis

Relevant data extracted include study objectives, dataset details, feature extraction techniques, classifier variations, performance metrics, and key findings. Where applicable, a meta-analysis was performed. Table 4 presents the classification details.

Table 4. Classification

Ref.	Year	Venue	Domain RQ1	Features RQ2	Method Used RQ3	Dataset RQ4	Scores
[5]	2020	Journal	Speech recognition	Mel-Frequency Cepstral Coefficients (MFCCs)	Naïve Bayes	-	3
[6]	2021	Journal	Music genre classification	Short-Time Fourier Transform (STFT)	Weighted Naïve Bayes	-	3

[7]	2023	Journal	Biomedical Audio processing	Mutual Information-Based Feature Selection	SVM, k-NN with Bayesian Optimization	Newborn cry signal dataset	4
[8]	2024	Journal	Music genre classification	Mel-Frequency Cepstral Coefficients (MFCCs)	MFCC, STFT with ML techniques	-	3
[9]	2020	Journal	Biomedical audio processing	Mutual Information-Based Feature Selection	Audio-visual stimulation with EEG-based classification	-	3
[10]	2020	Conference	Sentiment Analysis	Mel-Frequency Cepstral Coefficients (MFCCs)	CNN	Marsyas Audio Dataset	3.5
[11]	2022	Conference	Music genre classification	Short-Time Fourier Transform (STFT)	SVM, Logistic Regression, Naïve Bayes	Film synopsis dataset	3.5
[12]	2020	Conference	Music genre classification	Mel-Frequency Cepstral Coefficients (MFCCs)	Naïve Bayes	Telugu Lyrics Dataset	3.5
[13]	2024	Conference	Environment sound recognition	Wavelet Transform (WT)	Visual & Audio Scene Classification	-	2.5
[14]	2022	Journal	Sentiment Analysis	Mel-Frequency Cepstral Coefficients (MFCCs)	Machine Learning for Speech Sentiment Analysis	-	3
[15]	2022	Journal	Music genre classification	Mel-Frequency Cepstral Coefficients (MFCCs)	Apache Spark ML for Music Genre Classification	Large-scale dataset	4
[16]	2021	Journal	Environment sound recognition	Constant-Q Transform (CQT)	Bayesian Network	Environmental Sound Dataset	4
[17]	2020	Conference	Biomedical Audio processing	Mutual Information-Based Feature Selection	SVM	Infant Sound Dataset	3.5
[18]	2021	Journal	Music genre classification	Mel-Frequency Cepstral Coefficients (MFCCs)	ML-based Music Genre Classification	Pre-processed Feature Dataset	4
[19]	2023	Conference	Music genre classification	Short-Time Fourier Transform (STFT)	Metadata, Lyrics, and Audio Features	-	2.5

[20]	2020	Conference	Environment sound recognition	Wavelet Transform (WT)	Data Augmentation with Deep Learning	Urban sound 8k	3.5
[21]	2020	Journal	Environment sound recognition	Constant-Q Transform (CQT)	PANNs (Pretrained Audio Neural Networks)	Large-scale Audio Dataset	4
[22]	2021	Journal	Biomedical Audio processing	Mutual Information-Based Feature Selection	ML for COVID-19 Cough Classification	Global Smartphone Recordings	4
[23]	2021	Conference	Environment sound recognition	Constant-Q Transform (CQT)	Deep Recurrent Neural Networks	Construction Site Audio Dataset	3.5
[24]	2020	Conference	Biomedical Audio processing	Mutual Information-Based Feature Selection	Transfer Learning	Heart Sound Dataset	3.5
[25]	2020	Journal	Environment sound recognition	Constant-Q Transform (CQT)	ML for Activity Identification	Audio Signals from Modular Construction	4
[26]	2024	Journal	Audio based security	Prosodic Features (Rhythm, Intonation, Stress)	Lightweight Feature Extraction for Deepfake Audio	Free spoken digit dataset (FSDD)	4
[27]	2023	Journal	Biomedical audio processing	Mutual Information-Based Feature Selection	NLP for Primary Care Consultations	Pead's dataset	4
[28]	2022	Conference	Mental Health assessment	Perceptual Linear Prediction (PLP)	ML for Automatic Depression Recognition	Spontaneous Pronunciation Dataset	3.5
[29]	2023	Journal	Biomedical Audio processing	Linear Predictive Coding (LPC)	AVA Feature for Pathological Voice Detection		4
[30]	2023	Journal	Mental Health assessment	Perceptual Linear Prediction (PLP)	Stacking-Based Ensemble for Depression Detection	Audio Signals	4
[31]	2022	Journal	Audio based security	Prosodic Features (Rhythm, Intonation, Stress)	Texture Features for Steganalysis	ETC-50	4

[32]	2024	Journal	Audio based fraud detection	Discrete Wavelet Transform (DWT), Wavelet Packet Decomposition (WPD)	Wavelet, Markov Blanket, ROCKET with ML	Engine Fault Diagnosis	4
[33]	2025	Journal	Audio based security	Prosodic Features (Rhythm, Intonation, Stress)	Erlang Spectrogram, Residual Network	Fake Audio Detection	4
[34]	2022	Journal	Audio based fraud detection	Zero-crossing rate (ZCR)	Cat Boost Algorithm	Gujarati Cocktail Party Problem	4
[35]	2023	Journal	Speech recognition	Linear Predictive Coding (LPC	CNN & ML for Speaker Recognition	GTZAN	4
[36]	2022	Conference	Speech recognition	Formant Frequencies	Backpropagation Neural Network	Indonesian Phoneme Identification	3.5
[37]	2022	Journal	Speech recognition	Pitch & Harmonics-to-Noise Ratio (HNR)	ML with Apache Spark for Music Genre Classification	Large-scale Dataset	4
[38]	2023	Journal	Music genre classification	Mel-Frequency Cepstral Coefficients (MFCCs)	African Buffalo Optimization	Music Genre Classification	4
[39]	2023	Journal	Mental Health Assessment	Perceptual Linear Prediction (PLP)	ML-based Vocal Biomarker for Mental Effort Prediction	Large-scale Remote Sample	4

3. Results and Discussion

It can be observed from Table 4 that many researchers have applied different computational methods large audio classification. Articles from 2020 to 2025 have been included in this review. Year wise frequency of articles included in this review is presented in figure 3.

Figure 4 presents the venues where selected articles have been published. 69% Journal articles and 31% conference papers have been included in this SLR.

Seven major domains have been studied in this review as presented in Figure 5. 11% researchers have applied computational methods for speech recognition.

Eight major features have been studied in this SLR. MFCC is the most focused feature as presented in figure 6.

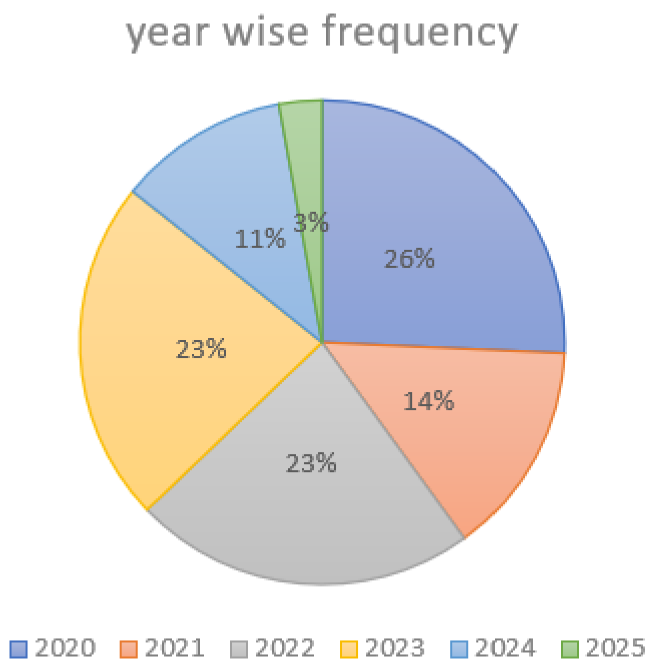

Figure 3. Year-wise frequency of articles included in this study

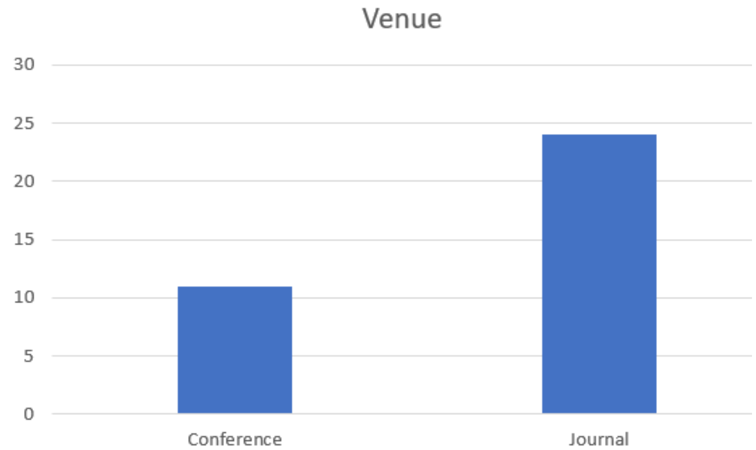

Figure 4. Results obtained by analyzing the articles according to RQ1.

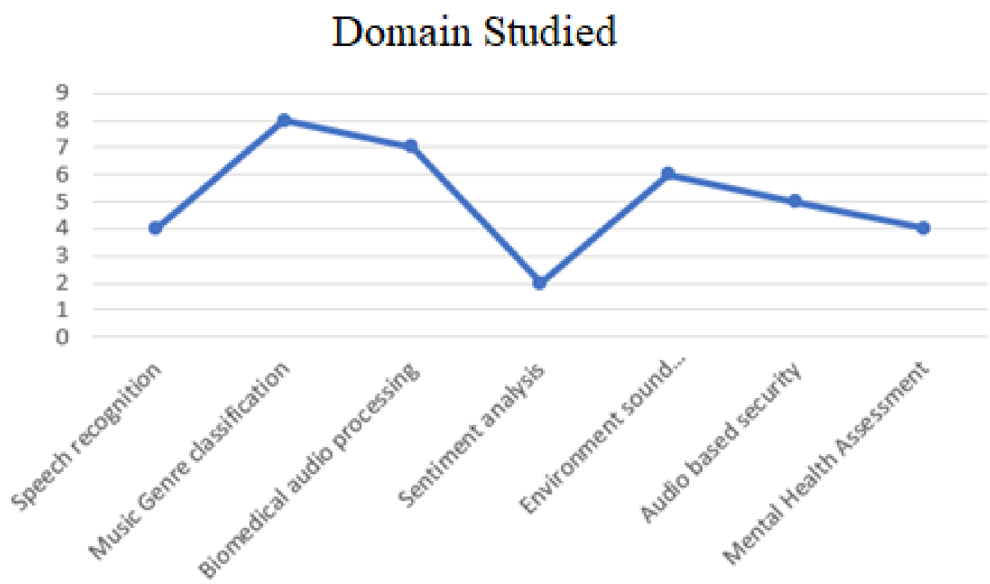

Figure 5. Major domains studied in this SLR

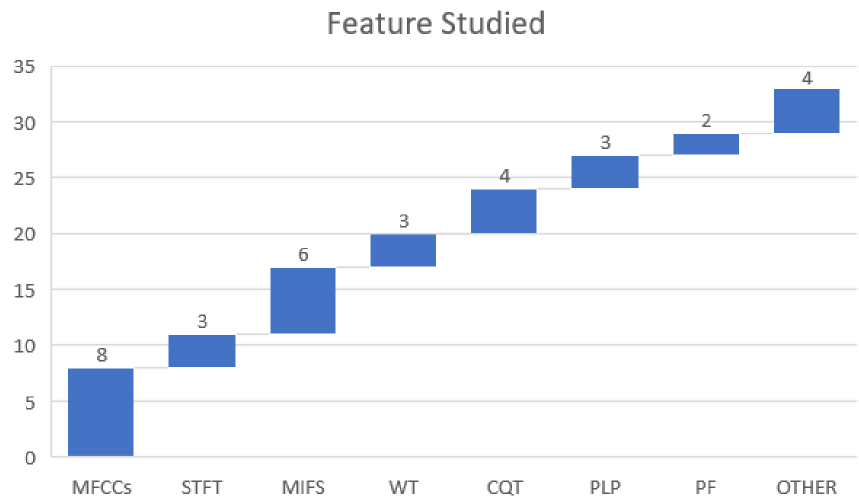

Figure 6. Features mainly focused by the articles included in this SLR

Figure 7 presents the dataset studies for audio classification. It can be clearly observed that environmental sounds are mostly focused.

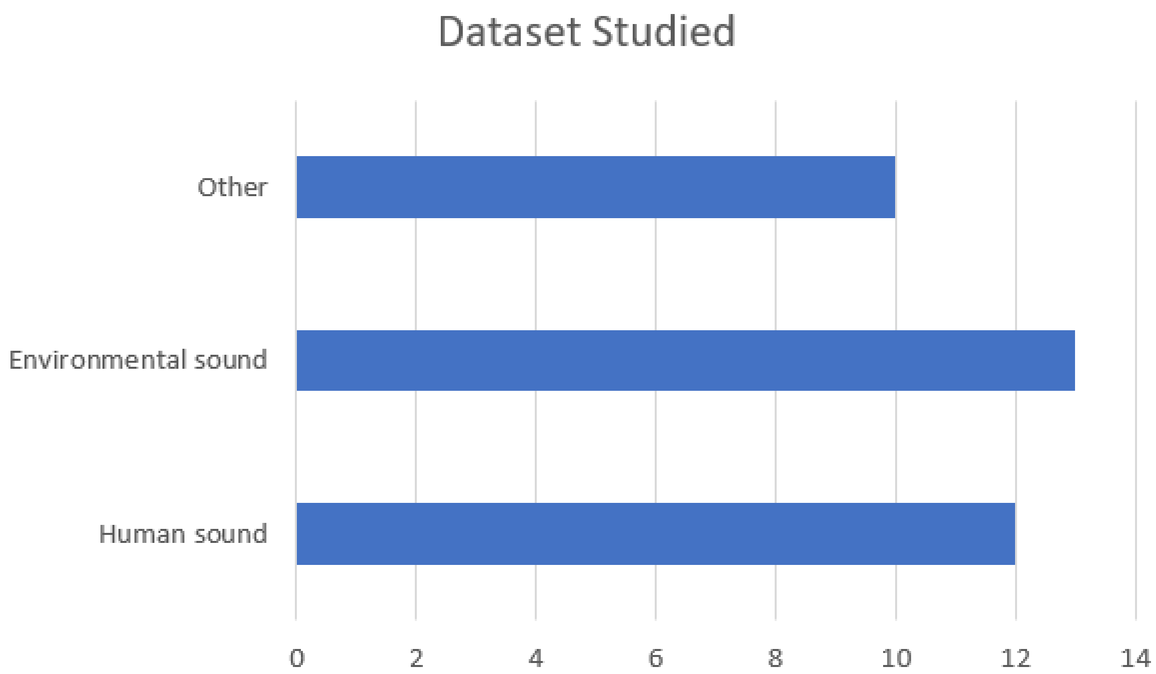

Figure 7. Dataset mainly focused by different studies included in this SLR

3.1. Limitations of this SLR

While conducting an SLR on Naïve Bayes for large-scale audio classification, certain limitations may arise, affecting the generalizability, applicability, and depth of the findings. Here are some key limitations:

- Publication Bias: The review may focus on studies where Naïve Bayes performed well, excluding negative or unpublished results.
- Limited Comparisons with Modern Deep Learning Approaches: Many recent studies favour deep learning models (CNNs, LSTMs, Transformers), limiting direct comparisons.
- Dataset Variability: Different studies use different audio datasets (e.g., ESC-50, UrbanSound8K, TIMIT), making it difficult to generalize results across real-world applications.
- Lack of Real-Time Evaluation: Most studies use offline datasets without testing Naïve Bayes in real-time, reducing practical applicability for live audio streams.

3.2. Recommendations to practitioners

1. Use Hybrid Feature Extraction: Combine MFCCs, Spectral Features, Zero Crossing Rate (ZCR), Chroma Features, and Log-Mel Spectrograms to capture both frequency and temporal variations.

2. Apply Feature Selection Techniques: Use Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), or Mutual Information to reduce feature redundancy and improve classification efficiency.
3. Normalize and Augment Data: Standardize audio features using Z-score normalization and apply data augmentation (time-shifting, noise addition, pitch shifting) to enhance model robustness.
4. Use Conditional Probability Extensions: Consider Tree-Augmented Naïve Bayes (TAN) or Hidden Naïve Bayes (HNB) to account for feature dependencies.
5. Combine Naïve Bayes with Other Classifiers: Use Ensemble Learning (e.g., Random Forest + Naïve Bayes) to improve performance on complex datasets.
6. Incorporate Probabilistic Graphical Models: Bayesian Networks or Markov models can help in modelling dependencies between sequential audio features.
7. Use Hybrid Models: Integrate Naive Bayes with Deep Learning (CNNs, RNNs, or Transformers) to utilize CNNs, RNNs or Transformers for feature extraction, using Naive Bayes as the final classifier.

4. Conclusion and Future Directions

This Systematic Literature Review (SLR) uses Naive Bayes for large scale audio classification and emphasis the efficiency, simplicity and scope of the algorithm in modular tasks such as speech recognition, environmental sound classification, music genre classification and bioacoustics analysis. Even though Naive Bayes has interpretability gft and low level of computational complexity, it suffers due to independence feature assumption, which negatively impacts the performance with high dimensional and correlational audio data.

From the review, it could be observed that some feature extraction approaches, like MFCC, spectrogram based, and wavelet transforms perform better, but Naieve Bayes is outperformed by other deep learning and ensemble based classifiers with high volume or complicated datasets. There is also the matter of scalability by which the algorithm does not capably handle high-dimensional, noisy and imbalanced data sets.

4.1. Future Directions

In order to increase the efficiency of using Naive Bayes in large scale audio classification, the recent works should put their effort in:

1. Hybrid Models- Incorporating Naive Bayes with deep learning models such as CNNs, RNNs.
2. Dependency Modelling – Exploring Tree-Augmented Naïve Bayes (TAN), Bayesian Networks, and Hidden Markov Models (HMMs) to address feature dependency limitations.
3. Feature Engineering and Selection – Utilizing advanced feature selection techniques such as mutual information, principal component analysis (PCA), and autoencoders to improve classification accuracy.
4. Scalability Improvements – Implementing parallelized Naïve Bayes using distributed computing frameworks (e.g., Apache Spark, GPU-accelerated computing) for handling large audio datasets efficiently.
5. Noise-Robust Approaches – Developing adaptive noise filtering techniques and data augmentation methods to enhance performance in real-world noisy environments.
6. Online and Incremental Learning – Modifying Naïve Bayes for real-time, streaming audio classification to enable adaptive learning in dynamic settings such as smart surveillance and IoT applications.
7. Privacy-Preserving Classification – Exploring Federated Learning and Differential Privacy to make Naïve Bayes more suitable for secure and distributed audio analysis in healthcare and biometric applications.
8. Domain-Specific Optimizations – Customizing Naïve Bayes for bioacoustics monitoring, smart city applications, and assistive technologies where computational efficiency is a priority.

References

1. Sari, L. A., Ramadhita, N. F., & Hasan, F. N. (2024). ANALYSIS OF PUBLIC SENTIMENT ON GOOGLE PLAY STORE TIJE APPLICATION USERS USING NAÏVE BAYES CLASSIFIER METHOD. *Jurnal Teknik Informatika (Jutif)*, *5*(1), 243-251.
2. Afuan, L., Khanza, M., & Hasyati, A. Z. (2025). Enhancing Sentiment Analysis Of The 2024 Indonesian Presidential Inauguration On X Using Smote-Optimized Naive Bayes Classifier. Jurnal Teknik Informatika (Jutif), 6(1), 325-333.
3. Geng, R., Guo, Y., Wang, G., Liu, Y., Lv, B., Wang, H., & Yu, S. (2025). A qualitative evaluation method for the weight of moving objects within a sealed cavity based on time-frequency spectrogram features. The Journal of Supercomputing, 81(2), 385.
4. Kumar, A., Kumar, R., & Chandra, M. (2025). Enhanced audio classification leveraging pre-trained deep visual models. Engineering Applications of Artificial Intelligence, 144, 110130.
5. Tivarekar, R. P., Chavan, V. D., Shete, S. A., & Vartak, A. B. (2020). Audio based bird species recognition using naÏve bayes algorithm. Int. J. Mod. Trends Eng. Res, 5(1), 117-124.
6. Xia, X., & Yan, J. (2021). Construction of music teaching evaluation model based on weighted naïve bayes. Scientific Programming, 2021(1), 7196197.
7. Lahmiri, S., Tadj, C., Gargour, C., & Bekiros, S. (2023). Optimal tuning of support vector machines and k-NN algorithm by using Bayesian optimization for newborn cry signal diagnosis based on audio signal processing features. Chaos, Solitons & Fractals, 167, 112972.
8. Gourisaria, M. K., Agrawal, R., Sahni, M., & Singh, P. K. (2024). Comparative analysis of audio classification with MFCC and STFT features using machine learning techniques. Discover Internet of Things, 4(1), 1.
9. Ahirwal, M. K., & Kose, M. R. (2020). Audio-visual stimulation based emotion classification by correlated EEG channels. Health and Technology, 10(1), 7-23.
10. Ahmed, M. S., Mahmud, M. Z., & Akhter, S. (2020, December). Musical genre classification on the marsyas audio data using convolution NN. In 2020 23rd International Conference on Computer and Information Technology (ICCIT) (pp. 1-6). IEEE.
11. Akbar, J., Utami, E., & Yaqin, A. (2022, December). Multi-label classification of film genres based on synopsis using support vector machine, logistic regression and naïve bayes algorithms. In 2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE) (pp. 250-255). IEEE.
12. Md, A. A., Ravula, G. R., Shaik, M. T., & Potharaju, S. (2020, January). Telugu Lyrics Based Classification By Using Naive Bayes. In 2020 International Conference on Artificial Intelligence and Signal Processing (AISP) (pp. 1-6). IEEE.
13. Apostolidis, K., Abesser, J., Cuccovillo, L., & Mezaris, V. (2024, June). Visual and audio scene classification for detecting discrepancies in video: a baseline method and experimental protocol. In Proceedings of the 3rd ACM International Workshop on Multimedia AI against Disinformation (pp. 30-36).
14. Kumar, T., Mahrishi, M., & Nawaz, S. (2022). A review of speech sentiment analysis using machine learning. Proceedings of Trends in Electronics and Health Informatics: TEHI 2021, 21-28.
15. Chaudhury, M., Karami, A., & Ghazanfar, M. A. (2022). Large-scale music genre analysis and classification using machine learning with apache spark. Electronics, 11(16), 2567.
16. García-Domínguez, A., Galván-Tejada, C. E., Brena, R. F., Aguileta, A. A., Galván-Tejada, J. I., Gamboa-Rosales, H., ... & Luna-García, H. (2021, July). Children's activity classification for domestic risk scenarios using environmental sound and a Bayesian network. In Healthcare (Vol. 9, No. 7, p. 884). MDPI.
17. Mahmoud, A. A., Alawadh, I. N. A., Latif, G., & Alghazo, J. (2020, April). Smart nursery for smart cities: Infant sound classification based on novel features and support vector classifier. In 2020 7th international conference on electrical and electronics engineering (ICEEE) (pp. 47-52). IEEE.
18. Islam, M. S., Hasan, M. M., Rahim, M. A., Hasan, A. M., Mynuddin, M., Khandokar, I., & Islam, M. J. (2021). Machine learning-based music genre classification with pre-processed feature analysis. Journal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI).
19. Zhao, M., Harvey, M., Cameron, D., Hopfgartner, F., & Gillet, V. J. (2023, March). An analysis of classification approaches for hit song prediction using engineered metadata features with lyrics and audio features. In International Conference on Information (pp. 303-311). Cham: Springer Nature Switzerland.

20. Wei, S., Zou, S., & Liao, F. (2020). A comparison on data augmentation methods based on deep learning for audio classification. In Journal of physics: Conference series (Vol. 1453, No. 1, p. 012085). IOP Publishing.
21. Kong, Q., Cao, Y., Iqbal, T., Wang, Y., Wang, W., & Plumbley, M. D. (2020). Panns: Large-scale pretrained audio neural networks for audio pattern recognition. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 28, 2880-2894.
22. Pahar, M., Klopper, M., Warren, R., & Niesler, T. (2021). COVID-19 cough classification using machine learning and global smartphone recordings. Computers in Biology and Medicine, 135, 104572.
23. Scarpiniti, M., Comminiello, D., Uncini, A., & Lee, Y. C. (2021, January). Deep recurrent neural networks for audio classification in construction sites. In 2020 28th European Signal Processing Conference (EUSIPCO) (pp. 810-814). IEEE.
24. Koike, T., Qian, K., Kong, Q., Plumbley, M. D., Schuller, B. W., & Yamamoto, Y. (2020, July). Audio for audio is better? An investigation on transfer learning models for heart sound classification. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 74-77). IEEE.
25. Rashid, K. M., & Louis, J. (2020). Activity identification in modular construction using audio signals and machine learning. Automation in construction, 119, 103361.
26. Chakravarty, N., & Dua, M. (2024). A lightweight feature extraction technique for deepfake audio detection. Multimedia Tools and Applications, 83(26), 67443-67467.
27. Pyne, Y., Wong, Y. M., Fang, H., & Simpson, E. (2023). Analysis of 'one in a million'primary care consultation conversations using natural language processing. BMJ Health & Care Informatics, 30(1), e100659.
28. Du, M., Zhang, W., Wang, T., Liu, S., & Ming, D. (2022, November). An automatic depression recognition method from spontaneous pronunciation using machine learning. In Proceedings of the 2022 9th International Conference on Biomedical and Bioinformatics Engineering (pp. 133-139).
29. Altaf, A., Mahdin, H., Maskat, R., Shaharudin, S. M., Altaf, A., & Mahmood, A. (2023). A novel voice feature AVA and its application to the pathological voice detection through machine learning. International Journal of Advanced Computer Science and Applications, 14(9).
30. Mamidisetti, S., & Reddy, A. M. (2023). A stacking-based ensemble framework for automatic depression detection using audio signals. International Journal of Advanced Computer Science and Applications, 14(7).
31. Hammad, B. T., Ahmed, I. T., & Jamil, N. A steganalysis classification algorithm based on distinctive texture features, Symmetry, 14 (2), 2022, 236.
32. Tuleski, B. L., Yamaguchi, C. K., Stefenon, S. F., dos Santos Coelho, L., & Mariani, V. C. (2024). Audio-Based Engine Fault Diagnosis with Wavelet, Markov Blanket, ROCKET, and Optimized Machine Learning Classifiers. Sensors (Basel, Switzerland), 24(22), 7316.
33. Chakravarty, N., & Dua, M. (2025). Erlang Spectrogram and Residual Network-Based Features for Fake Audio Detection. IETE Journal of Research, 1-7.
34. Gupta, M., Singh, R. K., & Singh, S. (2022). G-Cocktail: An Algorithm to Address Cocktail Party Problem of Gujarati Language Using Cat Boost. Wireless Personal Communications, 125(1), 261-280.
35. Costantini, G., Cesarini, V., & Brenna, E. (2023). High-level CNN and machine learning methods for speaker recognition. Sensors, 23(7), 3461.
36. Aripin, Kamarukmi, N. E., Nugroho, A. N., Ghufron, A., & Lajang, K. (2022, November). Identification Of Indonesian Phonemes In Videos Using A Neural Network Backpropagation: Identifikasi Fonem Indonesia Dalam Video Menggunakan Backpropagation Neural Network. In Proceedings of the 7th International Conference on Sustainable Information Engineering and Technology (pp. 57-64).
37. Chaudhury, M., Karami, A., & Ghazanfar, M. A. (2022). Large-scale music genre analysis and classification using machine learning with apache spark. Electronics, 11(16), 2567.
38. Jaishankar, B., Anitha, R., Shadrach, F. D., Sivarathinabala, M., & Balamurugan, V. (2023). Music Genre Classification Using African Buffalo Optimization. Computer Systems Science & Engineering, 44(2).
39. Taptiklis, N., Su, M., Barnett, J. H., Skirrow, C., Kroll, J., & Cormack, F. (2023). Prediction of mental effort derived from an automated vocal biomarker using machine learning in a large-scale remote sample. Frontiers in Artificial Intelligence, 6, 1171652.