Oleinikova Olesia Olegovna
Student
FSBEI of HE “Russian Timiryazev State Agrarian University”
Moscow
Korshunova Veronika Iurevna
Master's student
FGBOU VO "Sibirskii gosudarstvennyi universitet nauki
i tekhnologii im. akademika M.F. Reshetneva"
Krasnoyarsk, Krasnoyarsk region
Scientific adviser
Kostina Raisa Vasilevna
Candidate of Economic Sciences, Associate Professor
FSBEI of HE “Russian Timiryazev State Agrarian University”
Moscow

APPLICATION OF LOGICAL CLASSIFICATION ALGORITHMS FOR PREDICTING STUDENT DROPOUT IN SMART EDUCATIONAL ENVIRONMENTS

Abstract: *this study investigates the application of decision trees for predicting student dropout and identifying academic risk factors based on Learning Management System (LMS) attributes. A logical classification algorithm searching for conjunctive patterns was applied to a dataset comprising 112 structural and behavioral features of student engagement. The decision tree model achieved an error rate of 4.12%. Key significant factors identified include the frequency of login gaps (40.3%), time since initial enrollment (16.8%), and the number of incomplete module directories (14.06%). Comparative analysis demonstrates that rule-based methods outperform neural networks in interpretability, providing actionable insights for educators. The results confirm the high efficiency and transparency of logical classification algorithms for real-time educational analytics.*

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

Keywords: *educational data mining, student dropout prediction, machine learning, decision tree, logical classification, smart campus.*

Олейникова Олеся Олеговна
студентка
ФГБОУ ВО «Российский государственный
аграрный университет – МСХА им. К.А. Тимирязева»
г. Москва
Коршунова Вероника Юрьевна
студентка
ФГБОУ ВО «Сибирский государственный университет
науки и технологий им. академика М.Ф. Решетнева»
г. Красноярск, Красноярский край
Научный руководитель
Костина Раиса Васильевна
канд. экон. наук, доцент
ФГБОУ ВО «Российский государственный аграрный
университет – МСХА им. К.А. Тимирязева»
г. Москва

ПРИМЕНЕНИЕ ЛОГИЧЕСКИХ АЛГОРИТМОВ КЛАССИФИКАЦИИ ДЛЯ ПРОГНОЗИРОВАНИЯ ОТЧИСЛЕНИЯ СТУДЕНТОВ В ИНТЕЛЛЕКТУАЛЬНЫХ ОБРАЗОВАТЕЛЬНЫХ СРЕДАХ

Аннотация: *в статье рассматривается применение деревьев решений для прогнозирования отчисления студентов и выявления академических факторов риска на основе атрибутов системы управления обучением (LMS). Логический алгоритм классификации, осуществляющий поиск конъюнктивных закономерностей, был применен к набору данных, включающему 112 структурных и поведенческих признаков вовлеченности студентов. Модель дерева решений достигла уровня ошибки 4,12%. Ключевыми значимыми факторами, выявленными в ходе исследования, стали: частота перерывов между входами в систему (40,3%),*

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

время, прошедшее с момента зачисления (16,8%), и количество незавершенных разделов модулей (14,06%). Сравнительный анализ показывает, что методы, основанные на правилах, превосходят нейронные сети по степени интерпретируемости, предоставляя преподавателям практически применимые выводы. Полученные результаты подтверждают высокую эффективность и прозрачность логических алгоритмов классификации для аналитики образовательного процесса в реальном времени.

Ключевые слова: *интеллектуальный анализ образовательных данных, прогнозирование отчисления студентов, машинное обучение, дерево решений, логическая классификация, умный кампус.*

1. Introduction: in the modern digital educational landscape, retaining students and ensuring academic success remains a critical challenge for higher education institutions. Smart educational environments and Learning Management Systems (LMS) generate vast amounts of data reflecting student engagement, behavioral patterns, and academic progress. Traditional threshold-based alert systems are increasingly ineffective due to the complex, non-linear nature of student disengagement; students often exhibit subtle signs of attrition long before formally dropping out. Consequently, heuristic and machine learning (ML) approaches have become critical components of modern educational analytics.

Among various ML algorithms, decision trees offer a unique advantage for educational applications: interpretability. Unlike «black-box» models such as deep neural networks, decision trees produce explicit, human-readable rules that educators and administrators can audit, validate, and integrate into student support policies. This transparency is essential for understanding the underlying causes of academic failure and designing targeted pedagogical interventions.

2. Materials and Methods: the dataset encompasses an extensive compilation of attributes extracted from LMS logs, tailored to streamline the identification of at-risk students. Encompassing 112 columns, the dataset includes metrics such as login fre-

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

quency, assignment submission times, forum participation, and resource access patterns. The output parameter is a binary dropout indicator (0 = active, 1 = at-risk/dropout).

Data preprocessing included filling missing values, eliminating outliers, and correlation analysis with a threshold of 0.05 to remove insignificant factors. The primary algorithm employed is a logical classification method based on the search for conjunctive patterns, effectively implemented via decision trees. This method constructs a hierarchical set of IF-THEN rules, which is particularly well-suited for educational tasks where risk factors often manifest as specific combinations of behavioral attributes rather than isolated metrics.

3. Results and Discussion: the constructed decision tree model demonstrated high predictive accuracy, yielding an error rate of merely 4.12% on the test set. Using the «Attribute Significance» metric, the most critical factors influencing student retention were identified:

1. *Frequency of login gaps (qty_inactive_days):* Contributing 40.3% to the model. Prolonged periods of inactivity strongly correlate with disengagement and subsequent dropout.

2. *Time since initial enrollment (time_since_enrollment):* Accounting for 16.8%. The first few months of study are critical, and early behavioral anomalies are strong predictors of attrition.

3. *Incomplete module depth (incomplete_assignment_depth):* With a 14.06% weight. Accumulating unfinished tasks creates a psychological barrier that often leads to course abandonment.

Unlike neural networks, which require significant computational resources and lack transparency, the logical classification algorithm allows university administrators to extract explicit rules (e.g., «IF inactive_days > 5 AND incomplete_depth > 3 THEN risk = HIGH»). These rules can be directly embedded into LMS dashboards to trigger automated alerts for academic advisors.

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

4. Conclusion: the low error rate and high interpretability of the decision tree method make it a powerful tool for educational data mining. The identified factors allow universities to implement automated, rule-based early warning systems without requiring complex GPU infrastructure. Future research will focus on integrating these models into real-time LMS environments to provide instant, personalized feedback to students and pedagogical staff.

References

1. Development of an algorithm for automatic classification of lands disturbed by mining operations using remote sensing data / O. S. Ermolaeva, A. A. Stupina, S. I. Gorbachev, G. V. Korableva // Proceedings of the Tula State University. Sciences of Earth. – 2026. – No. 3. – P. 167–172. – EDN CTQESR.

2. Kleshko I. I., Vakhrusheva I. A., Korableva G. V. Application of machine learning algorithms for detecting anomalous readings of open-pit mine dewatering sensors // Proceedings of the Tula State University. Sciences of Earth. – 2026. – No. 3. – P. 235–248. – EDN QCBKUR.

3. Korableva G., Kucherova E. Adaptive method of generation of optimal production orders // E3S Web of Conferences: 2018 International Science Conference on Business Technologies for Sustainable Urban Development, SPbWOSCE 2018, St. Petersburg, December 10–12, 2018. – St. Petersburg: EDP Sciences, 2019. – Vol. 110. – P. 02127. – DOI 10.1051/e3sconf/201911002127. – EDN CFFBTL.

4. An intelligent machine learning model for forecasting waste generation volumes / T. A. Panfilova, V. S. Tynchenko, B. V. Potapov, M. S. Nikanorov // Sustainable Development of Mountainous Territories. – 2026. – Vol. 18, No. 2 (68). – P. 1107–1120. – DOI 10.21177/1998-4502-2026-18-2-1107-1120. – EDN GTTEPG.

5. Panfilov I. A., Zhuravlev M. V., Malykha E. F. A neural network method of early detection of pre-emergency overloads in mine distribution networks // Occupational Safety in Industry. – 2026. – No. 7. – P. 7–13. – DOI 10.24000/0409-2961-2026–7-7-13. – DOI 10.24000/0409-2961-2026-7-7-13 – EDN DTSUHM.

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

6. Matasova I. Yu., Potapov B. V., Stepantsevich M. N. Features of distribution of a number of chemical elements in rocks of the Black Sea coast of Russia // Proceedings of the Tula State University. Sciences of Earth. – 2026. – No. 2. – P. 356–361. – EDN ZTXOZP.

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)