**Educational Administration: Theory and Practice**
2024, 30(5), 1683-1692
ISSN: 2148-2403
https://kuey.net/
**Research Article**

# Exploration Of The Correlation Between Students' Musical Interest And Academic Performance Based On Deep Learning

Yan Yuan[1*]

[1*]International College, Krirk University, Bangkok 10220, Thailand

***Corresponding author:** Yan Yuan
E-mail: 18989240568@163.com

**Citation:** Yan Yuan, (2024), Exploration Of The Correlation Between Students' Musical Interest And Academic Performance Based On Deep Learning, *Educational Administration: Theory and Practice*, *30*(5), 1683-1692
Doi: 10.53555/kuey.v30i5.3150

**ARTICLE INFO**

**ABSTRACT**

This research delves into the intricate relationship between students' musical interests and their academic performance, leveraging deep learning methodologies for a nuanced analysis. By employing advanced computational techniques, the study aims to uncover patterns, correlations, and potential influences that exist between musical engagement and academic achievement. Key objectives include utilizing deep learning algorithms to analyze large datasets of students' musical interests and academic outcomes, identifying patterns that indicate a correlation, and exploring potential causal factors. The research aims to provide valuable insights into how fostering musical interests may positively impact academic performance, contributing to the broader discourse on the intersections between arts education and scholastic achievement. The findings of this exploration are anticipated to offer practical applications for educators, policymakers, and researchers seeking to understand and optimize the relationship between students' extracurricular interests, such as music, and their overall academic success.

**Keywords:** students interest, academic performance, music interest, deep learning techniques, kernel regressive model

## 1. Introduction:

Russian universities have a unique chance to address several key issues through active integration into the digital environment and the use of digital technologies: Data science techniques can be used to: (1) modernise or design algorithms for databases to be filled and then analytically processed in the primary areas of activity; (2) gather and analyse data on the current status of educational activity in a timely manner (descriptive analytics); and (3) use data mining techniques to create predictive analytics as a separate direction. Currently, classic static data processing methods serve as the foundation for academic (educational) analytics in Russian universities. However, during the past ten years, worldwide practice has accumulated a great deal of expertise in implementing the findings of academic (educational) analytics through the use of multidimensional databases (including retrospective) [1]. It is quite significant that various analysts recognize scholastic investigation and information mining. The substance of the distinctions is that scholastic investigation is assessed as a spellbinding way to deal with information handling that permits tackling functional issues of dealing with the schooling system, and information mining is pointed toward uncovering stowed away examples that should be considered while pursuing key choices in view of information [2]. Countless examination works in field of information mining are focused on utilization of bunching, characterization, perception techniques. One of the promising, effectively creating regions is the advancement of prescient models of understudies' scholarly achievement in view of machine learning [3]. Prescient models foresee the understudy's instructive course some time before he graduates. A significant job in the development of such models is played by indicators, that is to say, key guaging boundaries. As of late, research interest in the improvement of prescient models in light of the examination of "computerized follows" has escalated. quantitative information on the movement of understudies in PC instructive conditions. Machine learning could be a specialization under the tremendous computer based intelligence. Machine learning makes progress toward understanding the intricacy of arranged kinds of gathered information and distinguishing the right model for the information by attempting a few models. This can be really systemized with simpler translation and use by individuals [4]. Machine learning exists in the designing science field however is not the same as fundamental processing calculations which are utilized for critical thinking. Inside the course of machine learning, the calculations are planned in a way which permits the framework or PC to handle the info data, make preparing sets and produce the ideal reach determined yield utilizing factual assessment. The greatest resource for many universities is an understudie. Colleges and students have a critical role in producing graduates with outstanding academic accomplishments through their intellectual presentations. The degree to which students have met their learning objectives through evaluation, appraisal, and other forms of estimation is known as scholarly execution accomplishment. However, the degree of instructional execution accomplishment varies for different reasons, therefore different understudies may have varying degrees of accomplishment. In every one of the basic pieces of an understudy's private and expert improvement is execution assessment. Execution assessments underline's major areas of strength for understudies and their strong point. This goes about as a crucial device in enlarging their assets and recognizing regions that require improvement as objectives. By being able to investigate the presentation of their understudies, educators can redirect their regard for the required regions, exhortation and guide the researchers along the appropriate way and recognize and compensate their accomplishments [5]. Understudies' presentation is a term utilized for estimating not exclusively understudies' accomplishments yet additionally the nature of instructive foundations. Some writers characterise student academic performance as solely likelihood of achieving a long-term objective, like graduation or potential for future employment opportunities, while others define it as the value derived from comparing a specific student learning assessment with the study curriculum, grade point average (GPA), or final grades. In order to address academic underachievement, higher university dropout rates, graduation delays, other issues, it is crucial for educational institutions to analyse and forecast student performance. This can be done by identifying students with low academic achievement early in their studies. Understanding the possibility of leveraging acquired data to enhance individual and institutional learning efficacy and academic success is crucial for educational institutions. The field of learning analytics encompasses the collection and analysis of student data, the identification of optimal student/teacher performance, and a deeper knowledge of the learning environment via examination and analysis [6]. To better understand and optimise learning and its settings, learning analytics involves gathering, measuring, reporting data on students and their circumstances. It also covers how the organisations are creating new plans. Other components of learning analytics include predicting students' academic progress, spotting patterns in their system access and navigational activities, and identifying students who could be at risk of failing. To evaluate possible student behaviour, digital data left by web-based education systems such as massive open online courses (MOOCs), student information systems (SIS), intelligent teaching systems (ITS), and learning management systems (LMS) may be evaluated. These data may be utilized, with help of EDM technique, to assess actions of both successful as well as at-risk students, to create remediation plans based on academic performance of the students, ultimately to support teachers in creating pedagogical approaches [7].

## 2. Contribution:

Instruments are critical gadgets for learning music. A shrewd instrument is made by integrating machine learning innovation into the instrument. It can store various music data of various present day instruments. Especially when it is utilized in music console, ability of music console is enhanced essentially, which has set off an arrangement of showing modes in music teaching. It has been shown through pertinent analyses that the start to finish brain organization and score arrangement used in this paper can reasonably effectively extricate the attributes of old style music. This article employs a two-level gathering model that is more accurate in identifying vintage instruments. Compared to P.563 music showing quality appraisal, the objective evaluation technique for explanation showing quality is more objective and definitive. With effective preprocessing, the important association can learn crucial level components that are useful for planning from the fundamental properties of the data; in monophonic music, of course, there is only one playing instrument. Such a neural network as we utilised in this article can achieve long-term improvement in erasing the qualities of old music after extensive preliminary work. This article employs a two-level gathering model that is more accurate in identifying vintage instruments. Removing exact symphonious plan from spectrogram with mean worth as well as time-repeat change is basic. Analyzing energy dissemination of consonant can truly perceive various instruments.

## 3. Related works:

Education institutions now have access to interactive learning resources including virtual reality, game-based simulation apps, and e-learning platforms thanks to learning management systems. Researchers are now able to gather and examine student data thanks to these platforms [8]. Authors [9] utilized SVM, Decision Tree, Neural Network classification machine algorithms to predict students' academic achievement based on their online usage habits. Findings demonstrated that, with an accuracy of 71%–76%, student internet usage behaviours accurately predict academic success; however, authors only took accuracy into account as a performance criterion. The authors of [10] paper presented a method that classifies students into excellent or poor groups using ML methods trained to predict students' academic achievement. K-nearest neighbour as well as decision tree classifier were used to create method, which was trained using data collected from a university source. Decision tree classifier's accuracy was 94.44%, according to the results, although the author only took accuracy into account when evaluating its effectiveness. Similarly, to predict students' academic achievement, the authors [11] created a classification machine learning model that uses logistic regression and SVM classifiers. The model classified student academic performance as poor, average, or good by extracting characteristics from preprocessed dataset that was received from an online learning platform. The results show that the SVM outperformed the logistic regression, with an accuracy rate of 79%. Method performance was assessed by authors using confusion box measurements, which comprised f1-score, accuracy, recall, and precision. Utilizing a dataset of 887 instances of 19 first-year student characteristics, authors [12] employed the RF classifier as well as Ensemble learners classification ML method to predict academic achievement. With a 93% accuracy rate, RF classifier outperformed the other methods. Confusion box measures were utilized to evaluate the model's performance in order to determine recall, precision, and f1-score. Field of machine learning (ML) in education is still in its infancy, and there are still a lot of issues that need to be resolved, including poor prediction accuracy, overfitting, underfitting, and model deployment. Thus, by comparing many ML models to DL methods, our suggested technique provides an accurate and efficient academic performance for students. A very accurate classification method for predicting student performance is support vector machines [13]. Accuracy of Naïve Bayes Simple, Multilayer Perceptron, SMO, J48, and REP Tree approaches for forecasting student performance was investigated in work [14]. The outcomes demonstrated that, while SMO is a competitive method, the Multilayer Perceptron approach was the best suitable. Few and infrequent EDM research have been conducted in Vietnam. One such study, [15], examined accuracy of decision trees as well as Bayes networks in predicting students' learning outcomes between Can Tho University and the Asian Institute of Technology in Thailand. Even though the dataset variety varied in this investigation, the approaches used produced comparable findings [16]. Creator [17] involved multivariate relapse to make forecasts of learning effectiveness in math or math-related subjects in view of learning techniques and segment factors. The outcome exhibited that surface methodology and affirmation mark impacted on scholastic result; yet orientation, parental training and math interest in secondary school didn't. Work [18] estimated understudies' learning result in view of multistrategy relapse approach and adds data about subject-related abilities. The exploratory outcome showed that the relapse based strategy gives preferable execution over the recommender framework - based one. The sluggish students were distinguished and shown by [19] by using characterization put together calculations based with respect to prescient displaying. The understudy socioeconomics connected with their prosperity were found by [20] using choice trees. The work by [21] performed early expectation of understudy's presentation utilizing the ML strategies like brain organizations, direct separate investigation, and choice trees. Instructive information digging was used for removing information from the alumni understudy's information that was gathered between years 1993 and 2007 from school of Science and Innovation. Dropout probability was anticipated by [22] using the choice tree classifiers. Point was to lessen dropout recurrence among graduates as well as increment understudy maintenance. In [23], the creators removed information on understudy's exhibition from result sheets utilizing choice trees. This made it easier to identify dropouts and gave the tutor the opportunity to provide the pupils guidance. By removing the knowledge discovery from the end-of-semester grades, Cortez and Silva [24] used four classification methods—DT, RF, ANN, and SVM—to enhance student performance as well as assist in achieving objective.

## 4. System model:

Online video lectures might not provide a real-time connection with the learner, which is essential for learning music. Nevertheless, since they don't offer an interactive learning environment, one-on-one video chats cannot be helpful in music training when used for larger audiences. Therefore, the creation of music classes may be made more engaging, scalable, and automated with the use of machine learning models. Fig. 1 illustrates a framework for assisting MOOCs that will enable them to reach large audiences.

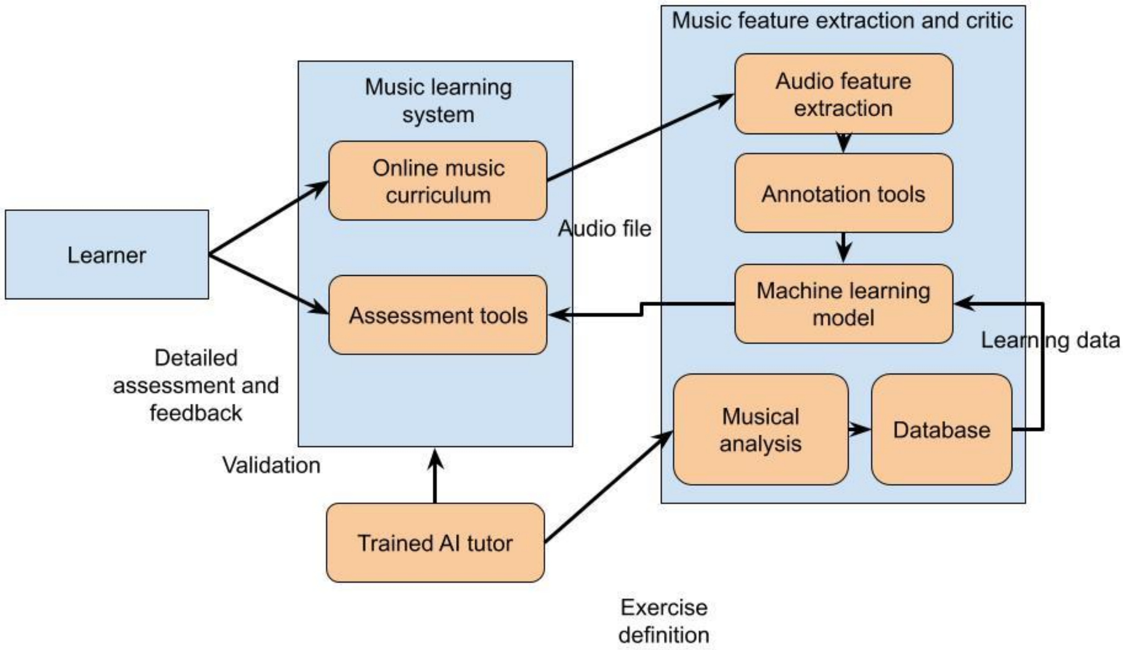


**Figure. 1. Music education framework**

The foundation of the framework is as follows: (1) Learner utilises interfaces for practice and learning; (2) Trained AI Tutor uses the Music Learning System (MLS) to give a music exercise. The student then uploads the audio recording of the activity to the MLS; (3) MLS forwards the sound file to Music Component Extraction and Pundit, where it is analysed and then shared with the ready AI mentor for evaluation. It is necessary to retrieve some crucial information from the music record, such as pitch, harmony, rhythm, duration, cadence, and components, in order to create automated illustration plans based on class accounts. These data might be helpful aspects when developing a model to assess a student's performance while learning an instrument. The understudy practice and recording connection point is one component of the concept that the schooling content architect (music instructor) may successfully tailor to specified activities. Our underlying approach produces better results when applied to a simple woodwind instrument, where the musician may play each note individually. Our fundamental examinations demonstrate that with such points of involvement, it is possible to effectively mimic the eye-to-eye conveyance of educator execution trailed by a few understudy redundancies. This course's main meeting is scheduled for the summer of 2022. During the meeting, the audience will be given access to exhibits showcasing points of engagement, results, and impressions regarding customer experience.

## Probabilistic reinforcement kernel regressive model with spatio convolutional neural networks:

This article uses a two-level characterisation approach that is more accurate in identifying vintage instruments. Compared to P.563 music showing quality calculation, objective assessment technique for articulation showing quality is more precise and goal-oriented. The impact of the convolution activity may be influenced by the time-recurrence sequential geography in the recurrence area. Additionally, the chance of overfitting is significantly decreased by changing the quantity of model parameters for the first layer. We aim to forecast the pitch combinations that correspond to the sung notes in audio frame X, given frame. In binary label vector form, the pitch combination is represented as y ∈ {, 0 1} 128. Its 128 dimensions match 128 pitch frequencies, it will have nth component in vector y if nth tone, yn = 1, is present in frame. We first train a multiple linear regression using a multilayer convolutional network to learn a feature map f X P θ: →, and then we apply a cross-discipline loss function to optimise parameters of regression to predict a given f. label vector j. A single sample is separated into $M + 1$ autonomous blocks, $x0$, $x1$, $\cdots$, $xM$, and it is associated with CPU; $[x1m, \cdots, xNm]$ refers to operation assignment of 1- N to CAP $m$, in order to reduce the duplication of created instances. Presume $\mathcal{G}$ is an alternative set for storing the instances to satisfy the constraints for every iteration. It is derived using the Bernoulli distribution of parameters for all $ul$ in $xg$. The residual block indicator $um$ in $M \setminus \mathcal{G}$ is adjusted according to $xg$, so that if $xig = 1$, then $uim = 0$ for $m \in M \setminus \mathcal{G}$. Given that the useful sample has been generated, it may be assumed that the cardinality of is equal to $M$, as indicated by $|\mathcal{G}|$. The cardinality of $\mathcal{T}$, denoted as $|\mathcal{T}|$, comes after sampling, and reached $S$. The valuable samples are gathered in $\mathcal{T}$. It shows that smaller distance from $(x)$ and $P(x)$, the smaller the minimum $(q, p)$. It indicates that by eqn (1)

$$\min H(q,p) = max \sum q(x) \ln p(x) = \max \frac{1}{S} \sum \ln p(x,u) \qquad (1)$$

If probability of an autonomous outcome in a collection of samples is $1/S$, where $S$ represents cardinality of a set, and $q(x) = 1\ S$. The goal is to find the best optimal indicator $u$ for lowering $(q, P)$ based on issues in (10). A candidate is selected in the kth iteration from a sequence of random examples, with the probability (k, p) determining the retrieval. It is established what sample could be obtained by adaptive sampling. The objective $\{\Psi(xs\ )\}s$=1 $S$ is assessed and arranged as follows: $\Psi(x[1]) \leq \Psi(x[2]) \leq \cdots \leq \Psi(x[S])$. Then, elites are selected from among samples like $x$ [1], [2], ⋯, and $x[elite]$, which provides the lower objective. Next, the best indicator $u$ for strategy $x$ is calculated as follows:$u *= \arg\max \frac{1}{s} \sum \ln p(x, u)$

By applying (9) and (11) by encouraging $\frac{\partial H(q,p)}{\partial ul} = 0$ saddle point c by eqn (2)

$u_l = 1/S_{e1lite} \sum_{s=1}^{Se1lite} x_l^{[s]}$ (2)

The CE-related metric, generated from the proposed model, is used to enlarge the likelihood. In the $(t + 1)$℗ iteration, the function $u$ $(t$+1) is upgraded based on $u$ ∗, which is regulated using (11) and (12). In the final iteration, $u$ $(t)$ is learned by considering randomness of sampling and a minimal count of samples. It's pertinent by eqn (3)

$u^{(t+1)} = au * +(1 - \alpha)u^{(t)}$ (3)

where the learning value is implied by $\alpha \in [0, 1]$. In general, the iterations of a CE-aided model converge to the most optimally solved problems by eqn (4)

$\Delta \mathbf{w}_i = \eta \left( \text{ true }_j - \text{ pred }_j \right)$ (4)

where truej is real class label, predj is anticipated class label, and η is learning rate. $\sum_n \tilde{\pi}(a \mid s) A_\pi(s, \bar{a}) \geq 0$ is ensured to build approach execution η, or leave it consistent for situation that normal benefit is zero all over place. This infers exemplary outcome that update performed by precise approach emphasis, which utilizes deterministic arrangement $\tilde{\pi}(s) = \arg\max_a A_\pi(s, a)$ works on strategy on the off chance that there is something like one state-activity pair with a positive benefit esteem as well as nonzero state appearance likelihood, generally calculation has joined to ideal arrangement. Nonetheless, in rough setting, it will commonly be undeniable, because of assessment and estimation mistake, that there will be a few states s for which the normal benefit is negative, or at least, $\sum_a \tilde{\pi}(a \mid s) A_\pi(s, a) < 0$ by eqn (5)

$L_\pi(\tilde{\pi}) = \eta(\pi) + \sum_s\ \rho_\pi(s) \sum_a\ \tilde{\pi}(a \mid s) A_\pi(s, a)$ (5)

Lπ utilizes the appearance recurrence ρπ as opposed to ρπ˜, disregarding changes in state appearance thickness because of changes in the strategy. Nonetheless, in the event that we have a defined strategy πθ, where πθ(a|s) is a differentiable capability of boundary vector θ, then, at that point, Lπ matches η to initially arrange. For any boundary esteem θ0 by eq. (6),

$$L_{\pi_{\theta_0}}(\pi_{\theta_0}) = \eta(\pi_{\theta_0})$$

$\nabla_\theta L_{\pi_{\theta_0}}(\pi_\theta)\Big|_{\theta=\theta_0} = \nabla_\theta \eta(\pi_\theta)|_{\theta=\theta_0}$ (6)

$$\max_\theta L_{\theta_{\text{dd}}, \mu}(\theta) \text{s.t. } D_{\text{KL}}^{\text{max, sqrt}}(\mu, \theta_{\text{old}}) D_{\text{KL}}^{\text{max, sqrt}}(\theta_{\text{old}}, \theta) + D_{\text{KL}}^{\max}(\theta_{\text{old}}, \theta) \leq \delta$$

where δ denotes the trust region restriction bound. The aforementioned problem is unsolvable because constraints are bounded at every point in state space. We utilise average KL divergence as a constraint: $D_{\text{KL}}^{\text{max, sqrt}}(\mu, \theta_{\text{old}})$ Thus, the trust region constraint optimisation problem in (7) is roughly represented by

$\max_\theta L_{\theta_{\text{dd}}, \mu}(\theta) \text{s.t. } D_{\text{KL}}^{\text{max, sqrt}}(\mu, \theta_{\text{old}})$ (7)

We then consider finite samples based on derived optimisation issue under parameterized policies in (23). We enlarge L_(θ_"dd ",μ) (θ) to employ finite samples to evaluate objective in restricted optimisation issue. Definition of a Markov process is where we begin. If and only if probability of going from one state to next, St+1, depends only on current state St and not on prior states S1, S2, · · ·, St−1, then a series of states is Markov. In other words, for every t by eqn (8)

$\mathbb{P}[S_{t+1} \mid S_t] = \mathbb{P}[S_{t+1} \mid S_1, S_2, \cdots, S_t]$ (8)

In reinforcement learning, we frequently discuss time-homogeneous Markov chains, where transition probability is independent of t by eqn (9)

$$\mathbb{P}[S_{t+1} = s' \mid S_t = s] = \mathbb{P}[S_t = s' \mid S_{t-1} = s]$$
$$\mathcal{P} = p_1 p_2 \dots p_n, \text{ where } p_i \in \{H, P\}, \forall 1 \leq i \leq n$$
$$\mathcal{B} = \{\mathcal{P} = p_1 p_2 \dots p_n \mid p_i \in \{H, P\}, \forall 1 \leq i \leq n, n \in \mathcal{N}\}$$
$$\mathcal{G} = \{G = (x_i, y_i) \mid x_i, y_i \in \Re, 1 \leq i \leq n\}$$
$$C: \mathcal{B} \to \mathcal{G}$$
$$\mathcal{P} = p_1 p_2 \dots p_n \mapsto \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$$
$$\forall 1 \leq i, j \leq n, \text{ with } |i - j| = 1 \Rightarrow |x_i - x_j| + |y_i - y_j| = 1$$

$\forall 1 \leq i, j \leq n, \text{ with } i \neq j \Rightarrow (x_i, y_i) \neq (x_j, y_j)$ (9)

If we define function I as by eqn (10)

$$I: \{1, \dots, n\} \times \{1, \dots, n\} \to \{-1, 0\}$$
$$\text{where } \forall 1 \leq i, j \leq n, \text{ with } |i - j| \geq 2$$

$$I(i, j) = \begin{cases} -1 & \text{if } p_i = p_j = H \text{ and } |x_i - x_j| + |y_i - y_j| = 1 \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

then the following definition of the energy function for a valid conformation C is given by eqn (11)

$E(C) = \sum_{1\leq i\leq j-2\leq n} I(i,j)$ (11)

The Bellman equation for Q-learning is as follows by eqn (12)

$$Q(s,a) = r(s,a) + \gamma \cdot \max_{a'} Q(s',a')$$

$Q(s,a) = (1-\alpha)\cdot Q(s,a) + \alpha \cdot \left(r(s,a) + \gamma \cdot \max_{a'} Q(s',a')\right)$ (12)

State space S will consist of $\frac{4^{n-1}}{3}$states, i.e $\bar{S} = \left\{s_1, s_2, \dots, s_{\frac{a^{n-1}}{1}}^{1}\right\}$. Initial state of agent in environment is s1 by eqn (13)

$\delta(s_j, a_k) = s_{4,j-3+k} \forall k \in [1,4], \forall j, 1 \leq j \leq \frac{4^{n-1}-1}{3}$ (13)

We mention that Q∗ is reward that one receives immediately after executing action a from state s, plus value of continuing to follow best course of action after that by eqn (14)

$Q^*(s,a) = r(s,a) + \gamma \cdot \max_{\alpha'} Q^*(\delta(s,a),a')$ (14)

Let us denote by $Q^*(s,a)$ agent's evaluate of $Q^*(s,a)$at n-th training episode. We will prove that $\lim_{n\to\infty} Q_n(s,a) = Q^*(s,a), \forall s \in \mathcal{S}, a \in \delta(s,a)$ by eqn (15)

$$0 \leq r(s,a) \leq \frac{(n-1)\cdot(n-2)}{2}, \forall s \in \mathcal{S}, a \in \delta(s,a)$$

$0 \geq E \geq \sum_{i=1}^{n-2}\sum_{j=i+2}^{n}(-1) = -\frac{(n-1)\cdot(n-2)}{2}$ (15)

During training process, evaluate Q(s, a) grow for every state action pair (∀s ∈ S, a ∈ δ(s, a)) by eqn (16)

$Q_{n+1}(s,a) \geq Q_n(s,a), \forall n \in N^*$ (16)

Our goal is to demonstrate that Inequalities (9) also apply to n+1, that is by eqn (17)

$$Q_{n+1}(s,a) \geq Q_n(s,a).$$

$$Q_{n+1}(s,a) - Q_n(s,a) = \gamma \cdot (\max_{a'} Q_n(s',a') --\max_{a'} Q_{n-1}(s',a'))$$

$$Q_{n+1}(s,a) - Q_n(s,a) \geq \gamma \cdot (\max_{\alpha'} Q_{n-1}(s',a') --\max_{a'} Q_{n-1}(s',a')) = 0$$

$Q_n(s,a) \leq Q^*(s,a), \forall s \in \mathcal{S}, a \in \delta(s,a)$ (17)

We are aware that Q∗ (s, a) is discounted total of the benefits attained by beginning at s, carrying out action a, and adopting the best course of action to reach a destination state. The number of parameters in the network can be significantly reduced by employing this layer. Convolutional layers, which are made up of a few convolutional neurons, are believed to be feature extraction layers. Size of convolution kernel finds size of output feature map. Following formula calculates feature map's size after kernel function has been convolved and transported by eqn (18)

$$\begin{cases} N_x^l = \frac{N_x^{l-1} - K_x^l + 2P_x^l}{S_x^l}, \\ N_y^l = \frac{N_y^l - K_y^l + 2P_y^l}{S_y^l} \end{cases} \quad (18)$$

where K is the convolution kernel's size, P is the fill pixel value, S is the step size, and l is the number of layers at present. After convolution process, activation function changes nonlinearly. Backpropagation approach yields weights as well as parameters of every neuron. Convolutional layer neuron is represented as follows by eqn (19)

$x_j^l = \text{Relu}\left(\sum_{i\in M_j} x^{l-1} w_{ij}^l + b_j^l\right)$ (19)

Values of specific features in input layer are computed as well as combined by using a subsampling layer to reduce the variance of the modified data. This can preserve properties while reducing the number of neurons. The formula for pooling layer is as follows by eqn (20)

$y = \max(x_i),\ x_i \in x$ (20)

where xi represents neuron's output in the region, which is x on the feature map. The final layer in our neural network is composed of three fully connected layers. Full connection layers are susceptible to overfitting issues. We employ the dropout function to lessen the first two layers' overfitting in order to remedy this. The output layer is the final fully connected layer. It is swapped out with an output layer with two neurons that determines whether or not smoke will be the output outcome. Probability is defined as follows using the Softmax function by eqn (21)

$y_j = \frac{\exp(f_j)}{\sum_{i=1}^{2}\exp(f_i)}, j = 1,2,$ (21)

Let $\{x_i\}_{i=1}^n$ be a collection of independent random variables on X with uniform distribution. Give by eq (22)

$I_n(g) = \frac{1}{n}\sum_{i=1}^{n} g(\boldsymbol{x}_i)$ (22)

Result of a quick evaluation is by eq (23)

$\mathbb{E}(I(g) - I_n(g))^2 = \frac{\text{Var}(g)}{n}$, $\text{Var}(g) = \int_X g^2(\boldsymbol{x})d\boldsymbol{x} - \left(\int_X g(\boldsymbol{x})d\boldsymbol{x}\right)^2$ (23)

As NN consists of an input layer, an output layer, hidden layer, figuring out output of the hidden layer is a prerequisite to figuring out the total output of the network. The hidden layer output, denoted as $e^{(H))}$, is computed using Eqs. (24, 25), where nf stands for activation function, I for hidden neuron, j for input neurons, and $w_{(\bar{\varepsilon}i)}^{(H)}$ for bias weight. The NN technique is offered by

$e^{(H)} = nf\left(W_{(Ri)}^{(i)} + \sum_{j=1}^{n} w_{(ji)}^{(i)} F_D\right)$ (24)

$\hat{G}_{\hat{o}} = nf\left(W_{(B\hat{O})}^{(G)} + \sum_{i=1}^{n_0} W_{(\bar{\iota})}^{(G)} e^{(i))}\right)$ (25)

Equation (19) illustrates how weight space as well as corresponding biases for ERN optimisation are produced using weight matrices given in (26) and (27).

$W_n = U_n = \sum_{n=1}^{N} a \cdot (\text{ rand } - \frac{1}{2}),$ (26)

$B_n = \sum_{n=1}^{N} a \cdot \left(\text{rand } - \frac{1}{2}\right)$ (27)

$\left|\mathcal{R}(\hat{f}) - \hat{\mathcal{R}}_n(\hat{f})\right| \leq \sup_{f \in \mathcal{H}_m} \left|\mathcal{R}(f) - \hat{\mathcal{R}}_n(f)\right| = \sup_{f \in \mathcal{H}_m} |I(g) - I_n(g)|$ (28)

in the weight matrix, where Wn is the Nth weight. Equation (29) generates a random number between 0 and 1, which is referred to as the rand in (1).

$W^* = [W_n{}^1, W_n{}^2, w_n{}^3, \ldots, W_n^{N-1}]$ (29)

Sum of square errors are readily predicted for every weight matrix in a NN operation. One input layer, one hidden or "state" layer, "output" layer make up three-layer network employed in the ERN structure.

## 5. Results and discussion:

Authors used Digital Electronic Education and Design Suites (DEEDS) dataset for this study because it uses a proprietary Enhanced Learning Technology (ELT) to monitor student behaviour and interactions in real time in the classroom. This is regarded as one of the few databases with this kind of information that is publicly available worldwide. The data that DEEDS makes available—which includes the learning environment, the amount of time spent on every issue, average idle time, the average number of keystrokes used, performance attained every session, kinds of online activities that every student was engaging in, etc.—is what makes it unique. Because of the abundance of data, it was possible to create a classification method that performs better than other methods used on the same dataset. Table 1 displays description of detailed dataset characteristics.

Table 1. Summary of DEEDS dataset log statistics

| DEEDS data description | Statistics |
|---|---|
| Total students | 115 |
| Average number of students in each sessions | 87 |
| Total online lab sessions | 6 |
| Non participating students in any activity | 7 |
| Total exercise | 6 |
| Total log entries for all session | 230,318 |
| Online log features captured | 13 |
| Total online activities | 15 |

DEEDS dataset created at College of Genoa, Italy in 2015 was created by logging understudies' connections while doing class exercises as well as tests. Understudies gain proficiency with a specific subject in meetings as well as for every single one of the 6 lab points, understudies were expected to perform practices that reach in number from 4 to 6. <eeting greatest grade was set at one or other 4 or 6 relying upon the action and subject. During every movement and test, all understudies were logged for a total and far reaching perspective on their ways of behaving while at the same time playing out the exercises and tests. For example, in a review meeting, understudy exercises are being recorded by DEEDS stage. These exercises are an assortment of a 1 or a considerable lot of the 15 potential exercises, for example, utilizing word processor, utilizing a reproduction timing outline instrument, checking on concentrate on material, and so on. In the wake of going to all lab meetings, understudies will be taking a test where questions are decided to cover every one of the six lab points. DEEDS makes a different record for every understudy going to a meeting and adds another section each one-second span. It is critical to specify that quantity of meetings in DEEDS dataset is 6. Understudies were tried with a halfway and last, most important test in all meetings aside from the first, where no activities nor tests were taken. Meeting 1 information was in the end separated from our handled dataset. The achieved grades will be utilized to mark the pre-owned information.

Table-2 Comparative analysis between proposed and existing technique

| Parameters | SVM | ANN | KGVNN_RTQRM |
|---|---|---|---|
| Random precision | 91 | 93 | 97 |
| RMSE | 81 | 83 | 85 |
| Computation time | 88 | 91 | 95 |
| F-measure | 85 | 88 | 92 |

Table-2 shows comparative based of number of epochs. Here the parameters compared are random precision, RMSE, computation time, F-measure. The existing technique compared are SVM and ANN with proposed technique.

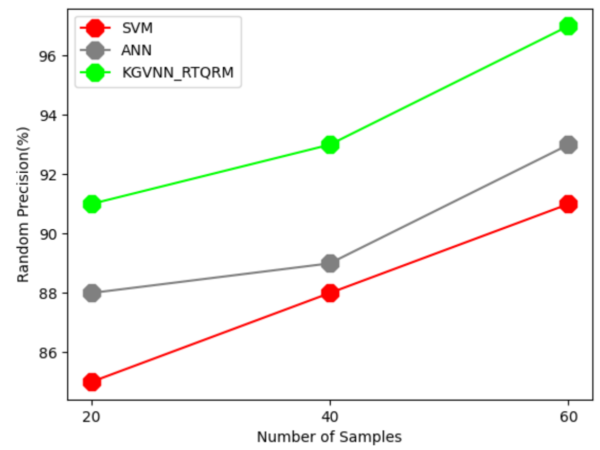


**Figure-2 Comparison of Random precision**

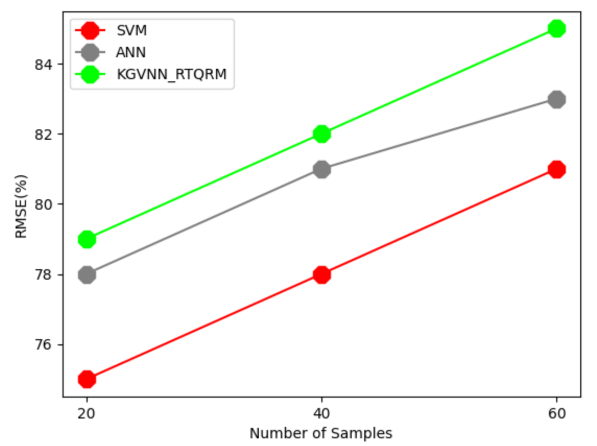


**Figure-3 Comparison of RMSE**

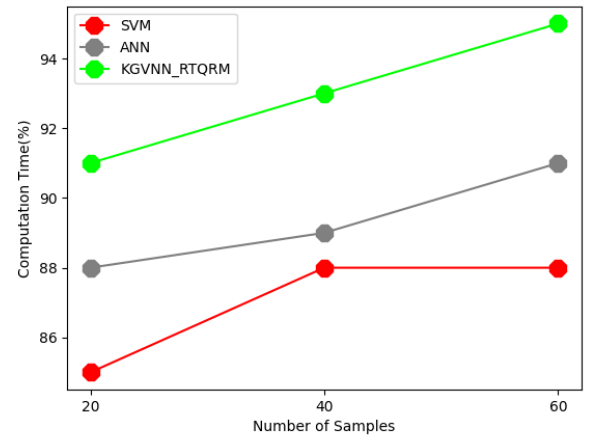


**Figure-4 Comparison of computation time**

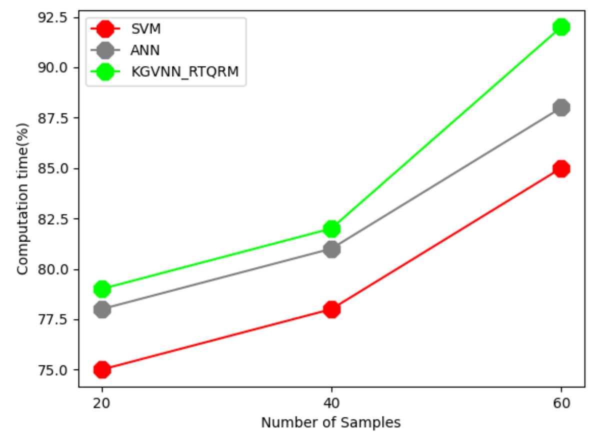


**Figure-5 Comparison of F- measure**

Figure 2-5 shows comparison of random precision, RMSE, computation time, F-measure. Proposed technique attained random precision of 97%, RMSE of 85%, computation time of 95%, F-measure of 92%, while existing SVM attained random precision of 91%, RMSE of 81%, computation time of 88%, F-measure of 85%; ANN attained random precision of 93%, RMSE of 83%, computation time of 91%, F-measure of 88%. We take into account a longer list of variables, all of which fall into one of the three categories below, in order to forecast students' performance: three categories of features: (1) features derived from the number of each of the nine activity types; (2) features derived from the number of timing statistics; and (3) features derived from the number of peripheral activities. Per student and per session, these characteristics were gathered. We suggest adding 252 additional elements, including the 86 features that were thought to better characterise each of the nine activities the students engaged in.
The simulation structure is as follows: (1) For every conceivable value of an attribute ($2 \leq |A| \leq 10$): Create a set P of 100K randomly distributed categorical probability distributions, each with a size of $|A|$. By selecting $|A|$ i.i.d. samples from the uniform distribution over (0, 1) and normalising sum to equal 1, every probability distribution $Pj \in P$ is created. A potential intended distribution over set A of attribute values is represented by each Pj. (b) For every $Pj \in P$, create 100 random candidates for every attribute value in A. These candidates' scores are selected one at a time, i.i.d. from uniform distribution over (0, 1), they are arranged in decreasing order of score. We repeat this process ten times, yielding 1M unique ranking tasks for every option in $|A|$. (ii) Apply every suggested fairness-aware ranking technique to obtain a fairness-aware re-ranked list of size 100. Inputs for each attribute value's generated random candidate list and the intended distribution Pj are used. Utilizing just the characterization correctnesses don't ensure that the model is awesome. It actually intends that on the off chance that the characterization precision is high, it isn't required that the model is great. For this reason, we determined the Kappa values, which makes sense of how much the anticipated worth concurs with the genuine qualities. At the point when we ascertain the arrangement exactness, it thinks about just the corner to corner upsides of the disarray lattice and doesn't think about the non-askew qualities. Notwithstanding, when we ascertain the Kappa, it likewise considers the nondiagonal values. We utilized the renowned Kappa translation to perceive how much there is a concurrence with the genuine qualities. The well

known translation is as 0-0.2: slight understanding, 0.21-0.4: fair arrangement, 0.41-0.60: moderate arrangement, 0.61-0.8: significant understanding, 0.81-0.99: practically amazing understanding.

The proposed model anticipated the final test grades of understudies with 73% exactness. As per this outcome, one might say that scholarly accomplishment can be anticipated with this model from here on out. By anticipating understudies' accomplishment grades in future, understudies can be permitted to audit their functioning strategies and work on their presentation. Te significance of the proposed strategy can be better perceived, taking into account that there is around 2.5 months between midterm tests as well as final tests in advanced education. Comparably thus, understudies' scholastic exhibitions were anticipated utilizing different indicators, different calculations and approaches. Results confirm that ML calculations are utilized to anticipate understudies' scholastic presentation. All the more significantly, the expectation was made exclusively with boundaries of midterm grade, staff and division. Educating staff can benefit from consequences of this exploration in early acknowledgment of understudies who have beneath or better than expected scholastic inspiration. Afterward, for instance as brings up, they can coordinate understudies with below average scholarly inspiration by understudies with better than expected scholastic inspiration and urge them to work in gatherings or task work. Along these lines, the understudies' inspiration is enhanced, their dynamic support in learning is guaranteed. Moreover, such information driven examinations ought to help advanced education in laying out a learning examination structure and add to dynamic cycles.

## 6. Conclusion:

Educational data mining, which employs ML techniques to analyse data from educational settings, has seen substantial study in the area of predicting students' academic progress. Student academic achievement is hard to measure since it depends on so many different factors. The system has generated a lot of data, which needs to be thoroughly analysed to identify the most important details for planning and future expansion. The prediction of students' grades as well as marks based on their previous educational records is one well-known and practical usage of the EDM. It becomes an excellent source of data that can be applied in many different contexts to enhance quality of education across country. Applicable examination study uncovers that various techniques for scholarly execution guaging are worked to carryout enhancements in managerial and educating staf of scholastic associations. In put sent method, procured informational index is pre-handled to cleanse information quality, named scholarly authentic information of understudy (30 ideal ascribes) is used to prepare. scholastic findings via web-based entertainment use and its effects on understudies' scholarly execution are not settled; a few examinations have found positive effects, others opposite, some have found no effects. Subsequently, research has not made it clear how much do online entertainment nature of application, pace of-purpose, period-ofuse predicts understudies' scholarly presentation. Outcomes acquired from this study uncovered conflict for certain areas of the writing. Consequently, while some writing sees web-based entertainment use as a terrible influence on understudies' scholarly execution, this study's results showed in any case. Subsequent expansion in expectation precision of understudies in danger permits scholarly foundations to be more effective in supporting those understudies while using minimal measure of assets. Future examinations might depend on engaging measurements to dissect the job of various psychographic factors as well as their effect on prescient method. It would be fascinating for impending examinations to test auto-produced group models in anticipating understudy profession achievement utilizing scholarly and psychographic information.

## Reference:


1. Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. Smart Learning Environments, 9(1), 11.
2. Balaji, P., Alelyani, S., Qahmash, A., & Mohana, M. (2021). Contributions of machine learning models towards student academic performance prediction: a systematic review. Applied Sciences, 11(21), 10007.
3. Fahd, K., Venkatraman, S., Miah, S. J., & Ahmed, K. (2022). Application of machine learning in higher education to assess student academic performance, at-risk, and attrition: A meta-analysis of literature. Education and Information Technologies, 1-33.
4. Hussain, S., & Khan, M. Q. (2023). Student-performulator: Predicting students' academic performance at secondary and intermediate level using machine learning. Annals of data science, 10(3), 637-655.
5. Zeineddine, H., Braendle, U., & Farah, A. (2021). Enhancing prediction of student success: Automated machine learning approach. Computers & Electrical Engineering, 89, 106903.
6. Hussain, S., Gaftandzhieva, S., Maniruzzaman, M., Doneva, R., & Muhsin, Z. F. (2021). Regression analysis of student academic performance using deep learning. Education and Information Technologies, 26, 783-798.
7. Nti, I. K., Akyeramfo-Sam, S., Bediako-Kyeremeh, B., & Agyemang, S. (2022). Prediction of social media effects on students' academic performance using Machine Learning Algorithms (MLAs). Journal of Computers in Education, 9(2), 195-223.
8. Rai, S., Shastry, K. A., Pratap, S., Kishore, S., Mishra, P., & Sanjay, H. A. (2021). Machine learning approach for student academic performance prediction. In Evolution in Computational Intelligence:

Frontiers in Intelligent Computing: Theory and Applications (FICTA 2020), Volume 1 (pp. 611-618). Springer Singapore.

9. Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. Computers in Human behavior, 104, 106189.
10. Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student'performance prediction using machine learning techniques. Education Sciences, 11(9), 552.
11. Pallathadka, H., Wenda, A., Ramirez-Asís, E., Asís-López, M., Flores-Albornoz, J., & Phasinam, K. (2023). Classification and prediction of student performance data using various machine learning algorithms. Materials today: proceedings, 80, 3782-3785.
12. Xu, X., Wang, J., Peng, H., & Wu, R. (2019). Prediction of academic performance associated with internet usage behaviors using machine learning algorithms. Computers in Human Behavior, 98, 166-173.
13. Issah, I., Appiah, O., Appiahene, P., & Inusah, F. (2023). A systematic review of the literature on machine learning application of determining the attributes influencing academic performance. Decision Analytics Journal, 100204.
14. Ha, D. T., Loan, P. T. T., Giap, C. N., & Huong, N. T. L. (2020). An empirical study for student academic performance prediction using machine learning techniques. International Journal of Computer Science and Information Security (IJCSIS), 18(3), 75-82.
15. Dabhade, P., Agarwal, R., Alameen, K. P., Fathima, A. T., Sridharan, R., & Gopakumar, G. (2021). Educational data mining for predicting students' academic performance using machine learning algorithms. Materials Today: Proceedings, 47, 5260-5267.
16. Poudyal, S., Nagahi, M., Nagahisarchoghaei, M., & Ghanbari, G. (2020, November). Machine learning techniques for determining students' academic performance: A sustainable development case for engineering education. In 2020 International Conference on Decision Aid Sciences and Application (DASA) (pp. 920-924). IEEE.
17. Maurya, L. S., Hussain, M. S., & Singh, S. (2021). Developing classifiers through machine learning algorithms for student placement prediction based on academic performance. Applied Artificial Intelligence, 35(6), 403-420.
18. Rastrollo-Guerrero, J. L., Gómez-Pulido, J. A., & Durán-Domínguez, A. (2020). Analyzing and predicting students' performance by means of machine learning: A review. Applied sciences, 10(3), 1042.
19. Kishor, K., Sharma, R., & Chhabra, M. (2021, September). Student performance prediction using technology of machine learning. In International Conference on Micro-Electronics and Telecommunication Engineering (pp. 541-551). Singapore: Springer Nature Singapore.
20. Yakubu, M. N., & Abubakar, A. M. (2022). Applying machine learning approach to predict students' performance in higher educational institutions. Kybernetes, 51(2), 916-934.
21. Ofori, F., Maina, E., & Gitonga, R. (2020). Using machine learning algorithms to predict students' performance and improve learning outcome: A literature based review. Journal of Information and Technology, 4(1), 33-55.
22. Sekeroglu, B., Dimililer, K., & Tuncal, K. (2019, March). Student performance prediction and classification using machine learning algorithms. In Proceedings of the 2019 8th International Conference on Educational and Information Technology (pp. 7-11).
23. Bhutto, E. S., Siddiqui, I. F., Arain, Q. A., & Anwar, M. (2020, February). Predicting students' academic performance through supervised machine learning. In 2020 International Conference on Information Science and Communication Technology (ICISCT) (pp. 1-6). IEEE.
24. Aslam, N., Khan, I., Alamri, L., & Almuslim, R. (2021). An improved early student's academic performance prediction using deep learning. International Journal of Emerging Technologies in Learning (iJET), 16(12), 108-122.