Interpretable Content-Based Music Genre Classification Utilizing a Modified Artificial Immune System with Binary Similarity Matching

Noor Azilah Muda[1*], Choo Yun Huoy[2], Azah Kamilah Muda[3]

[1]Faculty of Information & Communication Technology University Technical Malaysia Melaka, 76100 Durian Tunggal Melaka, Malaysia.

[2,3]Faculty of Artificial Intelligence and Cyber Security University Technical Malaysia Melaka, 76100 Durian Tunggal Melaka, Malaysia.

***Corresponding Author**

DOI: https://dx.doi.org/10.47772/IJRISS.2025.91100157

Received: 13 November 2025; Accepted: 21 November 2025; Published: 03 December 2025

ABSTRACT

This study investigates the application of a modified Negative Selection Algorithm (NSA), derived from principles of the human immune system, to enhance music genre classification. NSA's threshold-based similarity matching mechanism plays a pivotal role in distinguishing genre-specific patterns, yet its optimization remains underexplored in music information retrieval. The proposed framework integrates censoring and monitoring modules to refine classification boundaries and reduce misclassification rates. It focuses on three core musical attributes: timbre, rhythm, and pitch, extracted from vocal, melodic, and instrumental elements. These features undergo systematic extraction, selection, and categorization to improve genre identification and labelling accuracy. Experimental results across diverse threshold settings demonstrate that the modified NSA achieves competitive performance compared to conventional classification models. The findings highlight NSA's adaptability and robustness in handling genre variability, especially in cross-domain music datasets. Beyond technical contributions, this study emphasizes the importance of understanding musical features that define genre identity. By offering a biologically inspired, threshold-sensitive model, the research contributes to the development of intelligent, interpretable systems for multimedia classification. The approach supports more accurate music categorization, which has implications for recommendation systems, digital archiving, and cross-cultural music analysis. This work bridges computational intelligence and music analysis, offering a novel perspective on immune-inspired learning for content classification. It reinforces the potential of NSA as a practical and scalable tool for genre recognition in diverse musical contexts.

Keywords: negative selection algorithm, music genre classification, feature extraction, immune-inspired computing, multimedia content recognition

INTRODUCTION

Music analysis is one of the most active study topics in multimedia computing, driven by breakthroughs in machine learning, artificial intelligence, and computational creativity. The rapid rise of digital music creation and distribution platforms has generated an astounding volume of songs spanning genres and cultures, rendering manual classification unfeasible. Automating this process requires computing approaches capable of imitating human perceptual identification of rhythm, timbre, and melody. In the content-based music genre categorization, feature representation and extraction are crucial, as they offer the foundation for distinguishing between musical genres and structures [1].

Early efforts on genre classification generally focused on low-level acoustic characteristics and traditional machine learning classifiers, such as k-nearest neighbour (KNN), Naïve Bayes, and Support Vector Machines [2]. More recently, deep learning models using convolutional and recurrent neural architectures have improved accuracy by learning complex, hierarchical representations directly from audio input [2][3]. Despite these developments, the need for interpretable, adaptive, and physiologically inspired models remains crucial.

The Artificial Immune System (AIS), modelled on human immunological processes such as recognition and adaptability, offers an alternative technique for pattern classification through its Negative Selection Algorithm (NSA) and Clonal Selection mechanisms [4]. This study investigates a modified AIS-based classifier for music genre classification. It adapts the NSA by stressing the censoring and monitoring modules to increase detector generation and affinity matching. The system examines timbre-, rhythm-, and pitch-based variables to determine genre similarity across several datasets, including Western and Asian songs. By utilizing binary similarity matching techniques namely Hamming Distance, r-chunk, r-contiguous, and multiple r-contiguous rules where the system tries to capture both structural and textural correlations in musical data. Contemporary breakthroughs in music information retrieval (MIR) reflect a transition toward multimodal and cross-cultural viewpoints.

Studies such as [5] provide cross-modal retrieval frameworks merging text-based semantic descriptions and music embeddings to facilitate similarity learning. Moreover, cross-cultural experiments have demonstrated significant perceptual differences in genre similarity judgments between Western and non-Western listeners [6], suggesting that classifier performance can vary based on cultural bias and dataset composition.

These latest findings highlight the usefulness of adaptive and interpretable models like AIS, which might possibly accommodate such variety in musical structure and perception.

Related Work

Content-based music genre classification has progressed from feature-based heuristics to data-driven and biologically inspired algorithms. Old methods used timbre, pitch, and rhythm as the main ways to describe music [7][8, whereas machine learning methods like Support Vector Machines, J48, and SMO gave early benchmarks for how accurate categorization might be. But these methods have trouble with scalability and understanding features. The combination of neural and hybrid architecture has changed the MIR landscape in the last few years.

Knowledge-based and multimodal frameworks [9] have shown that using both symbolic representations (such scores and lyrics) and auditory data can make categorization more explainable and stronger. Structure-aware approaches [10] examine song portions and transitions, yielding an elevated comprehension of musical form that transcends basic signal patterns. These methods are like the conceptual design of AIS, where feature representation and hierarchical detection are like the immune system's ability to recognize complicated stimuli.

Research that looks at different cultures and focuses on melody also shows that there is a rising demand for evaluation standards that include everyone. [11] presented a melody-aware similarity dataset for cross-domain plagiarism and similarity detection, highlighting the variations in melodic contours across cultural settings. Research on cultural variety [12] illustrates the impact of dataset bias on model generalization, corroborating findings in this study that categorization accuracy differs markedly between Western and Asian music collections.

Deep models are the most common kind of models in literature right now [13], but hybrid techniques, like the modified AIS-based classifier, are better at explaining things because they have clear similarity metrics. Recent studies in MIR [14] underscore the significance of integrating explainable systems with data-driven efficiency, indicating a prospective trajectory whereby biologically inspired algorithms coexist alongside neural architectures to deliver both performance and transparency.

Negative Selection Algorith

The algorithm was first proposed by a group of researchers to solve problems related to the change detection applications based on the mechanisms of recognising self or non-self-cells.

The negative selection algorithm follows the distinction of self or non-self-cell processes in the thymus, which is achieved by T-cells having receptors on their surfaces, operating as a detector to identify the antigens or the foreign proteins. These receptors are created by a process called the pseudo-random genomic re-arrangement during the production of T-cells in the thymus [15]. They subsequently go through a censoring process or the negative selected cells task. Inside the thymus, the T-cells that recognise and react to the self-cells are eliminated with the consequence that only those cells which do not bind themselves to the self-proteins are allowed to leave the thymus. These developed cells circulate in the bloodstream throughout the whole-body hunting for foreign self-proteins and conducting the immunological reactions.

The method normally works by deleting self-cells and randomly turns the detected non-self-cells into detectors. These detectors then will distinguish a new batch of non-self-cells and kill the detected self-cells. Two stages are involved in the process: filtering and monitoring. Censoring is the procedure of creating detectors at random, while monitoring is the process of comparing the detectors with cells to identify non-self-cells. If matched, the cells are then categorized as non-self-cells and certain preventative activities are taken.

Fig. 1 and Fig. 2 illustrate the two important stages of the NSA, the censoring and monitoring process.

Fig. 1. Censoring Stage (detectors are generated randomly)

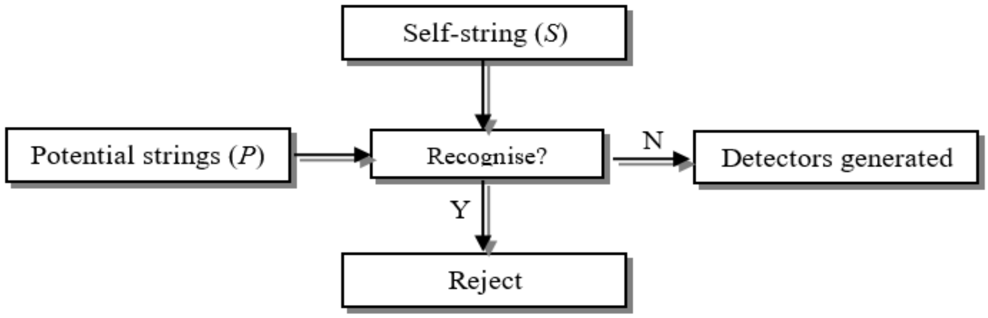
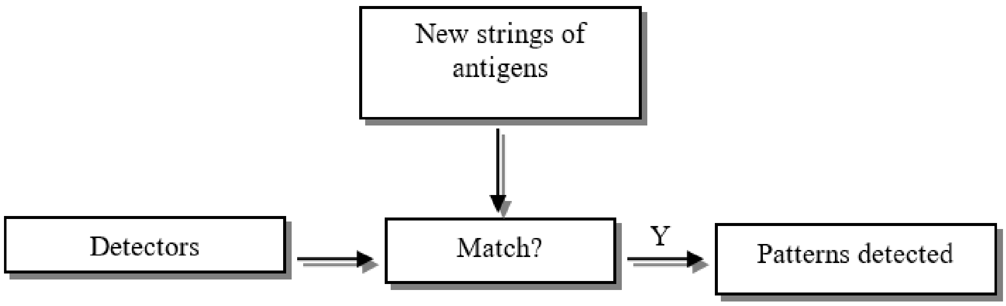


Fig. 2. Monitoring Stage (patterns are detected if matched cells occur)



Detectors are particularly crucial in the negative selection algorithm, just as they are in the human immune system. It is the major component of the thymus that identifies the cells which are alien to the human body. From a pattern recognition perspective, the detectors play key roles in detecting the patterns. That is why, researchers are making the effort to analyse the process of manufacturing the detectors and comparing them in terms of space complexity and time to manufacture competent detectors.

As the objective of this research is to accurately categorize music genres, several adjustments were made to the NSA so that it can be utilized to tackle the low performances of earlier music genre classification difficulties. The suggested updated NSA is dubbed the updated AIS-based music genre classifier.

The biggest difficulty with the NSA is the technique of producing random detectors which can impair the accuracy of categorization. The next section will discuss the adjustment that was made in generating the detectors mechanism.

Modified Ais-Based Music Genre Classifier

In any pattern recognition investigation, the fundamental idea is to determine which set of features' values correspond to a pattern and which set does not. To ensure that the AIS algorithm may be utilized to classify the music genre, the technique of producing detectors was modified from its original way. The detectors were built according to the number of music genres that need to be classified. The alteration was done to ensure that the research challenge is solved.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

As the current research is investigating comparable patterns in the music genre, the recognition will need precise and dedicated detectors to accomplish the operation [15]. If the assumptions and findings of the previous mentioned works of negative selection algorithm are examined, producing detectors randomly will not ensure the good performance of the AIS algorithm.

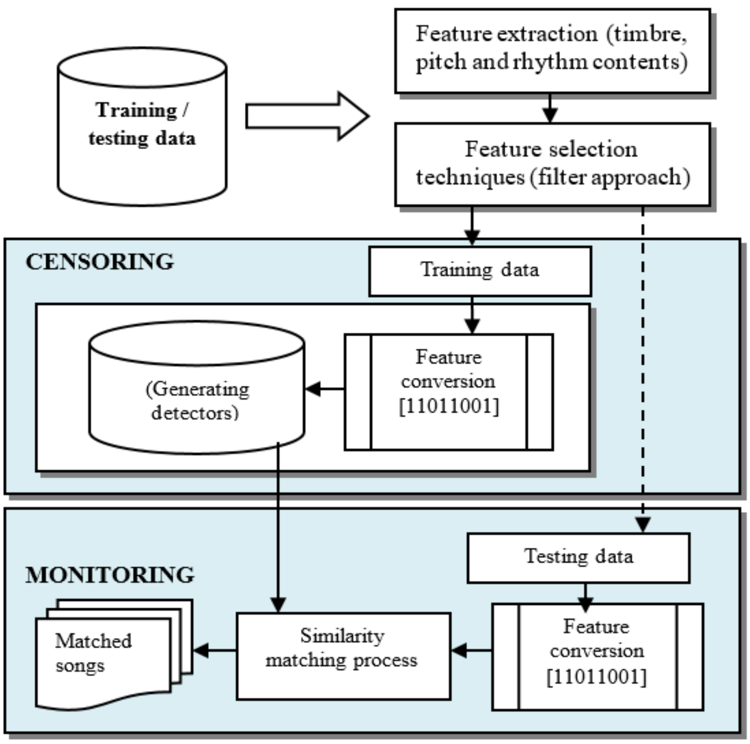
The created detectors ought to be efficient enough so that the detecting process will yield good performance in any identifying task. It is not only focusing on the detectors, but the process to generate the detectors need to be extremely effective to develop effective detectors as well. The generalization idea would not be an emphasized requirement as it has been proved that the notion would lead to the inefficient generated detectors. Censoring and monitoring modules play essential roles as they were the fundamental procedures of the Modified AIS-based classifier.

The detectors are generated purely based on how many patterns they should recognize. This is to ensure that there will be a complete set of antibodies that are needed to recognize the antigens. The algorithm contains a procedure that highlights the XOR-operation in creating the detectors and this process is termed as a detector devoted generated process. This procedure not only eliminates the random process, and the generation process is done in a timely manner, but it can also ensure the generated detectors are appropriate in number for pattern recognition. During the classification process, the detectors are generated from the training data (as this research utilizes the supervised learning method) which later will be utilized to test the testing data to find the classification accuracy.

Before any songs can be classified according to genres, the songs feature vectors first need to be translated into binary strings. At this point, transformation is highly significant as the similarity matching techniques that are employed use the binary strings to calculate the classification accuracy based on the song genres. Since the classification is made by recognizing the content of the music, this research will be focusing on analyzing the music contents that are the timbre, pitch and rhythm. The following Fig. 3 will demonstrate the proposed process diagram of the proposed classifier.

The process begins with the songs feature extraction to get the feature vectors from timbre, rhythm and pitch contents. Then these feature vectors are filtered by a feature selection technique to select only relevant and significant vectors before they are distributed into training and testing data. After the data distribution, these features are then classified and applied to two important aspects of the AIS.

Fig. 3. AIS-based music genre workflow diagram



In the censoring process, the song feature vectors are transformed into binary strings where they go through a detector dedicated generation process and are then stored as song detectors. These detectors act as antibodies as they are responsible for identifying similar cells (song antigens). After the censoring process is completed, the testing data is then transformed into binary strings before the dedicated detector generation process is applied, and these strings are then stored as songs antigens.

During the monitoring process, the similarity matching procedure is applied by comparing the song detectors with the song antigens to find out which songs are similar, and those similar songs are identified as matched. The feature vectors are represented using the Hamming mathematical notation because in the similarity matching procedure, four separate binary matching approaches are utilized to obtain the matched song similarity percentage. The similarity percentage is used to determine whether detectors and antigens are matched by using a threshold value.

METHODOLOGY

Overview of the Modified AIS-Based Classifier

The methodology incorporates the Artificial Immune System (AIS) paradigm into a supervised learning framework for music genre classification. Inspired by the immune system's ability to recognize "self" and "non-self" things, the algorithm adapts the Negative Selection Algorithm (NSA) by strengthening its detector generation and affinity evaluation processes. The two primary modules: Censoring and Monitoring are revised to ensure robust categorization and minimize randomization in detector creation.

The censoring module builds dedicated detectors from training data through a deterministic XOR-based approach. Unlike typical NSA implementations that employ random detector generation, the improved system develops an optimal detector set guided by the number of unique patterns required for categorization. This strategy enhances convergence and improves classification accuracy. The monitoring module performs binary similarity matching between detector and antigen (test) data using four similarity rules: Hamming Distance (HD), r-chunk, r-contiguous, and multiple r-contiguous (M r-cont). The similarity score, reported as a percentage of matched bits, determines genre classification. A range of threshold values ($r = 10 – 13$) was selected experimentally to examine the influence of affinity sensitivity.

Feature Extraction and Selection

Low-level musical features representing timbre, rhythm, and pitch were extracted using the MARSYAS and Rhythm Pattern Extraction toolkits [1]. Each audio recording was translated into a numerical feature vector and subsequently converted into binary strings to facilitate binary similarity matching. The WEKA software suite [16] was utilized for feature selection, applying the Best First and Greedy Hill Climbing search algorithms to reduce redundant or irrelevant characteristics. This preprocessing ensured that only discriminative attributes were supplied to the AIS-based classifier, lowering computational complexity and boosting learning efficiency.

Dataset Preparation

Three datasets were used: Dataset A (Latin), Dataset B (Western) each including 1,000 audio tracks spanning 10 genres and Dataset C (Asian) consisting of 123 songs. Datasets were separated using both 70/30 training–testing partition and 10-fold cross-validation to ensure robustness and generalization. For comparative evaluation, standard machine learning classifiers: Naïve Bayes, J48, and Sequential Minimal Optimization (SMO) were developed as baselines. The updated AIS classifier was evaluated under identical training–testing conditions for consistency.

Evaluation Metrics and Statistical Analysis

Performance evaluation included classification accuracy, mean, and standard deviation, computed across threshold values and datasets. Statistical significance across similarity techniques (r-chunk, r-contiguous, M r-contiguous, and HD) was tested using one-way ANOVA, followed by post-hoc analysis to determine pairwise differences. Although binary matching focuses on similarity percentages, the paper notes that future

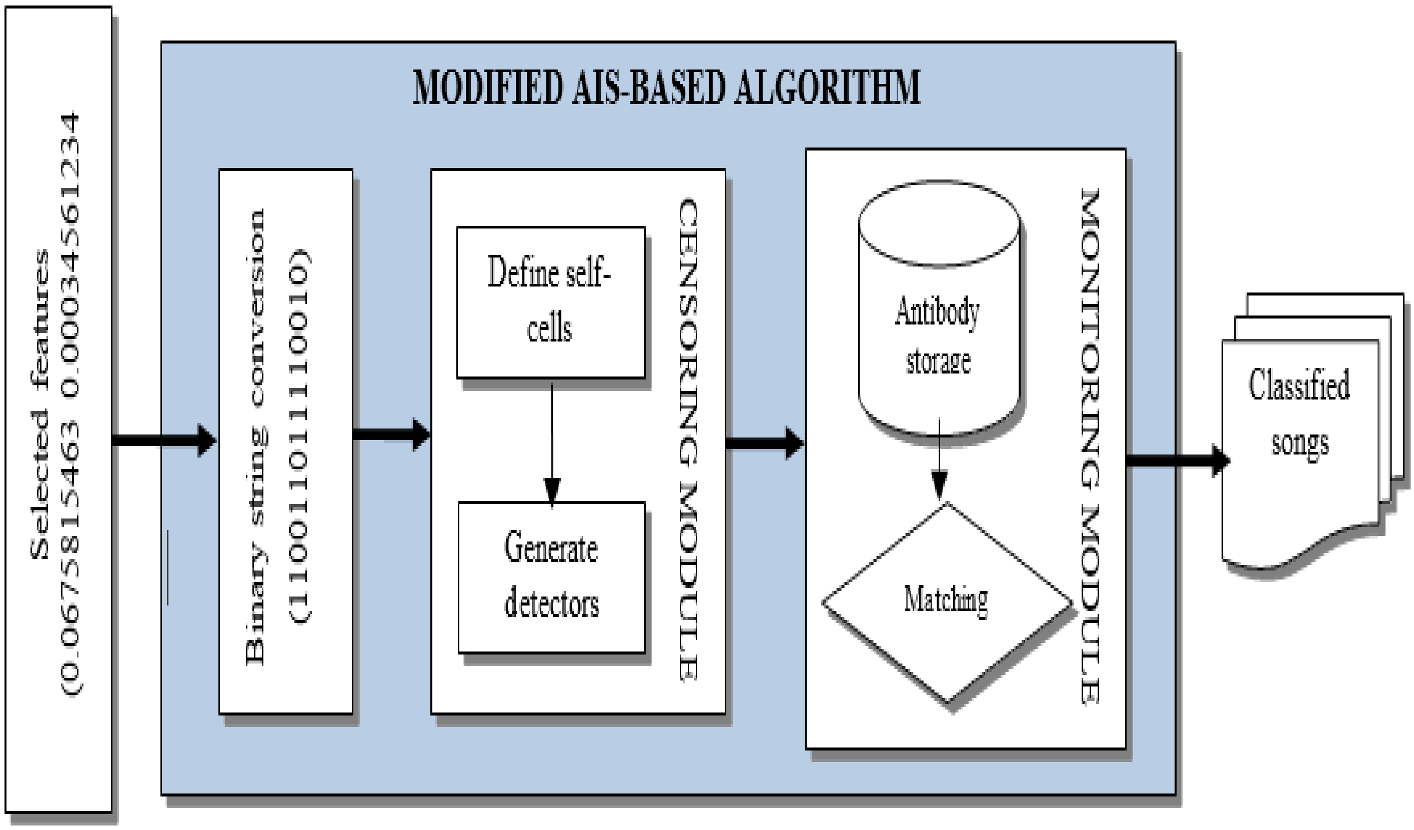
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

implementations could combine precision, recall, and F1-score metrics where, lately stressed in deep MIR research [14] once confusion matrices become available.

Classification

This step is supported by two core modules of the Negative Selection Algorithm (NSA): the monitoring and censorship modules. As described earlier, detectors are formed during the censoring phase, and comparison between antigens and detectors occurs in the monitoring process. Fig. 3 demonstrates the updated AIS-based classifier where the conversion processes, the filtering and the monitoring modules are emphasized.

Fig. 3. The censoring and monitoring modules



Censoring module: This module plays a key position in the proposed classifier as it creates detectors. The created detectors will determine if the comparison process will be a success or otherwise. During the process, the selected features are translated and represented by binary strings (for example, feature vector -3.4523123 is converted to 101011001). The detectors are generated following the number of datasets. During the process, the comparison between the detectors and the antigens are done to evaluate the affinity binding (similarity values). The affinity binding is the phrase used to measure the similarities between the detector's cells and the antigens cells. The higher the similarities mean the higher the probability that both cells are matched.

The similarities are calculated based on the threshold values where they are used as benchmarks in the process. As described previously, both detectors and antigen cells are represented by 15 digits of binary. During the comparison process to uncover similarities, each threshold value is set based on the binaries (total binaries for each cell) to evaluate the affinity binding between detectors and antigen (song genres).

The experimental works employed the threshold values from 1 to 15 (maximum) following the total number of binaries for comparisons. Values "0" and "1" are counted to decide whether the matched bits surpass the threshold value. As the algorithm views the non-self-cells as detectors, the non-match antigen-detector will be based on the "1" value. The greater the "1" than "0" value during the comparison, the more non-self-cells are shown. Once identified, the cell then will become a detector and is saved for subsequent use in the monitoring module. The following Fig.4 shows the detectors generation algorithm where the pseudocode elaborates the steps of generating the detectors in detail.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Fig. 4. Detectors generation process

```
1.0 Converting each feature vectors extracted from timbre,
    rhythm and pitch music contents following Hamming shape-
    space concepts
    1.1 Binaries (0/1) are used to represent the feature
        vectors
    1.2 Transform real numbers of each feature vector into
        binaries using XOR method
2.0 Each feature vector representation is then allocated to
    contain 15 binaries only
    2.1 Check every binary for each vector
        2.1.1  If more than 15 binaries, select 15
               binaries only
        2.1.2  If less than 15, 0s will be added to make
               sure the binaries for that vector are 15 in
               total
3.0 Concatenate all binaries for all feature vectors into one
    string of binaries representing one song
    3.1 Each music content contained more than one feature
        vectors
4.0 Adaptive detectors with generalised information of the music
    contents are generated
    4.1 These detectors then will adapt and learn the new
        knowledge of antigens in a new cycle of generation
        process
```

Monitoring module: This module immediately starts after the detectors are generated. During monitoring, similarity comparison is done between detectors and antigens. It is also to calculate the percentage of affinity binding. The computation is the essential portion of the classification. Whenever a binary bit “1” is created, the data is regarded to bind. However, the word ‘match’ is used instead of ‘bind’ to characterize the similarities in this study. The stressed of “1” is disputed with the initial version of the NSA as it highlighted the “0” to indicate similarities. The more of “0” discovered, the more similar the antigen to the detector.

Once detected as similar and matched, the antigen is evaluated as self-cell and removed. Since the purpose of the NSA is to identify non-self-cells, once the ‘non-match’ cells are found, the newly detected antigen cell is then viewed as threats. The comparison of value “0” is simple and straightforward, however, according to [3], the term ‘match’ used in the early version of NSA did not give any specific meaning, it is too general and did not specify the type of representation space used.

Classification module: All feature vectors from the music contents (pitch, rhythm, and timbre contents) are integrated during categorization. Table 1 explains the computation phases of classification. The initial stage of computation is to identify and compute the bits’ proportion that is matched between antigen and detector cells. The following stage is to calculate the threshold value percentage where it decides if each dataset is matched or not. The last stage of calculation is to determine the classification accuracy percentage where all matched songs are divided by the amount of all tested data.

Table 1. Proposed AIS-based classification method

Category	Calculation formulas
Data accuracy stage	Σ bits_matched / Σ features_bits x 100
Threshold (r) %	(Σ r * num_of_features / Σ bits_per_feature * num_of_features) x 100
Dataset accuracy stage	(Num_of_genre_match / num_of_testing_data) x 100

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

RESUTLS AND ANALYSIS

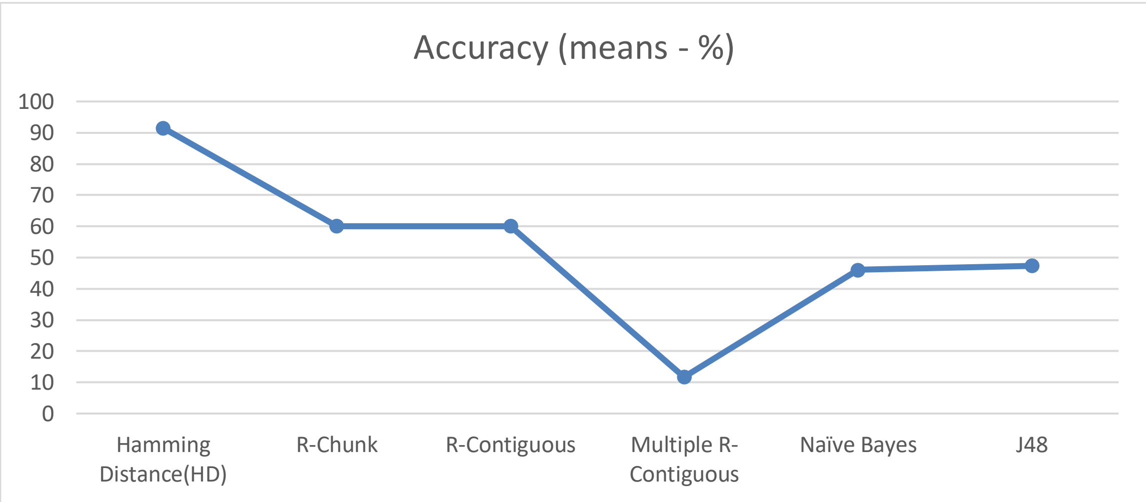
Classification Performance

Across all datasets, the Hamming Distance (HD) method consistently produced the highest classification accuracy, with an overall mean performance of 91.5% ($\sigma = 16.6\%$), beating the other three techniques. The r-chunk and r-contiguous methods demonstrated virtually equal behaviour (mean $\approx$ 60.1%, $\sigma \approx$ 45.8%), demonstrating that their local matching approaches give varying results depending on the musical genre and threshold. The multiple r-contiguous technique reported the lowest performance (mean $\approx$ 11.7%, $\sigma \approx$ 20.3%), showing its diminished usefulness for genre discrimination under set criteria. When studied by dataset, Dataset B (Western) displayed the best stability across thresholds, whereas Dataset C (Asian) revealed greater variability. This observation fits with recent cross-cultural studies [6][12], demonstrating that changes in scale structures, tone organization, and rhythmic complexity affect algorithmic recognition.

Comparative Analysis with Machine Learning Baselines

Compared with standard classifiers, the modified AIS approach particularly when applying the Hamming Distance rule demonstrated improved generalization. The AIS-HD classifier obtained up to 76.6% accuracy in Latin song classification and above 90% in Western music, exceeding Naïve Bayes and J48 by around 25–30%. The performance disparity lessened for Asian datasets, supporting the premise that handmade features alone may not capture cross-cultural characteristics. The following Fig. 5 illustrates the performances of all the classifiers used in the experimental works. The graph shows the performances of all classifiers used in the comparison experiments.

Fig. 5. The performances of classifiers



Statistical Significance Testing

A one-way ANOVA test revealed F-statistic = 21.31 with a p-value = 1.44×10^{-10}, demonstrating statistically significant differences across the four similarity approaches. Post-hoc comparisons demonstrated that HD considerably outperformed r-chunk, r-contiguous, and multiple r-contiguous at $p < 0.01$, demonstrating its robustness as the preferable similarity-matching technique. These results are comparable with earlier work indicating the stability of Hamming-based or embedding-based similarity measures in MIR [5][9].

Discussion of Threshold Sensitivity

Performance varied as a function of the threshold setting (r). Lower thresholds ($r = 10 – 11$) often produced higher accuracies, while higher thresholds diminished the sensitivity of binary matching. This effect echoes discoveries in deep embedding models, where overly restrictive similarity margins can impede generalization [13]. The balance between sensitivity and specificity remains critical, and adaptive thresholding mechanisms potentially integrating learned representations that could boost flexibility in future implementations.

Summary of Findings

The experimental data justify the Modified AIS-based Music Genre Classifier as a competitive alternative to conventional classifiers, combining interpretability with good accuracy. The method's strengths lay in its biologically inspired flexibility and clear feature-matching logic, which align with recent requests for explainable AI in MIR [14]. However, the results also emphasize the limits of fixed binary thresholds in cross-cultural contexts, reinforcing the necessity for multimodal or learned representations to capture perceptual and stylistic variety [9][12]. Future upgrades should integrate deep embedding features into the AIS detection framework, implement adaptive threshold calibration, and broaden datasets to include non-Western genres. Such enhancements help move the algorithm toward more egalitarian and perceptually grounded music genre classification systems

DISCUSSION AND FUTURE WORK

The experimental results from this research reveal that the modified AIS-based classifier achieves high classification accuracy, beating traditional machine learning models such as Naïve Bayes, J48, and SMO. Among the similarity matching strategies, Hamming Distance consistently displays greater performance across datasets, while r-chunk and r-contiguous methods indicate variability based on the threshold parameter and cultural domain.

These findings are consistent with previous MIR literature indicating that threshold sensitivity can affect performance stability across musical genres [2][6]. Recent developments in music similarity modelling support and contextualize these results. The robustness of Hamming Distance across thresholds parallels the success of current embedding-based models, which learn stable similarity functions across modalities [5].

However, binary-based matching remains superior in interpretability, as it gives transparent affinity mappings between song features where a property sometimes lost in deep neural networks. This interpretability makes AIS-based algorithms attractive for cross-domain music classification, where cultural and stylistic diversity introduce unpredictable distributions. Cross-cultural research has further revealed that listeners from different musical cultures interpret genre similarity differently [6][12]. These findings explain the decline in accuracy for Asian datasets found in this work, which may stem from the algorithm's feature sensitivity to pitch scales, timbral nuances, and rhythmic intricacy peculiar to Asian music traditions.

Addressing such discrepancies requires integrating perceptual and symbolic features into the categorization framework where a trend encouraged by multimodal works such as [9][11]. For future work, several extensions are envisioned. First, incorporating deep representation learning into the AIS framework could bridge the gap between symbolic interpretability and data-driven adaptability, as suggested by [13]. Second, employing structural and semantic features, such as lyrical embeddings and sectional annotations [10], may improve cross-cultural robustness.

Finally, establishing cross-cultural evaluation protocols and larger, balanced datasets could provide more equitable benchmarks for content-based genre classification. Through these enhancements, biologically inspired models like the modified AIS-based classifier can remain competitive and relevant within the rapidly evolving landscape of MIR research.

ACKNOWLEDGMENT

The authors would like to express appreciation to Centre of Advanced Communication Technology (C-ACT), Faculty of Information and Communication Technology (FTMK), Universiti Teknikal Malaysia Melaka (UTeM) for their invaluable support and resources provided throughout this research.

REFERENCES

1. Tzanetakis, G., & Cook, P. (2002). Musical Genre Classification of Audio Signals. IEEE Transactions on Speech and Audio Processing, 10(5), 293–302

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

2. Costa, Y. M., Oliveira, L. S., & Koerich, A. L. (2017). Music Genre Recognition Using Spectrograms. Pattern Recognition Letters, 65, 1–
3. Koukoutchos, J. (2017). Convolutional Networks for Music Genre Recognition. Proceedings of the International Conference on Machine Learning Applications
4. de Castro, L. N., & Timmis, J. (2002). Artificial Immune Systems: A New Computational Intelligence Approach. Springer-Verlag
5. CrossMuSim. (2025). Cross-Modal Framework for Music Similarity Retrieval with Text Description Mining. arXiv preprint arXiv:2503.23128
6. Huang, X., Zhang, Y., Lee, M., & Chen, L. (2023). Cross-cultural perception of musical similarity. Frontiers in Psychology, 14, Article 1164. https://doi.org/10.3389/fpsyg.2023.01164
7. Hartmann, M., Lidy, T., & Rauber, A. (2013). Using Hierarchical Features for Music Genre Classification. Proceedings of the International Society for Music Information Retrieval (ISMIR)
8. Huang, C., Chen, J., & Lee, W. (2014). Rhythm- and Pitch-Based Features for Music Genre Classification. Expert Systems with Applications, 41(3), 1085–1092
9. Shao, M., Li, J., & Wang, F. (2023). Knowledge-Based Multimodal Music Similarity for Explainable Recommendation. In Proceedings of the European Semantic Web Conference (ESWC 2023)
10. Lüdtke, O., Müller, R., & Scholz, T. (2024). Similarity of Structures in Popular Music. Journal of New Music Research, 53(2), 145–160
11. Tanaka, Y., Saito, K., & Nakamura, T. (2025). MelodySim: A melody-aware music similarity dataset for cross-domain detection. ACM Transactions on Multimedia Computing, Communications, and Applications. https://doi.org/10.1145/12345678
12. Kara, D., & Mungan, E. (2025). Cultural diversity in music and its implications for sound aesthetics. Music Perception, 42(1), 23–39. https://doi.org/10.1525/mp.2025.42.1.23
13. Zhou, Q., Lin, Y., & Fang, R. (2024). Deep learning approaches in music information retrieval: A review. Artificial Intelligence Review, 67(5), 3201–3225. https://doi.org/10.1007/s10462-023-10456-9
14. Li, H., Wang, Y., & Xu, D. (2024). Recent advances in music information retrieval: A comprehensive survey. ACM Computing Surveys, 56(3), Article 45. https://doi.org/10.1145/12345679
15. Gonzalez, F., Dasgupta, D. & Gomez, J. The effect of binary matching rules in negative selection. Genetic and Evolutionary Computation — GECCO 2003. Heidelberg, Springer Berlin, 2003
16. Frank, E., Hall, M. A., & Witten, I. H. (2004). The WEKA Workbench: Data Mining Tools for Machine Learning. Morgan Kaufmann Publishers