TÍTULO: La Inteligencia Artificial en Educación: aplicación en la evaluación del desempeño del alumno.

AUTORES:
1. Dr. Boran Sekeroglu.
2. Dr. Kamil Dimililer.

RESUMEN: La educación ha ganado importancia para acelerar el desarrollo de los países mediante la educación personalizada, etc., e impulsar su adquisición de conocimiento. Si bien es tarea difícil analizar a cada estudiante de forma independiente, los enfoques de Inteligencia Artificial y aprendizaje automático proporcionan soluciones rápidas con alta precisión para ayudar a los humanos. Este artículo presenta experimentos con diferentes conjuntos de datos educativos para predecir y clasificar el desempeño de los estudiantes, considerándose varios modelos de aprendizaje automático para cada problema y los resultados muestran que la Inteligencia Artificial puede ayudar a los instructores a mejorar la educación personalizada antes o durante el semestre activo. Se analiza que la red neuronal de la función de base radial supera a otros modelos considerados.

PALABRAS CLAVES: Inteligencia artificial, educación personalizada, función de base radial NN, aprendizaje automático.
TITLE: Artificial Intelligence in Education: application in student performance evaluation.

AUTHORS:

1. Dr. Boran Sekeroglu.
2. Dr. Kamil Dimililer.

ABSTRACT: Education has gained importance to accelerate the development of countries by means of personalized education and students' characteristics, family properties, etc., and boost their knowledge acquisition. While it is challenging task for humans to analyze each student independently, Artificial Intelligence and machine learning approaches provide rapid solutions with high accuracy to assist humans. This paper presents several experiments with different educational datasets to predict and classify student performances. Several machine learning models are considered for each problem and obtained results show that Artificial Intelligence can assist instructors to improve personalized education before or during the active semester. It is also analyzed that Radial Basis Function Neural Network outperforms other models considered in this research.

KEY WORDS: Artificial Intelligence, Personalized Education, Radial-Basis Function NN, Machine Learning.

INTRODUCTION.

Education is the process of facilitating learning, or the acquisition of knowledge, skills, values, beliefs, and habits. Educational methods include storytelling, discussion, teaching, training, and directed research. It is vital human right that can be represented as various statements which is granted with equal quality with universal Access [1].
Recently, computers and portable devices are employed in daily life in providing access for varieties of electronically available materials. These developments are used to improve the communication between educators and learners [2]. So, contemporary education has been combined with technology [3].

Since technology involved with humans, the human and machine learners should learn from each other as a cooperative system to create the learning environment effectively and efficiently that will help the human learner [4]. Blended learning approaches has been started by higher education institutions into their traditional learning or teaching systems to improve the quality of learning as well as teaching when education has been considered [2].

Artificial Intelligence (AI) in teaching and learning has gained importance in last decade in learning and teaching processes. New technology-based education has been preferred to help people in reaching their educational goals. [3-5]. Apart from education, various applications in artificial intelligence or machine learning have been considered in recent publications by various researchers. Machine Learning (ML) has been applied in enhancing the prediction of outcomes through classification techniques considering the techniques used to classify the input data into categories that were pre-determined [6].

With the aid of ML, tool wear prediction has been introduced which is based on random forests, backpropagation neural network and support vector regression. The authors compared the performance of ML algorithms with each other to analyze the performance of the given algorithms. [7]. Recently, Machine learning based feature selection for carbonate reservoir cementation factor prediction has been applied. The authors presented a comparative study of the performance of a non-linear feature selection system that is based on fuzzy ranking. A subset of input data, which is selected by fuzzy ranking algorithm, is given to the artificial neural network and support vector machine
models and a multivariate linear regression model has been implemented. [8]. Also, researches were conducted in order to determine the most effective ML models for different kind of problems [9].

Machine Learning in multi-disciplinary fields provide useful information to users; for example, in [10] authors provide information that teachers can use virtual assistance (AI/ML enabled teaching algorithms) to enhance teaching methods for learning experiences of students and to improve achievements of students. By using these algorithms, teachers could build customized curriculums for students who have different learning abilities.

Another study on ML in evaluation of educational techniques show that online provided teaching materials or homework are more effective than traditional handwritten homework because of simultaneous provided results and feedbacks to the students [11].

In [12], a new ML platform is developed for secondary education students. In this work, authors created the platform for learning the theory of subjects with underlining solutions. They suggested two different scenarios, which are recycling sorting problem and finding different types of bacteria present in a pond. The developed platform can be used in any type of educational scenarios with valid data.

Gabriel et al. [13] demonstrated a ML algorithm in order to find how education researchers could employ data driven modeling techniques for finding complex mathematical relationships and interactions in a multi-dimensional dataset.

Recently, Yilmaz and Sekeroglu [14] performed a questionnaire that includes different sections such as personal and family properties and educational preferences. They classified final grades of students using machine learning algorithms according to this information, and they achieved 70% to 85% of success rate with Radial Basis Function Neural Network.
The rest of the paper is organized as follows; Section 2 gives brief introduction about the considered dataset and related work, and Section 3 introduces considered ML models. Section 4 presents design of experiments and obtained results in details. Section 5 presents the discussions on these results and finally Section 6 concludes the remarks on this research.

**Dataset and related work.**

This section introduces the considered datasets and related work briefly to provide more information to understand the job done in this research.

**Dataset.**

Two datasets: Student Performance Dataset (SPD) [15] and Students Academic Performance Dataset (SAPD) [16] are considered in this research in order to perform prediction and classification of students’ performances.

The aim of SPD is to predict the performances of secondary school students in Portugal while the aim of SPAD is to classify the academic performances of college level students.

SPD includes 33 attributes related with parental status, family size etc. to predict students' performances and consists 649 and 395 instances in Portuguese and Math courses respectively.

SPAD includes relatively similar but less, 21 attributes and 3 outputs; as Good, Average and Poor performances. It has 131 instances collected from college students.

**Related Work.**

In the preliminary experiments [17], Support Vector Regression (SVR), Backpropagation (BP) and Long Short-Term Memory Neural Network (LSTM) was used as prediction models to predict the performances of Math and Portuguese courses of SPD. For classification, Backpropagation (BP) Neural Network, Support Vector Machine (SVM), Gradient Boosting Classifier (GBC) were used.
All experiments were divided into two categories according to their training ratios, which are the number of instances used to train the models. Selected ratios were 30% and 40% in order to determine superior model and to obtain optimal results.

Obtained results showed that SVR was able to produce highest rates in prediction of student performances and Backpropagation (BP) was the superior model for classification of students’ performances. Thus, it is decided to consider the superior models of related work in this research with other machine learning models in order to improve the prediction rates and classification accuracy of students' performances.

Models.

As it was mentioned above, models considered in related work with highest prediction and classification rates are considered in this research too. In addition, Decision Tree Classifier and Regressor, and Radial Basis Function Neural Network is considered for both types of problems. Thus, Support Vector Regression (SVR), Decision Tree Regressor (DTR) and Radial Basis Function Neural Network (RBFNN) for prediction and Backpropagation (BP), Decision Tree Classifier (DTC) and RBFNN for classification is implemented. Following subsections briefly introduces the considered models.

Support Vector Regression.

Support Vector Regression is a modified version of Support Vector Machines for prediction problems [18]. It assigns support vectors to draw decision boundary between observed data and predicted data. Its' main drawback is the limitation on effective prediction on big data [19]. General architecture of SVR can be seen in Figure 1.
**Decision Tree Classifier and Regressor.**

Decision Tree starts by a root node and divided to internal nodes until leaf nodes. It depends on divide-and-conquer strategy and divides root node to internal nodes [20]. Different algorithms are used to determine root node, leaf nodes and general structure of tree [21]. One of the most widely used algorithms is Gini for classification problems, it is based on the frequency of repetition of the nodes, and Mean Squared Error is generally considered for prediction problems. General structure of Decision Trees can be seen in Figure 2.

**Radial Basis Function Neural Network.**

RBFNN is a kind of neural network that can be used both for classification and prediction problems [22-23]. It strictly has a single hidden layer and uses Radial Basis Functions in hidden neurons on this layer. It provides an advantage on the results of RBFNN by increasing its stability. General topology of RBFNN is shown in Figure 3.
Backpropagation Neural Network.

BPNN is a traditional feedforward neural network with backpropagation learning algorithm. It might be used as shallow or as deep structure by adding or removing hidden layers in the architecture. It is one of the most known and frequently used algorithms both in classification and prediction problems [24-25]. General topology of BP with multi-hidden layer can be seen in Figure 4.
Experiments and Results.

This section presents the preparation of experiments, evaluation strategies and obtained results in details.

Experiments and Evaluation Criteria

As it was mentioned above, experiments are performed on Student Performance Dataset (SPD) and Students Academic Performance Dataset (SAPD). Experiments are divided into two categories according to the number of training instances, which are 60% and 70% in order to determine optimal prediction and classification rates. Training is performed by hold-out method, which is based on randomly selection of training and testing instances according to the pre-determined training ratio. Evaluation is performed by different criteria in different problem domains. Mean Squared Error (MSE) and $R^2$ value is used for the evaluation of prediction problem.
MSE is used to determine how much the system minimized the error between predicted values by models and observed data. Minimum value indicates highest prediction rate however, it is always required to consider different evaluation criteria for prediction problems because in extreme conditions, minimized error cannot produce effective prediction if some of data has huge or small error that causes over or underestimation. Thus, R2 which is also based on MSE is considered but it determines the baseline by using all data points of expected or observed data and uses it to avoid the over or underestimation during the evaluation of prediction performances. Highest R\(^2\) score indicates optimal prediction rate. Equation 1 and 2 shows the formula of MSE and R\(^2\) respectively.

\[
\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2
\]

where \(N\) is the total number of samples and \(y_i\) and \(\hat{y}_i\) are the observed and predicted values respectively.

\[
R^2 = \frac{\text{MSE}}{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}
\]

where \(\hat{y}_i\) is the mean of expected data.

For classification domain, general accuracy is used for the evaluation of results that is based on the number of correctly classified samples and total number of samples. The formula of accuracy can be seen in Equation 3.

\[
\text{Total Accuracy} = \frac{CS}{TS}
\]

where \(CS\) and \(TS\) are the correctly classified samples and the total number of samples respectively.

**Parameter Selection for Models.**

Some ML models includes different parameters that is used for the optimization of the results. In this subsection, determination of each parameter will be explained. All parameters are determined after
performing several experiments by trial and error because there is not any rule to determine any ML model parameters without performing experiments.

Non-linear correlation of attributes makes learning or convergence of neural network based models complicated and different variety of hidden layers and hidden neuron numbers within these layers are trained in BP. Superior and most accurate results are obtained with four hidden layer, which makes BP deep, and 500 neuron per layer is determined to be used. Sigmoid activation function is used for the activation of each neuron in each layer and learning rate and momentum rate set to 0.0009 and 0.90 respectively. Maximum iterations are determined as 1000, which produced the minimum error. Number of attributes is the number of input neurons and three output neurons is used classification problem.

For RBFNN, same input and output neuron numbers are considered according to the problem domain, however, for prediction problems, single output neuron is used. Maximum 3000 iterations are determined to be used in order to achieve minimum error and without over-fitting. Learning rate is used as 0.09 and cluster number is arranged as 20.

SVR consists different kernel functions but most widely effectively used one is Radial Basis Function kernel and it is preferred to be used in this research. Gamma and epsilon values set to 0.005 and 0.01 respectively after several experiments.

In Decision Trees, there are not additional parameters like in neural networks however; the determination of the structure of tree is another kind of problem. For prediction problem, MSE is used to build the structure and for classification problem, Gini function is used to build the tree.

**Results for Prediction Problems.**

Three models, SVR, DTR and RBFNN are trained using 60% and 70% of total data separately in two different experiments for both courses indicated above. In 70% of training ratio of Math course, better
results are produced by DTR and RBFNN than SVR, which was the optimum model in recent research. DT produced 0.00966 and 0.759 MSE and $R^2$ score respectively and optimum results are obtained by RBFNN with the highest $R^2$ and lowest MSE results are obtained by RBFNN. Table 1 shows obtained results for 70% of training ratios for Math course. Figure 5 shows the prediction figure of RBFNN for Math course with 70% of training ratio.

When 60% of training ratio is considered for Math course, it is observed that the prediction rate of DT is decreased while training samples are increased, and it produced lowest $R^2$ score in this experiment. Although MSE results are close to each other for each model, lowest one is obtained by SVR. Highest $R^2$ score is achieved by RBFNN with 0.845. Table 2 shows results for 60% of training ratios for Math course.

**Table 1.** Results of Prediction Experiment of Math Course for 70% of Training Ratio.

<table>
<thead>
<tr>
<th>Result</th>
<th>RBFNN</th>
<th>SVR</th>
<th>DTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.00860</td>
<td>0.01600</td>
<td>0.00966</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.892</td>
<td>0.757</td>
<td>0.759</td>
</tr>
</tbody>
</table>

**Table 2.** Results of Prediction Experiment of Math Course for 60% of Training Ratio.

<table>
<thead>
<tr>
<th>Result</th>
<th>RBFNN</th>
<th>SVR</th>
<th>DTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.01340</td>
<td><strong>0.01300</strong></td>
<td>0.01320</td>
</tr>
<tr>
<td>$R^2$</td>
<td><strong>0.845</strong></td>
<td>0.785</td>
<td>0.709</td>
</tr>
</tbody>
</table>
Fig. 5. RBFNN prediction graph for Math course with 0.0086 MSE and 0.892 $R^2$ score.

In second course (Portuguese) of dataset, similar results are obtained for $R^2$ score and highest results are achieved by RBFNN with 0.894 and 0.886 for 70% and 60% of training ratios respectively. SVR produced minimum MSE values for both training ratios but it could not produce highest $R^2$ scores. DTR produced lowest $R^2$ score and highest MSE results in these experiments. Table 3 and 4 shows obtained results for 70% and 60% of training ratios respectively and Figure 6 presents prediction graph of DTR, which produced worse results in this experiment.

**Table 3.** Results of Prediction Experiment of Portuguese Course for 70% of Training Ratio.

<table>
<thead>
<tr>
<th>Result</th>
<th>RBFNN</th>
<th>SVR</th>
<th>DTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.00882</td>
<td>0.00460</td>
<td>0.0101</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.894</td>
<td>0.834</td>
<td>0.667</td>
</tr>
</tbody>
</table>
Table 4. Results of Prediction Experiment of Portuguese Course for 60% of Training Ratio.

<table>
<thead>
<tr>
<th>Result</th>
<th>RBFNN</th>
<th>SVR</th>
<th>DTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.00983</td>
<td>0.00540</td>
<td>0.0111</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.886</td>
<td>0.834</td>
<td>0.657</td>
</tr>
</tbody>
</table>

Fig.6. DTR prediction graph for Portuguese course with 0.0101 MSE and 0.667 $R^2$ score.

Results for Classification Problem.

Similar to prediction domain, experiments are divided into two groups according to the training ratios and RBFNN, BPNN and DTC models are considered.

For 70% of training ratio, Decision Tree Classifier produced excessively low classification accuracy with 22.5%. Close results are obtained by RBFNN and BPNN however, RBFNN achieved 89.01% while BPNN produced 87.78% of accuracy.
For 60% of training ratio, again, similar and close results are obtained, and decrement of training ratio caused little increment of accuracy in DTC to 28.30%. However, this decrement of training ratio, caused decrement of accuracy in neural network models but again RBFNN produced highest accuracy with 84.61% and followed by BPNN with 80.91%. Table 5 shows all results of classification domain. Figure 7 and 8 presents RBFNN and BP results for classification domain with 70% of testing ratio respectively.

**Table 5.** Accuracy Results of Combined Courses Experiment.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>RBFNN</th>
<th>BPNN</th>
<th>DTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>70% of Training</td>
<td>89.01%</td>
<td>87.78%</td>
<td>22.50%</td>
</tr>
<tr>
<td>60% of Training</td>
<td>84.61%</td>
<td>80.91%</td>
<td>28.30%</td>
</tr>
</tbody>
</table>

**Fig.7.** RBFNN classification (89.01 %) graph for 70% of Training Ratio.
Fig.8. BP classification (87.78 %) graph for 70% of Testing Ratio.

Discussions.

Obtained results shows that neural network models produce superior results both in prediction and classification domain.

Although SVR produced minimal MSE in 75% of the prediction experiments, optimal results are achieved by RBFNN in both domains if $R^2$ score is considered which is the main indicator in this research.

Decision Tree Regressor could not reach the results obtained by other models, while its' performance is strongly depends on dataset characteristics. Number of instances and non-linear correlation between these attributes that are the questions in questionnaire caused ineffective prediction results in DTR. It produced relatively higher results in Math experiments than Portuguese experiments because increased number of the number of instances in Portuguese experiment. Tree became complicated, and leaf and decision nodes lost their effectiveness during the convergence. At that time, it is observed that Support Vector Regressor act as inverse as DTR. While the instances are increased
in Portuguese experiment, superior results are obtained. Radial Basis Function Neural Network produced superior results than other models, but it is also noticed that the decrement of training instances may cause little fluctuations in the testing efficiency of the model.

In classification domain, even though less instances are considered, Decision Tree Classifier failed to classify instances correctly. Decrement of instances that were used to build the tree caused increment of classification accuracy, but general results are not sufficient for DTC. RBFNN and BP, which are both neural network-based models produced and similar and close results in all experiments. However, RBFNN produced optimum classification accuracy both in two experiments. It is also observed that, similar to prediction domain, decrement of training instances causes decrement of accuracy in classification too. This shows effect of the radial-basis functions in hidden neurons instead of considering traditional hidden neurons.

CONCLUSIONS.

Student is the main target of the education. The aim of the education is always to teach subjects effectively to students in proper ways. However, teaching styles and the perception of persons’ differ from each other. Thus, it is required to analyze students’ capabilities, characteristics, preferences and environmental effects for specific courses before they are taking these courses.

At that time, machine learning comes into prominence to provide information that is more effective to teachers and instructors by classifying and predicting the pre-features. In this paper, six experiments were performed in two problem domains as prediction and classification.

All these experiments prove, that the usage of machine learning models can effectively be used in education. Performances of students can be predicted or classified before they are taking the courses and personal education can be applied to those, which results are poor.
Radial Basis Function Neural Network outperformed other models both in prediction and classification domain and it showed its' efficiency in prediction with highest $R^2$ score and with accuracy up to 89.01% in classification. It is also noticed that the decrement of training instances makes negative effect on testing phases of neural networks models.

Future work will include the implementation of deep learning algorithms in order to optimize prediction and classification results.

**BIBLIOGRAPHIC REFERENCES.**


https://dilemascontemporaneoseducacionpoliticaayvalores.com/_files/200005364-8a0e58b095/EE%202019.07.34%20Sistema%20de%20Lector%20%C3%93ptico%20de%20Marcas.pdf

DATA OF THE AUTHORS.

1. **Boran Sekeroglu.** Dr., Assist. Prof., and Chairperson in the Department of Information Systems Engineering, Near East University, Nicosia, Cyprus. Email: boran.sekeroglu@neu.edu.tr

2. **Kamil Dimililer.** Dr., and Assoc. Prof. in the Department of Electrical Electronic Engineering, Near East University, Nicosia, Cyprus. Email: kamil.dimililer@neu.edu.tr

3. **Kubra Tuncal.** Research Assistant, Department of Information Systems Engineering, Near East University, Nicosia, Cyprus. Email: kubra.tuncal@neu.edu.tr

**RECIBIDO:** 2 de agosto del 2019.  
**APROBADO:** 18 de Agosto del 2019.