Analyzing Final Grade Anomalies in Academic Performance Data

Sachin Don Siman
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
sachinshehan20@gmail.com

Nethmi Wijesinghe
Department of Computing
Infromatics Institute of Technology
Colombo, Sri Lanka
nethmi.wi@iit.ac.lk

Keywords—Anomaly Detection, Time-Series Analysis, Transformer Autoencoder, Machine Learning, Multi-Head Self-Attention

I. Introduction

Academic performance data, especially final grade records, is critical for measuring student achievement, informing policy decisions, and maintaining the integrity of educational assessments. However, anomalies within these records whether due to data entry mistakes, inconsistencies in grading criteria, or genuine outliers in student performance can adversely affect data quality and fairness. Identifying such anomalies efficiently and accurately is thus essential for ensuring reliability in academic evaluations.

This research focuses specifically on detecting anomalies within final grade datasets. By applying advanced anomaly detection methods, the study aims to achieve three primary objectives:

A. Data Quality Assurance

Identify and rectify records affected by entry errors or inconsistencies.

B. Enhanced Fairness

Highlight potential irregularities in grading practices to ensure equitable treatment of students.

C. Actionable Insights

Provide educators and administrators with actionable insights regarding anomalies, facilitating timely and informed interventions.

The significance of this work lies in its contribution to maintaining rigorous standards of data quality, transparency, and equity within educational institutions, ultimately promoting trust and reliability in academic assessments.

II. Content or Headings

A. Introduction

Student performance prediction remains a vital area of research in educational data mining, as accurate forecasting of at-risk learners enables timely interventions and improved academic outcomes. Traditional Learning Management Systems (LMS) often lack the capability to dynamically identify and support students who are struggling. Recent advances in deep learning particularly Long Short-Term Memory (LSTM) networks and Transformer-based models have demonstrated promising results in capturing the temporal and sequential nature of student interactions. This extended abstract focuses on the development of a grade prediction model for an online learning context, leveraging the insights from existing literature on student grade prediction and the implementation details of the Multi-Head Self-Attention Autoencoder with multiple layers and Dense bottleneck model.

B. Literature Review

Research on student grade prediction has frequently employed time-series and deep learning approaches to model student engagement and performance. LSTM networks excel at learning from sequential data, such as weekly clickstreams or continuous assessments, by retaining long-range dependencies [1]. More recent studies have introduced multilayer LSTM architectures enhanced with multi-head self-attention [2], aiming to integrate additional features such as demographic data and assessment histories. Although these advanced models often outperform traditional classifiers (e.g., logistic regression, random forests), challenges remain in handling imbalanced datasets, ensuring generalizability, and providing transparent interpretations of predictions [3].

To address these limitations, contemporary research has explored anomaly detection methods and reconstruction error metrics. By training an autoencoder solely on "normal" students, the model learns typical engagement patterns and flags deviations as potential at-risk cases. Integrating these anomaly detection approaches with LSTM or Transformer architectures can further enhance the early detection of underperforming students, enabling proactive interventions.

C. Materials and Methods

We employed the Open University Learning Analytics Dataset (OULAD)[4] as the primary data source, focusing on time-series data that was aggregated on a weekly basis, such as click counts and submission patterns. Demographic and registration data were merged with the activity records to label students who ultimately failed or withdrew as anomalies. To prepare the data for model training, all features were normalized using MinMaxScaler, ensuring consistent input ranges. Furthermore, students were grouped by their unique course enrollments, and each student's sequence was zero-padded to a fixed length corresponding to the maximum observed week, thereby standardizing the input sequence lengths.

https://doi.org/10.31705/ADScAI.2025.48

The proposed model is a Multi-Head Self-Attention Autoencoder with multiple layers and a Dense bottleneck, specifically designed for anomaly detection. This architecture integrates multi-head self-attention (MHA) blocks in both the encoder and the decoder. In the encoder, input sequences are compressed into a lower-dimensional latent space that captures the underlying structure of normal student behavior. The decoder then reconstructs the original input sequence from this latent representation. The model is trained by minimizing the mean squared error (MSE) between the reconstructed output and the original sequence. After training on data from normal students, the reconstruction error for each sample is computed, and a threshold set using a specific percentile of the reconstruction errors is applied to classify anomalies.

Implementation of the model was carried out using the TensorFlow/Keras framework. Key hyperparameters were chosen to optimize performance, including setting the model dimension (d_model) to 64, the number of attention heads (num_heads) to 4, the feed-forward dimension (ff_dim) to 128, and the dropout rate to 0.1. Additionally, early stopping was employed during training, terminating the process if the validation loss did not improve for 10 consecutive epochs, thereby ensuring efficient convergence and reducing the risk of overfitting.

D. Results and Discussion

After training exclusively on normal student data, the model was evaluated on a combined set of normal and anomalous sequences, yielding key performance metrics. As shown in Fig. 1. Confusion Matrix of the Multi-Head Self-Attention Autoencoder, the confusion matrix underscores the model's effectiveness in detecting at-risk students while maintaining a reasonable balance between precision and recall. Notably, the chosen threshold was 0.025132285503226395, and the model achieved a precision of 0.927, recall of 0.739, and an F1-score of 0.817. These figures underscore the model's effectiveness in detecting at-risk students while maintaining a reasonable balance between precision and recall.

Interpreting these metrics reveals that a precision of 0.927 means 92.7% of flagged students are genuinely at risk, while a recall of 0.739 signifies that 73.9% of all at-risk students were correctly identified. The F1-score of 0.817 indicates a robust balance between these two measures. Within the confusion matrix, the true positives (TP = 10607) demonstrate a substantial number of correctly identified at-risk students, while false negatives (FN = 3923) highlight the segment of students who were overlooked. False positives (FP = 837) remain relatively low, reflecting the high precision of the model.

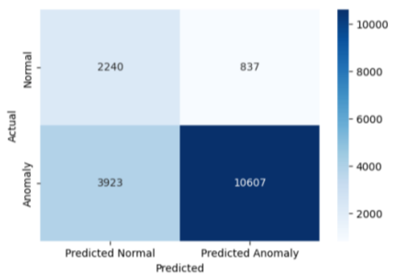

Fig. 1. Confusion Matrix of the Multi-Head Self-Attention Autoencoder

These findings suggest that the multi-head self-attention autoencoder effectively learned the normal engagement patterns of students, flagging deviations as potential anomalies. This aligns with existing research on attention-based models, which excel at capturing complex, time dependent relationships in sequential data. Future work could involve fine-tuning hyperparameters and integrating richer features, such as social interaction data, to further enhance the model's detection capabilities.

E. Conclusion

This extended abstract demonstrates how a Transformer-based autoencoder, focused on anomaly detection, can effectively predict at-risk students in an LMS context. Building on insights from the literature on student grade prediction specifically time-series modeling with LSTM and attention mechanisms our approach successfully identifies patterns of normal versus anomalous student engagement. The resulting precision (0.927) and F1-score (0.817) underscore the model's robustness, although a recall of 0.730 indicates room for improvement in detecting all struggling students.

Future work could explore more advanced thresholding strategies, additional features (e.g., discussion forum sentiment, peer interaction data), and interpretability frameworks (e.g., SHAP values) to provide clearer insights into why certain students are flagged. Overall, these findings support the feasibility of using anomaly detection with Transformer architectures for timely interventions in student performance prediction, ultimately helping educators tailor support and resources to maximize student success.

Acknowledgment

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

References

[1] N. R. Aljohani, A. Fayoumi, and S.-U. Hassan, "Predicting At-Risk Students Using Clickstream Data in the Virtual Learning Environment," *Sustainability*, vol. 11, no. 24, p. 7238, Dec. 2019, doi: 10.3390/su11247238.

[2] Z. Shou, M. Xie, J. Mo, and H. Zhang, "Predicting Student Performance in Online Learning: A Multidimensional Time-Series Data Analysis Approach," *Applied Sciences*, vol. 14, no. 6, p. 2522, Mar. 2024, doi: 10.3390/app14062522.

[3] Y. Liu, S. Fan, S. Xu, A. Sajjanhar, S. Yeom, and Y. Wei, "Predicting Student Performance Using Clickstream Data and Machine Learning," *Education Sciences*, vol. 13, no. 1, p. 17, Dec. 2022, doi: 10.3390/educsci13010017.

[4] M. H. Jakub Kuzilek, "Open University Learning Analytics dataset." UCI Machine Learning Repository, 2015. doi: 10.24432/C5KK69.