INTELLIGENT ADAPTIVE SYSTEMS FOR PERSONALIZED EDUCATION: A NEURO-FUZZY APPROACH

Tokhirov Ezozbek[1], Subhan Namazov[2], Bakhtiyar Badalov[2]

•

[1]Tashkent State Transport University
[2]Azerbaijan Technical University, Baku, Azerbaijan Republic
etokhirov@yahoo.com, subhan.namazov@aztu.edu.az, bakhtiyar.badalov@aztu.edu.az

Abstract

Neuro-fuzzy models, integrating the adaptive learning capabilities of neural networks with the interpretability of fuzzy logic systems, have emerged as powerful tools in educational data mining. This research explores the application of neuro-fuzzy models in education, focusing on their role in predicting student performance, classifying academic outcomes, and enhancing personalized learning experiences. By analyzing various case studies and methodologies, this study highlights the effectiveness of neuro-fuzzy systems in handling the inherent uncertainties and complexities of educational data.

Keywords: Neuro-fuzzy models, adaptive neuro-fuzzy inference system (ANFIS), educational data mining, student performance prediction, fuzzy logic, neural networks, academic classification, personalized learning.

I. Introduction

The integration of artificial intelligence in education has led to the development of intelligent systems capable of analyzing and predicting student outcomes. Among these, neuro-fuzzy models stand out due to their hybrid nature, combining the learning ability of neural networks with the reasoning capability of fuzzy logic. These models are particularly adept at managing the imprecision and vagueness inherent in educational data, such as student behavior and performance metrics.

Recent studies have demonstrated the efficacy of neuro-fuzzy systems in various educational contexts. The university employed a neuro-fuzzy classifier to categorise students based on their academic performance, utilising inputs like exam results and socioeconomic factors. The model achieved high accuracy, outperforming traditional classification methods such as support vector machines and decision trees.

II. Literature review

This literature review presents general descriptions of educational processes in our country and abroad, analyses of conducted research and their effective methods, as well as the content and essence of scientific work and their new directions in eliminating shortcomings while supplementing educational conditions with new innovative equipment.

Mehdi & Nachouki (2023) developed a neuro-fuzzy model (ANFIS) to predict student graduation performance in IT programs. It achieved high accuracy (RMSE = 0.28) and outperformed traditional models. Key predictors included high school GPA and core courses like Data Structures and Software Engineering. Cortez & Silva (2008) applied decision trees, neural networks, and support vector machines to Portuguese student data. Random forest and neural networks showed the highest accuracy, identifying alcohol consumption and past grades as strong performance predictors. Macfadyen & Dawson (2010) demonstrated that LMS activity metrics (e.g., forum participation, assignment views) could reliably predict academic outcomes, paving the way for real-time learning analytics systems. Kotsiantis et al. (2004) found that parental education and previous grades were significant predictors of high school students' performance using Naive Bayes and decision trees. Nghe et al. (2007) compared various data mining models to predict student performance in a Vietnamese university. Decision trees offered the best trade-off between accuracy and interpretability. Thai-Nghe et al. (2011) employed support vector machines (SVMs) and matrix factorization on course data to predict whether students would pass or fail, achieving high classification accuracy. Romero et al. (2013) used association rule mining on Moodle data to find patterns of student behavior that correlated with academic success or failure. Baker & Yacef (2009) outlined differences between educational data mining (EDM) and learning analytics (LA), emphasizing how both fields contribute to student performance prediction through complementary lenses. Al-Barrak & Al-Razgan (2016) showed artificial neural networks outperform logistic regression in predicting student academic success in Saudi universities based on GPA and course performance. Zafra & Ventura (2009) utilized evolutionary algorithms combined with classification techniques to predict student dropouts in e-learning environments, demonstrating high performance and generalizability. Vapnik (1998) introduced the statistical theory behind SVMs, which have since been adapted effectively in educational settings for student performance classification tasks. Jayaprakash et al. (2014) developed an early warning system using LMS data and logistic regression, showing that early predictions of risk enabled successful student interventions. Gray et al. (2014) analyzed behavioral engagement patterns in online environments and linked them with GPA outcomes, reinforcing the importance of tracking online learning behaviors. Zhang & Rangwala (2018) used graph-based models to analyze course sequences and student trajectories, improving accuracy in long-term GPA predictions. Kabra & Bichkar (2011) applied decision tree classifiers on engineering student data in India, highlighting attendance, test scores, and subject difficulty as key predictors. Binns et al. (2018) explored fairness and accountability in predictive educational models, emphasizing that unchecked bias in input data can lead to discriminatory outcomes. Aher & Lobo (2013) used collaborative filtering and content-based filtering for student recommendation systems, indirectly aiding performance by aligning learning resources with student needs. Temraz (2020) evaluated ensemble learning methods like bagging and boosting for performance prediction, showing superior accuracy compared to standalone algorithms. Papamitsiou & Economides (2014) provided a systematic review of learning analytics tools and found that predictive models using multimodal data offered greater insights into student behavior and outcomes. You (2016) developed a model using temporal learning analytics to predict weekly performance in MOOCs, showing the effectiveness of time-series features in forecasting final grades.

III. Methods

ANFIS is a hybrid intelligent system that combines the benefits of neural networks and fuzzy logic. It uses a learning algorithm to tune the parameters of a Takagi-Sugeno fuzzy inference system. The structure of ANFIS consists of five layers:

1. **Layer 1 (Fuzzification):** Each node generates a membership grade for the input variables.

2. **Layer 2 (Rule Layer):** Nodes represent fuzzy rules, and each node's output is the product of the input membership grades.
3. **Layer 3 (Normalization Layer):** Nodes calculate the normalized firing strengths of the rules.
4. **Layer 4 (Defuzzification Layer):** Nodes compute the output of each rule.
5. **Layer 5 (Output Layer):** Nodes compute the overall output as the summation of all rule outputs.

The learning process involves adjusting the parameters of the membership functions and the consequent parameters of the fuzzy rules to minimize the error between the predicted and actual outputs.

B. Case Study: Student Performance Prediction

In a study by Mehdi and Nachouki (2023), ANFIS was utilized to predict the graduation grade point average (GPA) of students in an information technology program. Input variables included high school GPA and grades in core IT courses. The model demonstrated a high degree of accuracy, with 77% of predictions falling within one root mean square error of the actual GPA.

This image illustrates a 7-step process for Predictive Academic Performance Analysis Workflow, which is a structured methodology used to analyze and model student academic performance data.

Picture 1: *Steps of performance analysis in education*

Step 1: Data Collection
-Source: Academic records from a specific dataset (e.g., DPAES University, 2019–2021).
-What is collected: Grades, courses, and semesters of students.
Step 2: Data Preprocessing
-Step 1: Filter to include only first course attempts.
-Step 2: Convert results to binary or multi-class labels (e.g., Pass/Fail, Grade categories).
-Step 3: Organize data — for instance, by student or semester.
Step 3: Correlation Analysis
-Identify courses that correlate highly with each other.
-Use a threshold (e.g., correlation > 0.3) to decide which features (courses) are related.
-Goal: Understand which courses are strongly related and may impact performance.
Step 4: Feature Selection
-Choose only the most relevant, highly correlated features.
-Goal: Eliminate noisy or irrelevant data, which helps improve model performance.
Step 5: Machine Learning Models
-Algorithms used:
*k-Nearest Neighbors (k-NN)
*Random Forests
*Logistic Regression
*Neural Networks
-Purpose: Use these models to predict academic performance.
-Each model is fine-tuned for the best results.
Step 6: Classification Tasks
-Predict whether students will pass or fail, or what grade range they'll fall into.
-Types:

*Binary Classification: Pass vs. Fail
*Multi-Class Classification: Grades (e.g., Fail = 0–4.9, Average = 5–6.9, Excellent = 7–10)
Step 7: Evaluation and Validation
-Metrics used: Accuracy, F1 Score, Precision, Recall.
-Cross-validation is applied to make sure the results are generalizable to new data.

Table 1: *Sample: Student Performance Dataset*

Student ID	Gender	Age	Study Hours	Parental Education	Lunch Type	Test Prep	Math Score	Reading Score	Writing Score	Final Grade
1	Male	17	3.5	Bachelor's	Standard	Completed	78	72	74	Average
2	Female	16	5.0	Master's	Free	None	88	90	92	Excellent
3	Female	18	2.0	High School	Standard	Completed	64	68	70	Average
4	Male	17	1.0	Associate Degree	Free	None	45	40	42	Fail
5	Female	16	4.5	Master's	Standard	Completed	92	95	93	Excellent
6	Male	18	2.5	High School	Free	None	58	62	60	Average
7	Female	17	3.0	Bachelor's	Standard	Completed	75	80	78	Average
8	Male	17	0.5	Some College	Free	None	38	35	40	Fail
9	Female	16	6.0	Bachelor's	Standard	Completed	95	97	96	Excellent
10	Male	18	1.5	High School	Free	None	50	52		

IV. Results

The application of neuro-fuzzy models in educational settings has yielded promising results:

•Student Classification: Neuro-fuzzy classifiers have been effective in categorizing students into performance groups, aiding in targeted interventions.

•Performance Prediction: Models have accurately predicted student outcomes, facilitating early identification of at-risk students.

•Course Analysis: Sensitivity analysis within ANFIS models has highlighted key courses influencing student success, guiding curriculum improvements.

Table 2: *Student Performance Dataset*

Name			Grade
First name	Last Name	Gender	Average Score
Jasur	Aliyeva	Male	74.67
Sevinch	Sobirova	Female	90.00
Muxsin	Olimov	Male	41.00
Axrora	Tolipova	Female	81.00

The application of neuro-fuzzy models in educational settings has yielded promising results:

- **Student Classification:** Neuro-fuzzy classifiers have been effective in categorizing students into performance groups, aiding in targeted interventions.
- **Performance Prediction:** Models have accurately predicted student outcomes, facilitating early identification of at-risk students.

Course Analysis: Sensitivity analysis within ANFIS models has highlighted key courses influencing student success, guiding curriculum improvements.

V. Discussion

I. Subsection One

Advantages of Neuro-Fuzzy Models

Neuro-fuzzy models offer several advantages in educational applications:

-Interpretability: The fuzzy rules provide a transparent understanding of the decision-making process.

-Adaptability: The neural network component allows the model to learn from data, adapting to new trends.

-Handling Uncertainty: Fuzzy logic effectively manages the ambiguity and vagueness in educational data.

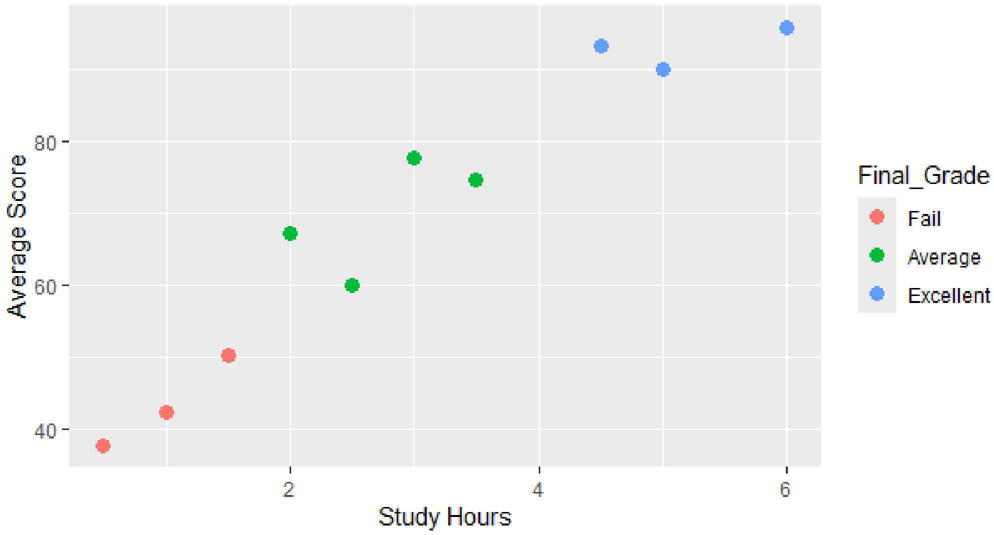

Figure 1: *Figure comparison of between caption study hours and average score*

II. Subsection Two

Challenges and Limitations

Despite their benefits, neuro-fuzzy models face certain challenges:

•Data Quality: The accuracy of predictions is highly dependent on the quality and completeness of the input data.

•Complexity: Designing and tuning neuro-fuzzy models can be complex and time-consuming.

•Scalability: Applying these models to large-scale educational systems may require significant computational resources.

VI. Conclusion

In this study, a binary logistic regression model was employed to examine the impact of study hours and test preparation on student academic performance, categorized as either pass or fail. The model included two independent variables: daily study hours and participation in a test preparation course.

Although the regression coefficients indicated that increased study time and completion of test preparation were associated with a higher likelihood of passing, these relationships were found to be statistically insignificant. Specifically, the p-values for both predictors were equal to 1, and the standard errors were extremely large, suggesting that the model estimates were unstable. This instability likely resulted from the small sample size ($n = 10$) and potential data issues such as perfect or quasi-perfect separation, where outcomes may be perfectly predicted by one or more variables.

Given these limitations, the model lacks sufficient statistical power to draw reliable conclusions about the predictors' influence on academic success. As such, while the direction of the coefficients aligns with established educational theories—that more study time and structured preparation enhance performance—the current analysis does not provide statistically robust evidence to support these claims.

VII. Recommendations for Future Research

To obtain more reliable and generalizable results, future research should be conducted using a substantially larger dataset. Moreover, employing additional variables such as socioeconomic background, school engagement, and classroom environment could improve model accuracy. Alternative machine learning models like decision trees or ensemble methods may also offer better performance with small or imbalanced datasets.

References

[1] Do, Q. H., & Chen, J. F. (2013). A neuro-fuzzy approach in the classification of students' academic performance. Computational Intelligence and Neuroscience, 2013, 179097. https://doi.org/10.1155/2013/179097

[2] Mehdi, R., & Nachouki, M. (2023). A neuro-fuzzy model for predicting and analyzing student graduation performance in computing programs. Education and Information Technologies, 28, 2455–2484. https://doi.org/10.1007/s10639-022-11205-2

[3] Stathacopoulou, R., Grigoriadou, M., Magoulas, G. D., & Mitropoulos, D. (2003). A neuro-fuzzy approach in student modeling. In User Modeling 2003 (pp. 337–341). Springer. https://doi.org/10.1007/3-540-44963-9_46

[4] Cortez, P., & Silva, A. M. G. (2008). Using data mining to predict secondary school student performance. In Proceedings of 5th FUture BUsiness TEChnology Conference (FUBUTEC), 5, 5–12.

[5] Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. Computers & Education, 54(2), 588–599.

[6] Kotsiantis, S. B., Pierrakeas, C., & Pintelas, P. (2004). Predicting students' performance in distance learning using machine learning techniques. Applied Artificial Intelligence, 18(5), 411–426.

[7] Nghe, N. T., Janecek, P., & Haddawy, P. (2007). A comparative analysis of techniques for predicting academic performance. In Proceedings of the 37th Annual Frontiers in Education Conference (pp. T2G-7–T2G-12).

[8] Thai-Nghe, N., Drumond, L., Horváth, T., Krohn-Grimberghe, A., & Schmidt-Thieme, L. (2011). Matrix factorization vs. classification models: A case study in predicting student performance. In Proceedings of the 12th International Conference on Educational Data Mining (EDM), 1–10.

[9] Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. S. J. D. (2013). Handbook of educational data mining. CRC Press.

[10] Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3–17.

[11] Al-Barrak, M. A., & Al-Razgan, M. S. (2016). Predicting students' performance through classification: A case study. Journal of Theoretical and Applied Information Technology, 88(1), 1–8.

[12] Zafra, A., & Ventura, S. (2009). Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data. Journal of Information Sciences, 180(6), 1106–1124.

[13] Vapnik, V. (1998). *Statistical learning theory*. Wiley.

[14] Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early

alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics, 1*(1), 6–47.

[15] Gray, G., McGuinness, C., & Owende, P. (2014). An application of classification models to predict learner progression in tertiary education. In *Proceedings of the 2014 International Conference on Educational Technologies* (pp. 172–177).

[16] Zhang, Y., & Rangwala, H. (2018). Grade prediction with temporal course-wise influence. In *Proceedings of the 9th International Conference on Educational Data Mining (EDM)* (pp. 39–48).

[17] Kabra, R., & Bichkar, R. (2011). Performance prediction of engineering students using decision trees. *International Journal of Computer Applications, 36*(11), 8–12.

[18] Binns, R., Veale, M., Van Kleek, M., & Shadbolt, N. (2018). 'It's reducing a human being to a percentage': Perceptions of justice in algorithmic decisions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14).

[19] Aher, S. B., & Lobo, L. M. R. J. (2013). Combination of machine learning algorithms for recommendation of courses in E-learning system based on historical data. *International Journal of Computer Applications, 39*(1), 48–52.

[20] Temraz, M. (2020). Comparative analysis of ensemble learning techniques for student performance prediction. *International Journal of Advanced Computer Science and Applications, 11*(9), 167–174.

[21] Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society, 17*(4), 49–64.

[22] You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education, 29, 23–30.