Journal of Engineering Science and Technology
Special Issue on INCITEST2025 Vol. 21, No. 2 (2026) 81 - 88


# DESIGNING A TRANSPARENT AI SYSTEM USING SHAP AND LIME TO IDENTIFY THE DOMINANT PREDICTOR OF STUDENT STUDY INTEREST

HANI IRMAYANTI[1,*], MEITA MAHARANI ISKANDAR[1],
TELMAN SADULLAYEVICH ACHILOV[2]

[1]Universitas Komputer Indonesia, Bandung, Indonesia
[2]Chirchik State Pedagogical University, Uzbekistan
[*]Corresponding Author: haniirmayanti@email.unikom.ac.id

**Abstract**

This study is to interpret the transparency of an explanation model (Artificial Intelligence, AI) in predicting student's interest on studying further education. Specifically, the goal of this work is to identify and describe the causes that lead a new feature extracted by a Decision Tree C4. 5 based predictive application to be understood by the final user. The methodology used is based on explainable artificial intelligence (XAI) and consists in a double approach: shapely additive explanation (SHAP) for global feature contribution, plus local interpretable model-agnostic explanations (LIME) for local explanation. This provides the positive result that an explanation technique for the AI model prediction has been developed, allowing us to move beyond "black-box" nature of the predictions. In the prediction, the extracurricular activity was found as the biggest contributor (SHAP = 0.182 and LIME = 0.391) then IQ and GPA. SHAP provides a detailed summary whereas individual cases are analysed through LIME. The XAI analysis particularly provides evidence that extracurricular activities constitute an "invisible factor" over academic potential. Such transparency frameworks XAI more responsibly and reliably than other predictive AI systems. More than anything, it provides a solid basis for tailored academic intervention and career guidance planning.

Keywords: Artificial intelligence, Decision tree C4.5, LIME, SHAP, Xpalanable AI.

## 1.Introduction

Choosing what to study is an important step that influences students' educational journeys and career achievements [1, 2]. Schools are beginning to implement Artificial Intelligence (AI) to try and predict what students will be interested in studying, looking at things like GPAs, IQs, and things like extracurriculars [2-5]. The combination of such types of data presents a problem of feature selection on heterogeneous data that begs a robust hybrid out of the box machine learning approach [6]. This trend is a good fit to the new business model for the private higher education in the digital age, where data fuelled services are at the core of improving the competitiveness and viability of an educational institution [7].

Nevertheless, Predictive Models' "Black Box" Issue results in end users - in the case, guidance counsellors, and students alike - not understanding how targets are decided and the recommendations made, and the resulting predictions are out of those recommendations [8-11]. The loss of such information contributes to eroding users' trust and undermining active planning for a meaningful academic intervention [12, 13]. This is the problem Designing a Transparent AI System strives to solve, focusing on explainable artificial intelligence (XAI) [14]. The purpose of deploying models based on the SHAP and LIME methods is to achieve transparency and accountability concerning decisions made by the models, and we be better dindan Rhemain to ignore the models modified by the users [15-17].

Previous researchers have previously recognized the contributions Explainable AI (XAI) have made to the conversations on fairness, accountability, and transparency in AI-assisted educational interventions [9, 18]. Various studies have incorporated XAI in the analysis of several models aimed at student adaptability predictions [19]. Other studies analysed the integration of predictive models with the SHAP technique to assess the various predictors of academic achievement [20, 21]. SHAP and LIME have also been reported to aid in the analysis of engineering domains predictive models [17]. Other studies have also demonstrated SHAP and LIME as significant approaches for black box problems and in producing clearer output for enhanced imbalanced classification models [22-24].

The purpose of this study is to construct a Transparent AI System whose purpose is to enhance the understanding of AI prediction models designed with the C4.5 Decision Tree algorithm to unpick the Dominant Predictors of student study interests [11]. This makes the prediction process understandable to end users. This is accomplished with a dual-XAI approach with SHAP and LIME [16, 17]. This research is novel first for its emphasis on systems engineering and the design and validation of a Transparent AI System framework with the embedding of SHAP and LIME modules and of the Decision Tree prediction model. Second, for dual method validation, this is the first of its kind systematic comparative validation of the SHAP-LIME dual approach on the C4.5 algorithm, which offers complementary global and local perspectives. Third, with Real Action Insights, this is the first of its kind contribution in that it demonstrates the explicit identification of Dominant Predictors and the provision of significant macro-to-micro real action insights to frame intervention policy design beyond the default academic metrics.

## 2. Research Methodology

The present research methodology submits a Transparent AI System Engineering instrument, which is aimed at adding Explainable AI (XAI) modules into the study interest prediction system. The stages of the research are shown in Fig. 1, the stages being: Preparing data, initial prediction model development, transparent XAI system design, interpretation of results and identification of dominant [25, 26].

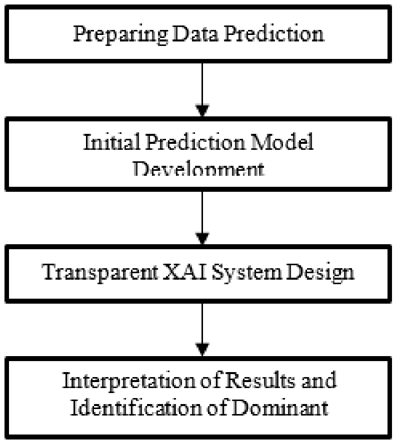


**Fig. 1. Research stages.**

## 3. Results and Discussion

### 3.1. Results

### 3.1.1. Preparing data prediction

The crude C4.5 prediction component is generally accurate to about 68% (see Table 1). The mismatch between the performances (recall Class 1 (Engineering) only 41.7%) is the reason why an XAI layer for explanation of the decisions made is absolutely necessary.

**Table 1. Results prediction.**

| | Precision | Recall | f1-score | Support |
|---|---|---|---|---|
| **0** | 0.631579 | 0.923077 | 0.750000 | 26.00 |
| **1** | 0.833333 | 0.416667 | 0.555556 | 24.00 |
| **accuracy** | 0.680000 | 0.680000 | 0.680000 | 0.68 |
| **macro avg** | 0.732456 | 0.669872 | 0.652778 | 50.00 |
| **weighted avg** | 0.728421 | 0.680000 | 0.656667 | 50.00 |

### 3.1.2. Implementation of the initial prediction model development

The first-stage prediction system is carried out in a user interface as shown in Fig. 2. The interface takes the values of the attributes "Extracurricular Activities (Sports)", "IQ Level (Gifted)", and "GPA (Low)" and produces a NON-TECHNICAL recommendation (63.64% probability). The presented system serves as a front-end, the result of which is further processed by the XAI Module.

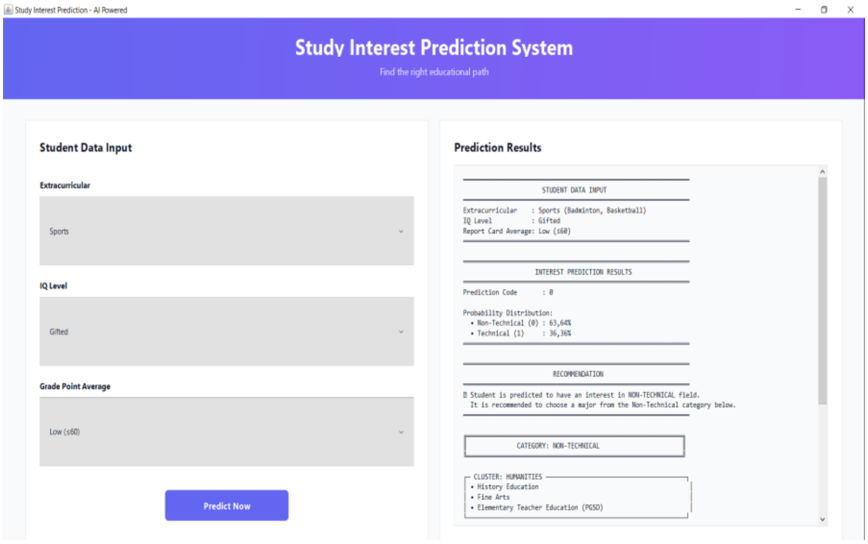


**Fig. 2. Interface application study interest prediction system.**

### 3.1.3. Transparent XAI system design

#### A. SHAP: Interpretation of the global features

The use of the SHAP module met with success in giving a broad understanding. The Bar Plot Feature Importance (see Fig. 3) illustrates that the first-hand influence of the Extracurricular Activities (0.178) is even more than that of IQ (0.104) and GPA (0.089). This result points out the role of Extracurricular Activities as the DPC (Dominant Predictor Component) in the whole dataset.

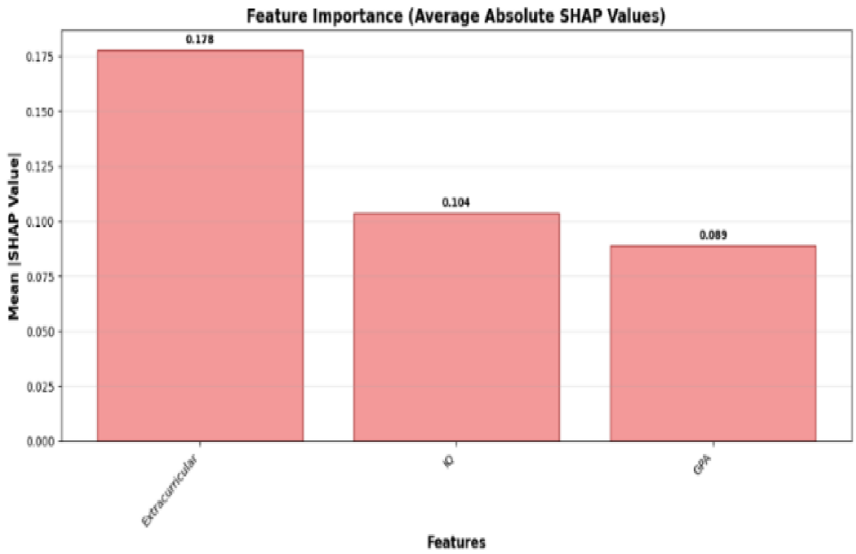


**Fig. 3. The bar plot feature importance.**

**B. LIME: Interpretation of the local prediction**

LIME generates user-friendly explanations for single cases. In the example of an occurrence 0, the main reason for the “Technical” (confidence 93.8%) prediction is Extracurricular Activities that make the most positive contribution (+0.391), whereas IQ makes a negative contribution (-0.102). The behaviour is going further deeply consistent (see Fig. 4): high extracurricular activities are powerful in leading the model to label Technical.

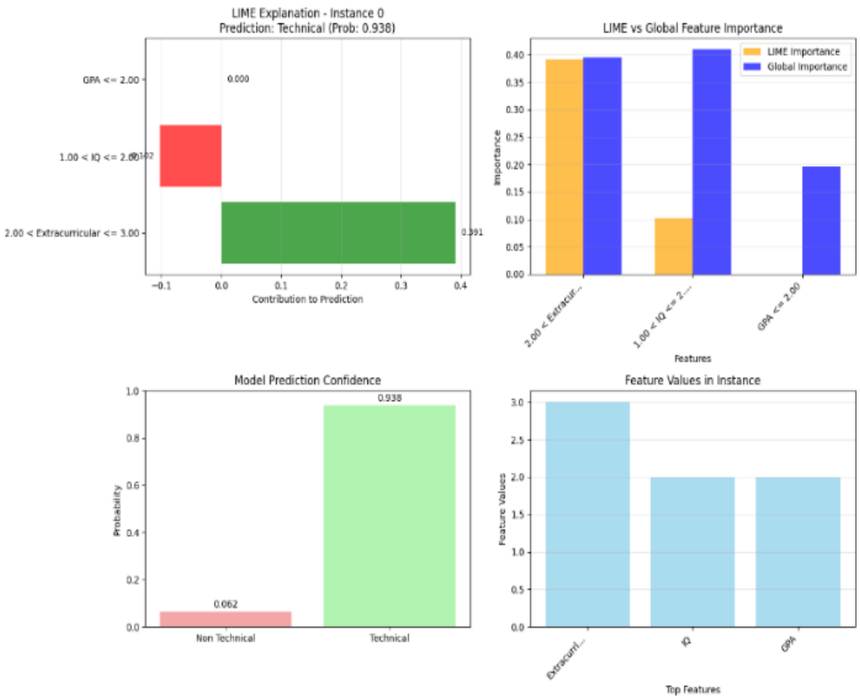


**Fig. 4. LIME explanation visualization results.**

**C. Interpretation of results and identification of dominant**

Dual-XAI integration serves as a validation tool for the Transparent AI System with the main discoveries: Out-of-school activities are the main factor (0.178 SHAP, 0.391 LIME), even going beyond academic ability as a second factor). XAI adeptly reveals these "invisible factors" that help in giving the required transparency.

## 3.2. Discussion

Dual-XAI integration is one of the key methods to substantiate the Transparent AI System through main outcomes: Extracurricular activities are the main driving factor (0.178 SHAP, 0.391 LIME), academic ability being just a secondary factor). XAI manages to reveal these "hidden factors" that are crucial for transparency. The Transparent AI System developed is a good example of how the limitations of black box models can be overcome, thus, it is in line with global ethical standards [18]. Although the C4.5 model is the one that mainly makes the predictions, the XAI module is the one that guarantees interpretation, fairness, and accountability [9].

Extracurricular Activities turned out to be the most influential factor, achieving the value of 0.182 (SHAP) and 0.391 (LIME). These figures have a deep and wide impact on Educational Data Mining (EDM), as they are evidence of non-academic factors having more weight than academic ones in the prediction process [17, 19]. The findings show that non-academic factors need to be considered in prediction models for a more thorough understanding of student tendencies [20].

The concurrent application of both SHAP and LIME was one of the main factors that not only complemented the validation of system but also contributed to the solution of black box problems [17]. The main functions of each are different SHAP (Global) Is a tool that supplies a wider view for supporting institute policy [21, 27], whereas LIME (Local): Pointing out the most important factors for career counsellors to be able to explain the reason of personal predictions [28].

The alignment of the main factors (0.182 SHAP, 0.391 LIME) is one of the elements that strengthens the trust in the system. Thanks to this transparency the stakeholders can come to the realization that academic grades, though still very important, cannot be the only focus of their attention and thus, the development of extracurricular activities as the key indicators can be initiated.

## 4. Conclusions

The study had a huge success in executing its objectives evidenced by the creation and the validation of a Transparent Artificial Intelligence (AI) System that can predict students' interest in further study which is a good way of getting over the “black box” problem of a C4.5 Decision Tree-based prediction model. A double explainable AI (XAI) approach was used to arrive at the conclusions represented here. XAI Method Validation, i.e., SHAP and LIME Integration, was the main vehicle towards the comprehensive and intuitive interpretation mechanism. SHAP summarizes feature importance at a global level, while LIME gives detailed explanations for single instances in the form of readily understandable visualizations (green-red).

Beyond that, the identification of the dominant predictors in particular, XAI analysis, discloses and corroborates that the extracurricular activities are a latent variable that is the largest dominant contributor to the prediction results with influence values being very high both at the global level (SHAP: 0.182) and at the local level (LIME: 0.391) - quite consistent. Conventional academic factors (IQ and GPA) are in the background playing a secondary role. The experiential implication is that the transparency afforded by XAI is a powerful and responsible platform that stakeholders can use to plan academic interventions and career guidance that would mainly focus on the development of non-academic skills thus, resulting in an increase in the trust and accountability of the AI prediction system.

## References

1. Guleria, P.; and Sood, M. (2023). Explainable AI and machine learning: Performance evaluation and explainability of classifiers on educational data mining inspired career counseling. *Education and Information Technologies*, 28(1), 1081-1116.
2. Sunge, A.S.; Abdurachman, E.; Heryadi, Y.; and Kartowisatro, I.H. (2023). Interpretable using feature importance, SHAP, and LIME for Caesar’s

prediction. *International Journal of Applied Engineering and Technology*, 5(2), 83-89.

3. Ginting, S.L.B.; and Rahman, M.A.F. (2021). Data mining, neural network algorithm to predict student's grade point average: Backpropagation algorithm. *Journal of Engineering Science and Technology*, 16(3), 2028-2037.
4. Nagy, M.; and Molontay, R. (2024). Interpretable dropout prediction: towards XAI-based personalized intervention. *International Journal of Artificial Intelligence in Education*, 34(2), 274-300.
5. Eteng-Uket, S.; and Njaka, R.C. (2023). Predictive power of grit and emotional intelligence on research anxiety experience of students. *American Journal of Education and Learning*, 8(1), 63-75.
6. Nursikuwagus, A.; Purwanto, H.; and Dewi, D.A. (2023). Feature classification based on heterogenous data using hybrid machine learning: A review. *Journal of Theoretical and Applied Information Technology*, 101(17), 6991-7010.
7. Luckyardi, S.; Suprayogi, Y.; Soegoto, E.S.; and Pesakovic, G. (2024). A business model for private higher education in a digitalized era. *Australasian Accounting, Business and Finance Journal*, 18(1), 130-155.
8. Fadillah, A.P.; Fachrizal, M.R.; and Reski, M. (2023). Decision support system to assist students in choosing thesis research topics. *Journal of Engineering Science and Technology*, 18(5), 2432-2443.
9. Khosravi, H.; et al. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074-100095.
10. Sokol, K.; and Flach, P. (2020). One explanation does not fit all: The promise of interactive explanations for machine learning transparency. *KI - Künstliche Intelligenz*, 34(2), 235-250.
11. Ali, O.; Murray, P.A.; Momin, M.; Dwivedi, Y.K.; and Malik, T. (2024). The effects of artificial intelligence applications in educational settings: Challenges and strategies. *Technological Forecasting and Social Change*, 199, 123076-123093.
12. Danish, M. (2025). Enhancing cyber security through predictive analytics: Real-time threat detection and response. *International Journal of Advanced Computer Science and Applications*, 16(8), 38-49.
13. Alfredo, R.; Echeverria, V.; Jin, Y.; Yan, L.; Swiecki, Z.; Gašević, D.; and Martinez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence*, 6, 100215-100229.
14. Johora, F.T.; Hasan, M.N.; Rajbongshi, A.; Ashrafuzzaman, M.; and Akter, F. (2025). An explainable AI-based approach for predicting undergraduate students academic performance. *Array*, 26, 100384-100406.
15. Airlangga, G. (2024). Decoding energy usage predictions: An application of XAI techniques for enhanced model interpretability. *Indonesian Journal of Artificial Intelligence and Data Mining*, 7(2), 275-284.
16. Hasan, M.M. (2023). Understanding model predictions: a comparative analysis of SHAP and LIME on various ML algorithms. *Journal of Scientific and Technological Research*, 5(1), 17-26.

17. Parisineni, S.R.A.; and Pal, M. (2024). Enhancing trust and interpretability of complex machine learning models using local interpretable model agnostic shap explanations. *International Journal of Data Science and Analytics*, 18(4), 457-466.
18. Chander, B.; John, C.; Warrier, L.; and Gopalakrishnan, K. (2025). Toward trustworthy artificial intelligence (TAI) in the context of explainability and robustness. *ACM Computing Surveys*, 57(6), 1-49.
19. Nnadi, L.C.; Watanobe, Y.; Rahman, M.M.; and John-Otumu, A.M. (2024). Prediction of students' adaptability using explainable AI in educational machine learning models. *Applied Sciences*, 14(12), 5141-5160.
20. Wang, Y.; Hu, L.; Hou, L.; Wang, L.; Chen, J.; He, Y.; and Su, X. (2024). A SHAP machine learning-based study of factors influencing urban residents' electricity consumption-evidence from Chinese provincial data. *Environment, Development and Sustainability*, 26(12), 30445-30476.
21. Fu, H.; Zhou, X.; Xu, P.; and Sun, D. (2025). Prediction of compressive strength of concrete using explainable machine learning models. *Materials*, 18(21), 5009-5018.
22. Widyanto, N.; Kim, J.W.; and Nursikuwagus, A. (2024). Best classifier over feature selection and delta error in time series dataset. *Journal of Engineering Science and Technology*, 19(2), 390-404.
23. Purnomo, A.; Putrada, A.G.; Habibi, R.; and Syafrianita. (2024). MDI and PI XGBoost regression-based methods: Regional best pricing prediction for logistics services. *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, 22(5), 1157-1166.
24. Fida, S.; Masood, N.; Tariq, N.; and Qayyum, F. (2022). A novel hybrid ensemble clustering technique for student performance prediction. *Journal of Universal Computer Science*, 28(8), 777-798.
25. Molnár, G.; and Kocsis, Á. (2024). Cognitive and non-cognitive predictors of academic success in higher education: A large-scale longitudinal study. *Studies in Higher Education*, 49(9), 1610-1624.
26. Almasri, A.R.; Yahaya, N.A.; and Abu-Naser, S.S. (2022). Instructor performance modeling for predicting student satisfaction using machine learning-preliminary results. *Journal of Theoretical and Applied Information Technology*, 100(19), 5481-5496.
27. Mingyu, Z.; Sutong, W.; Yanzhang, W.; and Dujuan, W. (2022). An interpretable prediction method for university student academic crisis warning. *Complex and Intelligent Systems*, 8(1), 323-336.
28. Nurhayati, S.; Wicaksono, M.F.; Syahrul; and Effendi, D. (2024). Prediction system for vitamin A requirements in West Java Indonesia using linear regression. *Journal of Engineering Science and Technology*, 19(1), 55-64.