J. Electrical Systems 20-10s (2024): 591-599

[1]Yu Wang

[2]Lin Zhu

The Collision and Fusion of Minority Music and Popular Music in Multimedia Network Real-Time Video Processing

Abstract: - The integration of minority music and popular music in multimedia network real-time video processing is being revolutionized by constant advancements in information technology (IT). The rapid development of data technology and artificial intelligence (AI) has introduced new interdisciplinary concepts to music education. This research presents a novel approach to music education, leveraging machine learning techniques. A comprehensive dataset, documenting students' diverse interests and activities, was analyzed and filtered using a Gaussian kernel filter. Subsequently, feature extraction was performed using a Boltzmann spatio convolutional neural network (GKF_BSCNN). Experimental investigations were conducted to evaluate the approach in terms of random accuracy, Area Under the Curve (AUC), precision, and recall across various student datasets. The proposed method demonstrated significant efficacy, achieving 97% random accuracy, 94% precision, 89% recall, and 40% AUC. This approach not only provides valuable assessment results with broad applicability but also mitigates the subjectivity inherent in traditional evaluation methods.

Keywords: Music Education, Artificial Intelligence, Machine Learning, Gaussian Kernel Filter, Convolutional Neural Network

1. Introduction:

Science and technology have never been separated from art. As Li Zhengdao said, "Science and workmanship are indistinguishable, very much like the different sides of a coin." Music as a type of craftsmanship is normally indistinguishable from science. With the coming of the Web period, advanced education organizations keep on further developing their own informatization showing level, bringing about a lot of information connected with the showing system [1]. The most effective method to utilize this data to work on the nature of educating and logical examination administrations in colleges and the degree of the board direction has turned into the greatest test confronting colleges. Utilization of AI innovation can really dissect and separate secret data in huge datasets, that is to say, finding examples and information in a lot of information and foreseeing results or ways of behaving. The optimization of educational resource allocation, the prediction of student academic performance, academic planning, and the enhancement of the future development of alumni are just a few of the applications of machine learning that have gained popularity in higher education over the course of recent years. It has additionally brought forth the introduction of new instructive exploration fields [2]. Unearthing inside and out data from instructive elements like understudies, educators, showing associates, graduated class, and training overseers can assist schools and colleges with allotting different showing assets and coordinate instructive exercises all the more successfully, and all the more actually work on understudies' fulfillment with courses, to improve the learning impact of understudies and increment the enlistment pace of majors [3]. The evaluation of teachers' curriculum performance is a common issue in higher education. Understudies are the main wellspring of data for the educational experience and the main individual who can assess the quality, adequacy and fulfillment obviously satisfied, showing techniques, showing materials, and tasks. Understudy assessment is predominantly used to work on the nature obviously educating and, simultaneously, as the reason for showing evaluation of instructing staff. Profound learning is for the most part founded on exact figuring out, inside and out investigation, and reflection and driven by students' natural inspiration, to fundamentally and independently learn new information and groundbreaking thoughts, take care of useful issues, and develop understudies' profound learning and investigation capacities [4]. The role that this model plays in encouraging students to

[1] School of Music and the Performing, Yibin University, Yibin 644000 China

[2]College of Music, Chong'qing Normal University ,Chong'qing, 401331, China

Corresponding Author: Lin Zhu

1.Yu Wang：wangyu19841@126.com

2.Lin Zhu：zhulin202003@163.com

develop their abilities, shape their character, and raise their quality is crucial. erefore, in middle school music training educating, educators ought to sensibly utilize the profound learning model to further develop understudies' extensive music capacity. A subfield of machine learning called "deep learning" takes its name from a neural structure. ese networks remove the elements consequently from the dataset and are equipped for learning any nonlinear capability. The term "universal functional approximator" refers to neural networks. Understanding learning, effectively addressing common problems, and gaining a fundamental grasp of new information and ideas while integrating them into an existing framework are all aided by the profound learning approach. Extreme emphasis is placed on foundational knowledge in this learning style, with students being instructed to experimentally assimilate data and then duplicate the data set [6].

The natural combination of AI and music education has enhanced the technological methods of instruction, increased the capabilities of intelligent instruments, and extended classroom teaching materials. Teachers will have more time to concentrate on their students if they use AI to manage classroom resources and simplify administrative duties. Grading assignments, creating progress reports, and organising classroom resources are just some of the mundane administrative activities that may be automated by AI-powered systems.

2. Contribution:

Instruments are significant devices for learning music. An intelligent instrument is created by incorporating machine learning technology into the instrument. It can store numerous music information of different modern instruments. Particularly when it is utilized in music console, the capability of music console is enhanced fundamentally, which has set off an assortment of showing modes in music educating. It has been demonstrated through relevant experiments that end-to-end NN and score alignment utilized in this paper are able to fairly successfully extract the characteristics of classical music. Two-level grouping model utilized in this paper is more precise in recognizing old style instruments. Objective assessment strategy for articulation showing quality is more goal and exact than P.563 music showing quality assessment. Profound organization can learn significant level elements that are useful for arrangement from information essential highlights; then again, there is just a single playing instrument in monophonic music, utilizing productive preprocessing. Two-level grouping model utilized in this paper is more precise in recognizing old style instruments. Extricating precise symphonious design from spectrogram with mean worth and time-recurrence transformation is simple. Dissecting energy circulation of consonant can really recognize different instruments.

3. Related works:

As of late, the constant improvement of simulated intelligence innovation has made electronic instruments more wise, refined and concentrated to deliver new [7]. Intelligent electronic instrument is capable of not only storing the timbre of any musical instrument, but also realizing effective combination of any and all timbre, allowing for the performance of any and all timbres in accordance with various action instructions. For traditional musical instruments, it is evidently difficult to perform this function. With these benefits, insightful electronic instruments are step by step acquainted into music instructing with guide understudies to learn new savvy electronic instruments. Work [8] created and concentrated on the main murmuring acknowledgment framework in view of profound realizing, which utilizes common string matching-based acknowledgment innovation and utilizations the letters U, D, or S to portray pitch change of sound sign to address murmuring sound sign utilizing a string made out of these three characters, and afterward utilize string matching calculation to compute matching likelihood of melodies in data set. Work [9] applied limited Boltzmann machine to order of music sorts and developed a 5-layer confined Boltzmann machine, yet this strategy has a conspicuous imperfection that it must be utilized in four. Classification of music genres is still more than 50% accurate. With the increment of the kinds of music sorts arranged, the order precision rate will likewise diminish [10]. Work [11] trusts that PCs, similar to individuals, are an actual image framework, and the age of insightful conduct should have the option to change explicit images or certain examples of physical science into different examples and image frameworks. The NNS created by [12] can impersonate the human cerebrum system for information investigation. Its ancestor is counterfeit NNS, and its fundamental component is to attempt to mimic example of data transmission as well as handling between neurons in cerebrum. Work [13] recommend that full of feeling figuring is a calculation that is connected with feelings and starts from or can influence feelings. Human emotions and moods cannot be understood or accommodated by traditional human-computer interaction. This need and limit of the capacity to comprehend and communicate feelings make it hard to anticipate that PCs should have similar knowledge as people from here on out, and it is likewise challenging to see that human-PC collaboration can be really regular and agreeable later on. Work [14] utilizing artificial intelligence innovation introduced nonstop examination of learning results given by the framework to educators and understudies, remembering important reports for execution, understudies' learning status and learning demeanor, and any mix-ups made by understudies in educational experience as well as digressed comprehension of learning content. Work [15] investigated the perspectives of music education major students regarding course curriculum planning. The author of [16] argued that natural language and music share numerous fundamental processing methods. They assessed the viability of developing music majors' independent learning in execution practice through a progression of educational plan changes. Work [17] fought that applying mixed media advanced application research and recovering explicit data from countless music-related information had turned into a test in music data recovery. To predict rise in popularity of music recording, HitMusicNet, a multimode end-to-end Deep Learning (DL) architecture, was proposed. work [18] assisted understudies with learning music by planning another strategy in view of Mixed Reality (MR) and games to animate understudies' advantage in learning. Therefore, understudies' learning inspiration enhanced, and their presentation of music style was generally perceived.

4. Proposed model:

Online video lessons may lack the one-on-one interaction with the learner that is so important while learning an instrument. One-on-one video chats are great for teaching music privately, but cannot be scaled up to teach larger groups because they lack the interactivity necessary for effective learning. This suggests that machine learning models might be used to facilitate interactivity, increase scalability, and streamline the process of creating musical instruction. Supporting MOOCs with the infrastructure shown in Fig. 1 will allow these courses to attract massive numbers of students.

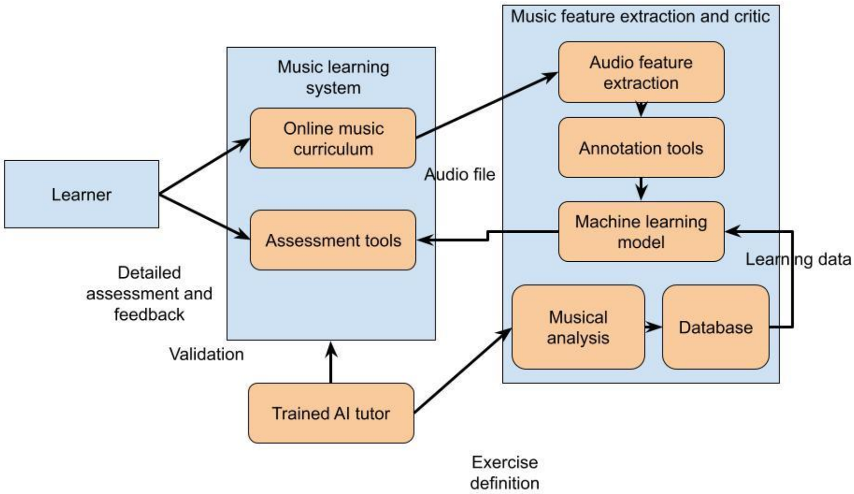

Figure 1. Music education framework

To make mechanized illustration plans in light of class accounts, some essential data, including pitch, harmony, beat, length, cadence, and elements, ought to be recovered from the music record. When training a method to evaluate a student's performance while learning an instrument, these data can serve as useful features. One part of the model is the understudy practice as well as recording connection point that can be effectively custom-made to explicit activities by schooling content architect. Our underlying model gives improved results a straightforward instrument, for example, a woodwind in which the player can play each note in turn. Our underlying tests show that eye to eye conveyance of educator execution followed by a few understudy redundancies can be effectively imitated with such points of interaction. Main meeting of this course will be presented during Summer 2022. Exhibits of points of interaction, results, and perceptions on client experience is imparted to crowd during the meeting. The two-level characterization method utilized in this paper is more exact in distinguishing old style instruments. Objective assessment strategy for articulation showing quality is more goal and exact than P.563 music showing quality assessment. Time-recurrence consecutive geography in recurrence area might work on the impact of the convolution activity. Additionally, replacement of number of model specifications for the first layer greatly reduces the likelihood of overfitting.

4.1 Gaussian kernel filter with Boltzmann spatio convolutional neural network:

A class of convolutional filters known as "Gaussian filters" choose its weights based on the characteristics of a Gaussian function. A extremely effective filter for eliminating noise derived from a normal distribution is the Gaussian smoothing filter. One-dimensional zero-mean Gaussian function is

$$g(x) = e^{-\frac{x^2}{2\sigma^2}} \tag{1}$$

A 3-D isotropic Gaussian has the following shape:

$$G(x, y, z) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} \tag{2}$$

Gaussian separating aims to include this spread as a "point-spread" capability, and it does so with the use of convolution. The cluster was rescaled such that each of its four corners would equal 1, however the resulting integrals are not integers. In the end, the sum of the cover's relative diversity of values is 273. It was filtered utilizing Gaussian Blur built-in function from ImageJ after noise was added using the Add noise built-in function. The "Gaussian Blur" results from the convolution of an image's pixels and a kernel that can be characterised by a Gaussian function. In discrete example, the convolution is given by

$$\mathrm{f} * \mathrm{g}[\mathrm{n}] \stackrel{\mathrm{def}}{=} \sum_{m=-\infty}^{\infty} \mathrm{f}[\mathrm{m}] \cdot \mathrm{g}[\mathrm{n}-\mathrm{m}] \tag{3}$$

$$f(x, y) = A \cdot e^{-\left(\frac{(x-x_0)^2}{2\sigma x^2} + \frac{(x-y_0)^2}{2\sigma y^2}\right)} \tag{4}$$

A low-pass filter, the convolution with the Gaussian kernel smoothes the edges of the image. The variance of the Gaussian distribution serves as the parameter that we employ to define the filter because it significantly influences the filtering results. As the predicted effects of the Gaussian blur are the focus of the current work, we will base our conclusions on a comparison of filtered image with original image utilizing quality factor provided by

$$Q(f, g) = \frac{\sigma_{f,g}}{\sigma_g^2 \sigma_f^2} \cdot 2 \cdot \frac{Tg}{f^2+g^2} \cdot 2 \cdot \frac{\sigma_g^2 \cdot \sigma_f^2}{\sigma_f^2+\sigma_g^2} \tag{5}$$

Signal-to-Noise Ratio, or SNR, is the measurement we utilised to assess the noise level in the photos.

$$SNR = 20.\log\left(\frac{\sigma_{signal}}{\sigma_{noise}}\right) \tag{6}$$

4.2 Symphonic Intelligence: Revolutionizing Music Education with AI and Machine Learning

By using this filtering technique, image data will be improved by reducing background noise and distortions as well as background noise. For processing and analysis, it improves image features. The RGB-formatted photos are scaled down to conventional dimensions. One auto-encoder is trained during the training procedure. Figure 2 shows testing and training framework for stack auto-encoder-based objective evaluation technique of speech teaching quality.

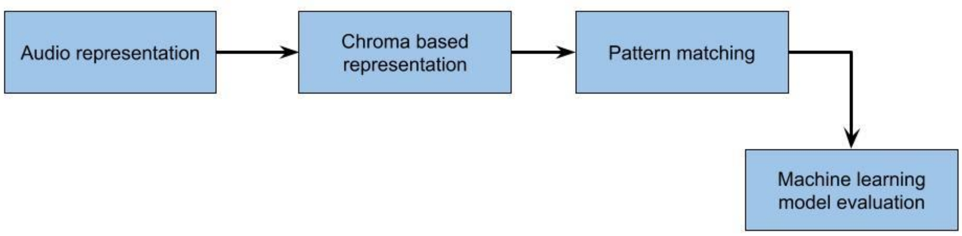

Figure-2 Proposed music playlist and students interaction analysis

Assume that MRF that $\hat{x}_t$ represents is realised in Xt. The posterior probability distribution presented in equations (7) can be maximised to determine the label field for any feature extraction problem.

$$E(|y_t \cap A_1||\gamma_{t+\Delta t} \cap A_2|) = \int_{A_3}\int_{A_t} k_{t,\perp}(x_1,x_2)dx_2 \,\mathrm{d}x_1 + \int_{A_1\cap A_s} k_{t,L}(x)\mathrm{d}x,\ \frac{\mathrm{d}}{\mathrm{d}t}k_t(\eta) = (L^{\Delta}k_t)(\eta), \tag{7}$$

a differential equation with the form eq. (8)

$$\frac{\mathrm{d}}{\mathrm{d}t}k_t^{(1)}(x) = -mk_2^{(\mathrm{d})}(x) - \int_{R^t} a^-(x-y)k_t^{(2)}(x,y)dy + \int_{R^t} a^+(x-y)k_t^{(1)}(y)dy. \tag{8}$$

RBMs are energy-based techniques that aim to learn as much as possible. The energy is provided by eq. (9) when its hidden variables h and visible variables v are specified simultaneously.

$$E(\mathbf{v},\mathbf{h};\theta) = -\sum_{ij}\mathbf{W}_{ij}\mathbf{v}_i\mathbf{h}_j - \sum_i \mathbf{b}_i\mathbf{v}_i - \sum_j \mathbf{a}_j\mathbf{h}_j \tag{9}$$

$$P_\theta(\mathbf{v},\mathbf{h}) = \frac{1}{Z(\theta)}\exp\left(-E(\mathbf{v},\mathbf{h};\theta)\right). \tag{10}$$

$$P_\theta(\mathbf{v},\mathbf{h}) = \frac{1}{Z(\theta)}\exp\left(\sum_{i=1}^{D}\sum_{j=1}^{F}\mathbf{W}_{ij}\mathbf{v}_i\mathbf{h}_j + \sum_{i=1}^{D}\mathbf{v}_i\mathbf{b}_i + \sum_{j=1}^{F}\mathbf{h}_j\mathbf{a}_j\right) \tag{11}$$

$$P_\theta(\mathbf{v}) = \frac{1}{Z(\theta)}\sum_{\mathbf{h}}\exp\left[\mathbf{v}^T\mathbf{W}\mathbf{h} + \mathbf{a}^T\mathbf{h} + \mathbf{b}^T\mathbf{v}\right] \tag{12}$$

$$L(\theta) = \frac{1}{N}\sum_{n=1}^{N}\log\ P_\theta\left(\mathbf{v}^{(n)}\right)$$

$$\frac{\partial L(\theta)}{\partial \mathbf{W}_{ij}} = \frac{1}{N}\sum_{n=1}^{N}\frac{\partial}{\partial W_{ij}}\log\left(\sum_{\mathbf{h}}\exp\left[\mathbf{v}^{(n)T}\mathbf{W}\mathbf{h} + \mathbf{a}^T\mathbf{h} + \mathbf{b}^T\mathbf{v}^{(n)}\right]\right) - \frac{\partial}{\partial W_{ij}}\log\ Z(\theta) = E_{P_{\mathrm{dax}}}\left[\mathbf{v}_i\mathbf{h}_j\right] - E_{P_o}\left[\mathbf{v}_i\mathbf{h}_j\right] \tag{13}$$

The first half of the formula is easy to calculate. The average vi and hj values across all datasets are calculated. Formula's second half (38) equals

$$\sum_{v,h}\mathbf{v}_i\mathbf{h}_j P_\theta(\mathbf{v},\mathbf{h}) \tag{14}$$

$$\Delta\mathbf{a}_i = v_i^{(0)} - v_i^{(k)}$$

$$\Delta\mathbf{b}_i = P\left(\mathbf{h}_j = 1 \mid \mathbf{v}^{(0)}\right) - P\left(\mathbf{h}_j = 1 \mid \mathbf{v}^{(k)}\right)$$

$$\Delta\mathbf{W}_{ij} = P\left(\mathbf{h}_j = 1 \mid \mathbf{v}^{(0)}\right)\mathbf{v}_i^{(0)} - P\left(\mathbf{h}_j = 1 \mid \mathbf{v}^{(k)}\right)\mathbf{v}_i^{(k)} \tag{39}$$

Eqn (15) offers final specification update equation.

$$\mathbf{a}_i = \mathbf{a}_i + \Delta\mathbf{a}_i,$$

$$\mathbf{b}_j = \mathbf{b}_j + \Delta\mathbf{b}_j,$$

$$\mathbf{W}_{ij} = \mathbf{W}_{ij} + \Delta\mathbf{W}_{ij} \tag{15}$$

$$P\left(V,h^{(i)},h^{(2)}\right) = \frac{1}{Z(\theta)}\exp - E\left(V,h^{(i)},h^{(i)};\theta\right)$$

$$P\left(V,h^{(1)},h^{(2)}\right) = -V^T W^{(1)} h^{(i)} - V^T W^{(2)} h^{(2)} + b \tag{16}$$

5. Results and discussion:

The dataset and trial system for single-name kind arrangement from sound are depicted in this part. More specifically, we experimented with a variety of data modalities to classify track genres: just sound, just collection cover fine art, and both. Dataset portrayal:

MSD-I dataset: Million Tune Dataset contains metadata and pre-registered sound qualities for 1 million melodies. This dataset was released alongside a dataset that included annotations for 15 top-level genres and a single label for each song (Schreiber, 2015). Last assortment contains 30,713 MSD accounts and their comparing collection cover photographs, each marked with an unmistakable type characterization from one of

15 classes. In light of a starter assessment of the photographs, we found that this assortment of music is connected to 16,753 collections, with a normal of 1.8 tunes per collection.

GTZAN Dataset : A gathering of ten classifications, each with 100 sound accounts, each enduring 30 seconds. In AI research (MGR), the GTZAN dataset is most often utilized public dataset for music classification acknowledgment assessment. The records were acquired in 2000 and 2001 from a number of sources, including individual Compact Discs, radio, and receiver accounts, to handle a range of recording settings.

Table 2: Analysis of Proposed technique with existing techniques

Datasets	**Techniques**	**Random Accuracy**	**Precision**	**Recall**	**AUC**
MSD-I dataset	**SVM**	93	92	82	50
	HitMusicNet	96	95	86	45
	GKF_BSCNN	97	94	89	40
GTZAN Dataset	**SVM**	93	92	84	54
	HitMusicNet	96	95	83	55
	GKF_BSCNN	98	98	90	36

Comparative investigation of music genre classification for multiple datasets is shown in table 1 above. Based on the classification of music education, a comparison has been done to determine the type of music education from the input dataset. In examining the input music education enhancement dataset, the suggested GKF_BSCNN based classification outperformed all other methods in terms of output quality and precision.

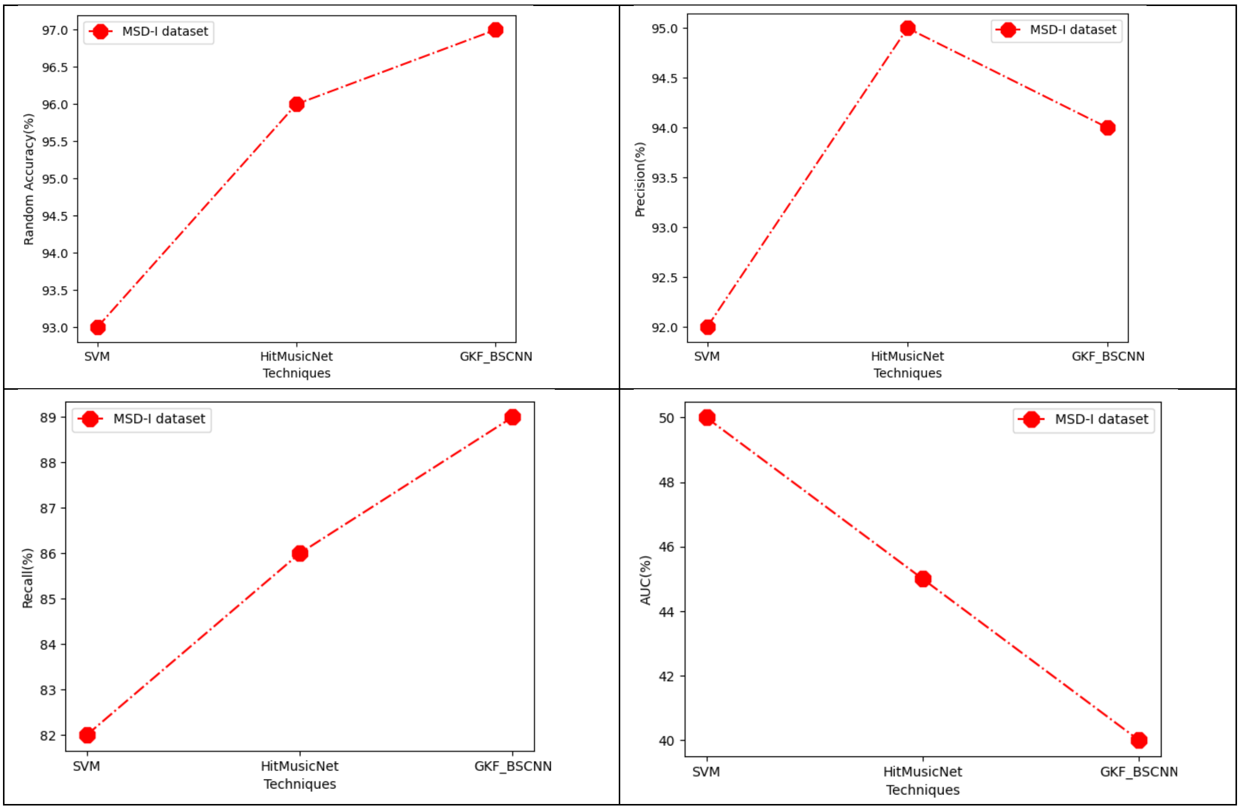

Figure 3: Comparison of parameters for MSD-I dataset (a) random Accuracy, (b) Precision, (c) Recall, (d) AUC

Figure 3 above compares metrics in terms of recall, accuracy, precision, and AUC. The MSD-I dataset used to create the aforementioned graph has random accuracy, precision, recall, and AUC values of 97%, 94%, 89%, and 40%, respectively. From this comparison, it can be shown that the suggested technique produced the best results for the MSD-I dataset's parameters. In most cases, data collected is not labelled, but it may be separated into groups based on how similar the data inside them is. Furthermore, the data shared across clusters is innovative, and the majority of clustering computations are related to parcel-based clustering, thickness-based clustering from tables, thickness-based bunching, and occurrence clustering. The methods used in a bunch inquiry are well-known for being useful, simple, and very effective. It has found widespread use, particularly in areas such as archive union.

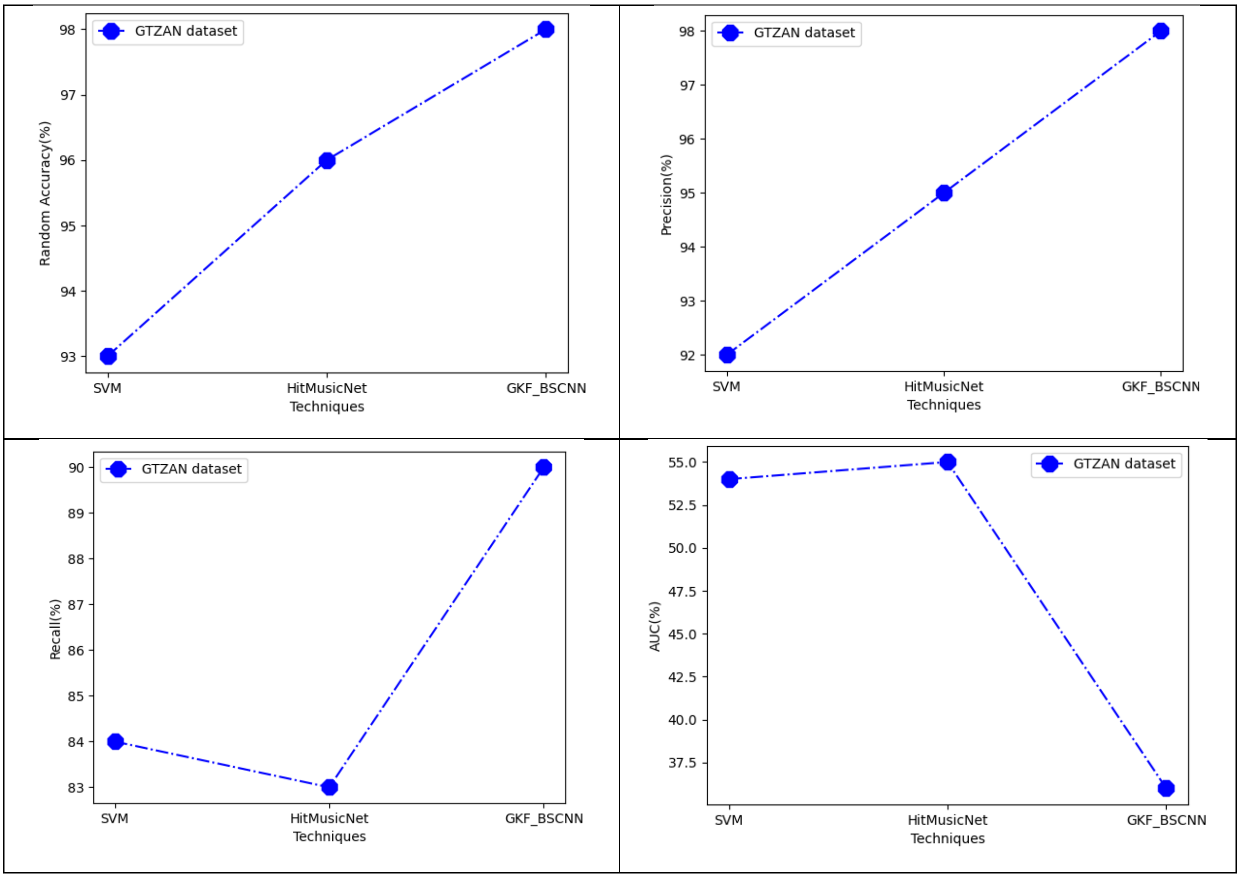

Figure 4: Comparison of parameters for GTZAN Dataset (a) Accuracy, (b) Precision, (c) Recall, (d) AUC

Figure 4 above compares metrics in terms of recall, accuracy, precision, and AUC. The accuracy of the proposed method, which produces better results in music genre categorization than the existing method, is shown in the graph above, which was produced using the GTZAN Dataset. Its accuracy measures 98%, its precision 93%, its recall 89%, and its AUC 36%. Utilizing the two-layer organization can recognize the greater part of the pitches of explicit frequencies that should be distinguished; in any case, those that don't exist are additionally distinguished, prompting over-fitting. Choices, similar to two-layer organizations, channels, and start to finish brain organizations, two-level grouping models, objective assessment strategies, should in any case be thought of. We see execution of above calculation in traditional music training, which includes ID of old style instruments, component extraction and recognizable proof of traditional music, and quality assessment of traditional music schooling. Through research, center arrangement and start to finish brain networks are nearly fruitful in extricating the highlights addressed by the circle focuses. We ought to take note of that the even hub compares to the emotional MOS score. By surveying the relationship between result results as well as emotional assessments, we can measure viability of a music quality assessment framework. A corner to corner line in a scatterplot delineates an ideal match between the goal and emotional evaluations.

6. Conclusion:

This research proposes novel technique in music education enhancement based on students interest using Gaussian kernel filter and Boltzmann spatio convolutional neural network (GKF_BSCNN). Using machine learning techniques, we classified the data from online course evaluations and examined the significance of student evaluation characteristics. +e various components of school courses and instructor viability are estimated by the web-based course assessment. Experiments have demonstrated that the machine learning approach accurately categorizes teachers' "satisfaction" and "dissatisfaction." +e results showed that presentation signs of the relative multitude of referenced procedures were around 90%. Among them, understudies' very own advantages and information acquired are the fundamental reason for understudies to assess courses. +ese results can be utilized to work on the foundation and improvement of online assessment markers, in order to set more sensible and logical assessment pointers. +e investigation consequence of the component significance of the classifier shows that there are numerous potential upgrades in the plan of the estimation pointers utilized in the instructor execution assessment, and it gives a specific reference to the administration division to further develop the assessment markers. By and large talking, during the time spent showing work, common schools and colleges change the showing assessment arrangement of musicology, which can do the showing work of musicology in a logical and ordinary way, and coordinate the principal errands of musicology showing in schools and colleges. Finished without a hitch, so during the time spent coordinating homeroom showing exercises, music educators ought to regard the understudies' characters, make sensible references to the showing assessment framework, and give full play to the understudies' learning excitement, so through study hall showing exercises, each understudy can get a specific measure of music information. Along these lines, countless music gifts in accordance with social advancement can be developed and the advancement of music training can be advanced. Proposed technique attained random accuracy of 97%, precision of 94%, recall of 89%, AUC of 40%.

Declarations

Availability of data and Materials: All the data’s available in the manuscript

Ethical Approval: This article does not contain any studies with animals performed by any of the authors.

Conflicts of Interest: The authors declare no conflict of interest

Funding: No Funding

Reference:

[1] Hong Yun, Z., Alshehri, Y., Alnazzawi, N., Ullah, I., Noor, S., & Gohar, N. (2022). A decision-support system for assessing the function of machine learning and artificial intelligence in music education for network games. *Soft Computing*, *26*(20), 11063-11075.

[2] Wei, J., Karuppiah, M., & Prathik, A. (2022). College music education and teaching based on AI techniques. *Computers and Electrical Engineering*, *100*, 107851.

[3] Yu, X., Ma, N., Zheng, L., Wang, L., & Wang, K. (2023). Developments and applications of artificial intelligence in music education. *Technologies*, *11*(2), 42.

[4] Xia, Y., & Xu, F. (2022). Design and application of machine learning-based evaluation for university music teaching. *Mathematical Problems in Engineering*, *2022*, 1-10.

[5] Shuo, W., & Ming, M. (2022). Exploring online intelligent teaching method with machine learning and SVM algorithm. *Neural Computing and Applications*, 1-14.

[6] Kim, H., Ramoneda, P., Miron, M., & Serra, X. (2022). An overview of automatic piano performance assessment within the music education context.

[7] Wang, D. (2023). Application of Machine Learning Technology in Classical Music Education. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, *18*(2), 1-15.

[8] Zhaoran, S. (2021). Wireless processor application in home music teaching based on machine learning. *Microprocessors and Microsystems*, *80*, 103359.

[9] Zhu, Z., Xu, Z., & Liu, J. (2022). Application of lightweight deep learning model in vocal music education in higher institutions. *Computational Intelligence and Neuroscience*, *2022*.

[10] Jamshidi, F., Marghitu, D., & Chapman, R. (2021, July). Developing an online music teaching and practicing platform via machine learning: a review paper. In *International conference on human-computer interaction* (pp. 95-108). Cham: Springer International Publishing.

[11] Carey, G., Coutts, L., Grant, C., Harrison, S., & Dwyer, R. (2018). Enhancing learning and teaching in the tertiary music studio through reflection and collaboration. *Music Education Research*, *20*(4), 399-411.

[12] Chen, J. C. W. (2023). Mobile learning as deep learning: Content analysis of in-service primary school music teachers’ lesson plans in mobile music creation. *International Journal of Music Education*, 02557614231191520.

[13] Guo, T. (2022). Application of internet of things technology in vocal music teaching recording equipment assisted by machine learning. *Wireless Communications and Mobile Computing*, *2022*.

[14] Zhu, Z., Xu, Z., & Liu, J. (2023). Flipped classroom supported by music combined with deep learning applied in physical education. *Applied Soft Computing*, *137*, 110039.

[15] Liu, X. (2022). Analysis of psychological characteristics and emotional expression based on deep learning in higher vocational music education. *Frontiers in Psychology*, *13*, 981738.
[16] Clements, A. (2018). A postdigital future for music education: Definitions, implications, and questions. *Action, Criticism & Theory for Music Education*, *17*(1).
[17] Li, Y. (2022). Machine Learning-Based Aesthetic Music Education Informatics Assessment Method. *Wireless Communications and Mobile Computing*, *2022*.
[18] Li, H. (2021). Piano education of children using musical instrument recognition and deep learning technologies under the educational psychology. *Frontiers in Psychology*, *12*, 705116.