International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 15, Issue 03, March 2026)

An Interpretable Multi-Agent Framework for Student Performance Prediction and LLM-Based Academic Advising

Atharv Dhole[1], Vaibhav Gade[2], Seema Chowhan[3], Sudarshan Lakhdive[4]

[1,2,3,4]*Department of Computer Science, Baburaoji Gholap College, Sangvi, Pune, Maharashtra, India*

***Abstract*—The fast development of online and blended learning environments has allowed educational institutions to gather large amounts of academic and behavioral data. The successful use of this data is critical to identifying at-risk students during their early stage of academic growth and contributing to the effective and timely intervention. But most current learning analytics and educational data mining methods are based on black box models of machine learning, which are only predicative but have small amounts of interpretability and little actionable advice. deficiency makes them less useful in real-life contexts of academics.**

The presented paper introduces a student performance prediction model based on an interpretable multi-agent framework and a controlled academic advisory system on the basis of a significant language model (LLM). The proposed architecture uses a modular architecture whereby special agents will perform data preprocessing, engagement analysis, performance prediction, and generation of advisory. A regression-based predictive model of academic and engagement-related features, namely the logistic regression, is used to produce probabilistic and interpretable academic risk responses in accordance with categories, i.e. Pass, At-Risk, and Fail.

To resolve this disparity between analytic information and guidance, prediction results are organized into a specialized direction at a central planning point into governable inputs to an advisory agent that can be driven by a large language model. Notably, the LLM can be applied on a highly advisory basis and is not involved in the prediction results, which guarantees reliability, transparency, and responsible use of AI.

Experimental analysis shows that the combination of the engagement indicators with the classic academic characteristics raises the predictive consistency, and results in an improved early risk identification. The suggested multi-agent architecture is scalable, interpretable, and ethically deployable, which provides a valuable framework of student-centred academic intervention and decision support in contemporary learning settings.

***Keywords*—Large Language Model, Multi-Agent System, Logistic Regression, Engagement Analysis, Learning Management System, Student Performance Prediction, Academic Advising, Responsible AI.**

I. Introduction

Due to the emergence of online learning sites, colleges can now collect enormous data on the performance of students, their attendance, and the online activities. The analysis of this data can make it possible to understand the learning behavior and identify the risk of failure among students, and making interventions at the right time. Consequently, learning analytics and educational data mining have become important approaches for improving educational outcomes [1], [4].

The existing literature on the same has largely used machine learning algorithms to predict performance of students based on their academic, demographic or learning management system data. Although such methods may achieve high predictive accuracy, many rely on complex models that lack interpretability [3], [5]. This renders them inapplicable in fields of practical use since both teachers and learners can struggle to make out predictions and formulate viable support plans.

Engagement indicators such as attendance, participation, and LMS usage have been shown to correlate strongly with student performance [2], [5]. But the majority of predictive models do not even focus on personalized support measures and just predict the risks. Conversely, recent advances in Large Language Models (LLMs) have allowed offering natural-language academic feedback and advising. However, the direct use of large language models in education raises concerns regarding reliability, interpretability, and responsible deployment [9], [11], [12].

In order to address these weaknesses, this paper suggests an interpretable multi-agent model of student performance prediction and academic advising. The developed system is a combination of engagement analysis, interpretable machine learning, and a controlled LLM-based advisor agent implemented in a modular system. The main innovations of this paper include the design of a multi-agent system to facilitate the early academic risk identification, the use of a clear-cut predictive regression model that is weakly supervised, as well as the combination of a locally deployed LLM that generates customized and privacy-aware academic guidance based on analytical data.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 15, Issue 03, March 2026)

II. Background And Related Work

A. Literature Review

Data-driven-based methods of student learning outcomes analysis and improvement have become the focus of increasing attention in recent years. Learning analytics and educational data mining will seek to derive actionable information based on academic data and learning management system (LMS) data in order to assist with the prediction of students at academic risk. Earlier studies show academic predictors like grades and past performance can be successfully applied in prediction of student outcomes [1], [3].

As more and more digital learning data are filled with data, researchers have integrated behavioral and engagement-related variables, such as attendance, participation and online learning activities, into predictive models [2], [5]. These research works highlight that the student engagement has a close relation with academic performance and may enhance the reliability of prediction in case of combination with the traditional academic characteristic. Nonetheless, current methods mainly concentrate on predictive accuracy and are not that many interpretable and provide limited actionable intervention support [3], [4].

Recent research has covered smart tutoring systems and agent-based learning systems to facilitate individualized learning assistance. Multi-agent systems are suggested to deal with the complexity of educational analytics by allocating the efforts of monitoring, analysis, and providing feedback to specialized agents [6], [8]. Simultaneously, Large Language Models (LLM) have been deployed to provide feedback in natural language and academic advising [7], [9]. Although these strategies demonstrate high potential, explainability, reliability, and uncontrolled advisory behavior are still a cause of concern in the context of deploying LLMs that have no formal analytical base. Moreover, adaptive multi-agent-based frameworks of academic advising have been suggested to facilitate intelligent decision-making and personal intervention in higher education settings [14].

B. Algorithm and Technologies

The suggested framework consists of interpretable machine learning, multi-agent system, and academic advisory based on LLM to assist in predicting and directing student performance.

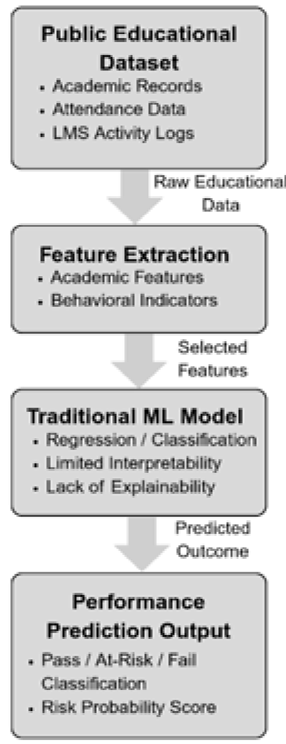

Fig.1.Traditional Student Performance Prediction Workflow.

Fig. 1 represents the conventional student performance prediction pipeline. Educational data including academic history, attendance history, and logs of LMS activities are initially loaded and analyzed via a feature extraction phase to create relevant indicators of academic and behavioral conducts. The student academic outcomes are then estimated by applying traditional machine learning models (regression or classification algorithm) which are based on the analysis of these features. Nevertheless, these systems are primarily concerned with the quality of predictions and they offer little interpretability or actionable scholarly advice.

1. *Student Performance Prediction Based on Regression:* The machine learning part of the student performance prediction for academic risk includes the use of a regression-based learning model, namely, logistic regression. Regression analysis is very much popular due to its simplicity, scale in computation and interpretability [5]. Since pass, at-risk or fail, as the possible outcomes, are categorical values, Logistic regression is a good choice in the transparent academic risk estimation since it is a good fit for the kind of dependence between the input features and the likelihood of an outcome (Pass, At-Risk or Fail).

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 15, Issue 03, March 2026)

Within the framework, the academic records and the engagement characteristics are used as the input variables whereas the probabilistic score of the risk is used as an output. The fact that the model is a probabilistic one allows the teacher to assess how predictable or predictable the driving elements are, which helps the teacher to make informed and clear academic choices. Regression based methods provide much better explanation and are more appropriate in implementing in real world as compared to the complex black box models [3].

2. *Multi-Agents System Architecture:*The system is created on the pattern of multi-agent system architecture where each agent performs a certain task enhancing modularity, scalability and transparency. The major agents are a Data and Profile Agent, Engagement Agent, Performance Prediction Agent and Advisor Agent. This separation of analytic and advisory services is guaranteed so that the predictive logic is readable, but system extension and adaption to the needs of a number of educational environments is possible [6], [10].
3. *Large Language Model Academic Advisory:*To get academic feedback and advising, the framework uses a Large Language Model which use LLaMA architecture. The LLCM is implemented in the form of only an advisory element and has no involvement with the prediction and decision-making process. It is fed with structured data which includes the status of the forecasted academic performance, level of engagement, and the probability of risks in order to come up with a customized academic advice. Such use is under control that the advisory outputs are rooted in analytic findings, and the likelihood of giving invalid or speculative recommendations is minimized [9], [11], [12]. This method of controlled deployment is consistent with established ethical design principles for autonomous and intelligent systems, which include transparency, accountability, and well-being in AI-driven applications [13].
4. *LangChain for Agent Orchestration:*The framework uses LangChain as an orchestration layer in case of a coordination of interactions between analytical agents and LLM-based advisor. LangChain helps in Structured Prompt Engineering, for managing the flow of data and Integration of LLMs with external flow of logic and analysis output.

This coordination allows the teamwork to acquire uniform, repeatable and secure production of academic recommendations and ensures firm that there is a rigid segment of separation between the predictive modeling and natural language generation.

III. PROPOSED FRAMEWORK AND METHODOLOGY

This section explains the design and implementation of the proposed multiagent framework of student performance prediction and academic advising. The framework takes a modular approach combining elements of engagement analytics, interpretable machine learning and LLM based advisory support. The separation of the analytical and advisory parts, the system promises transparency, scalability and coordination.

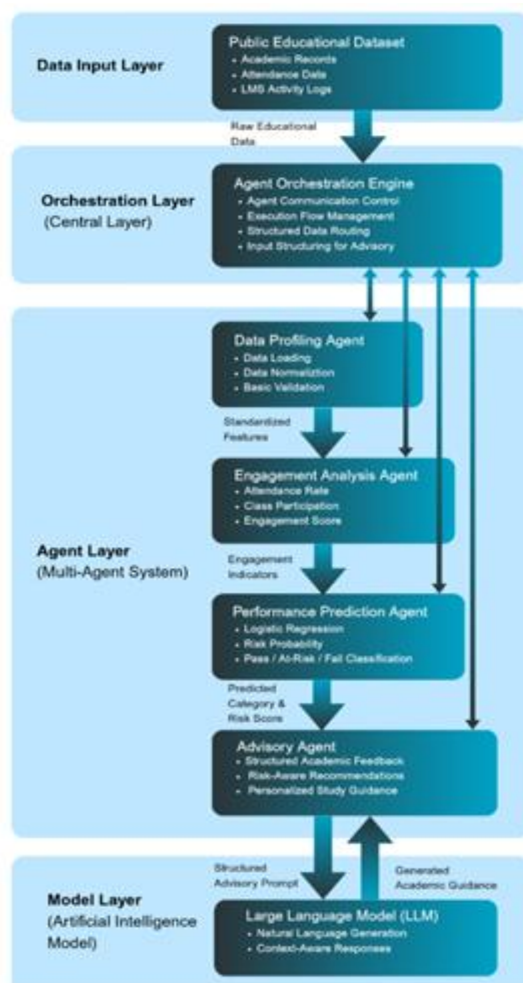

Fig.2.Multi-agent framework architecture for Student Performance Prediction.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 15, Issue 03, March 2026)

Fig. 2 illustrates this multi-agent model of predicting interpretable student performance and academic advisory. Its architecture is defined as a data input layer, orchestration layer, agent layer, and AI model layer that coordinate the processing of data and its support to make decisions. Data profiling, engagement analysis, and performance prediction are carried out by specialized agents on the basis of an interpretable logistic regression model. An advisory agent and a huge language model (LLM) then look at the predicted outcomes to produce individualized academic feedback and study recommendations without intermingling the elements of prediction and advisory.

A. Multi-Agents System Design and Workflow

The proposed framework is based on a modular architecture of multi-agents which each have a specific analytical or advisory function. This separation of responsibility provides for better system interpretability, scalability and reliability as it can be ensured that individual components can be independently validated and extended. The roles of the agents are as described below.

1. *Data and Profile Agent:*Data and Profile Agent are in charge of the raw student data received from academic records and learning management systems. Its main jobs are loading data, cleaning data, normalizing data and basic validation of data. This agent is responsible for ensuring that all the input features are standardized and appropriate for further processing while avoiding any data inconsistencies from propagating throughout the system.
2. *Engagement Agent:*The Engagement Agent calculates engagement-related indicators that are corelate with student participation and involvement in the learning process. It takes normalized behavioral characteristics such as attendance rate, class participation and online activity to create an interpretable engagement score. It is this score which allows easily a compact representation of student engagement and represents an important input for the prediction of performance.
3. *Performance Prediction Agent:*The Performance Prediction Agent uses an explainable machine learning model to predict student academic risk. The agent develops probabilistic forecasts of students performing in each category e.g. Pass, At-Risk or Fail using a technique called logistic regression. The use of a transparent model enables academic stakeholders to gain an understanding of the confidence in predictions, and contributing factors in order to support informed decision-making.
4. *Advisory Agent:*The Advisory Agent is charged to produce the personalized academic advice that relies on analytical outputs. It takes input in a structured form, e.g. predicted performance category, engagement level, risk probability, etc. and uses a Large Language Model to help it generate context-aware feedback and study recommendations. Importantly, this agent only works in an advisory role and does not affect the outcomes of predictions.
5. *Orchestration Layer:*The orchestration layer is the centralized coordination engine of the framework. It is a rule in communication between agents, execution flow, and data routing in structured data between advisory and analytical segments. Specifically, it interprets analytical results such as projected category and engagement score and risk probability into formatted notifications that are subsequently sent to the advisory agent. This controlled routing mechanism makes it reproducible, and also ensures that LLM generation is not out of control, and that there is a clear distinction between predictive modelling and natural language generation.

B. Dataset Description

The framework uses a publicly available data on education that has been used on academic and engagement related attributes. The data set will contain academic indicators in the form of grade point average (GPA) and course outcome and behavioral indicators based on learning management system (LMS) logs. The features associated with engagement will also encompass the attendance rate, the level of participation in the classroom, and the number of online learning events. Demographic attributes are not used as predictive modeling because the use of such attributes must be done in a responsible and ethical manner, in line with accepted best practices in learning analytics [1], [4]. The data is anonymized and utilized with the purpose of academic assistance and research.

C. Data Preprocessing

Data preprocessing by Data and Profile Agent is a technique of preparing raw student records for analysis purposes. Numerical features like attendance percentage, number of activities are normalized on a common scale hence we need not worry about different attributes having different scales. Missing values handled using omitting or default value assignment using the feature relevance. Categorical variables are encoded numbers wherever necessary.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 15, Issue 03, March 2026)

This is a preprocessing step that ensures that the input data is clean and standardized and ready for feature extraction and training the model [3], [5].

D. Feature Extraction

Feature extraction is interested in extracting meaningful indicators that reflect both the academic achievement and the engagement of the student. Academic features are GPA and course completion status and engagement features are calculated using normalized measures of attendance, participation, and LMS activity. These engagement indicators are combined to yield an overall engagement score; an interpretive representation of student involvement in the learning process. Prior works have found that the use of engagement features with academic measures enhances the reliability of performance prediction models [2], [5].

E. Model Training and Evaluation

First, A stratified split was used to partition the dataset into training and testing datasets so that the dataset was balanced with classes. Standard classification measures such as accuracy, precision, recall, F1-score, and confusion matrix analysis have been used to evaluate the model. This method will make sure that performance measures reported are based on generalization capabilities and not overfitting.

The Performance Prediction Agent uses a regression approach model, specifically logistic regression, to estimate the risk of student academic. Logistic regression is chosen because of its interpretability and probabilistic output properties [3], [5]. In case ground truth labels are unavailable, the weak supervision strategy is taken up, by generating performance labels with the rule-based academic criteria. The model is so trained on the mix of academic and engagement-related features to type the students as being returned to Pass, At-Risk and Fail.

Model performance is evaluated using standard classification values such as accuracy, confusion matrices and recall. These metrics can give an insight into both the overall predictive behavior and class-wise too. The probabilistic nature of logistic regression makes it possible to interpret the prediction confidence in a transparent way, which is key to academic decision-making. Evaluation results show that the integration of engagement features with academic indicators increases the consistency of the prediction, which supports the results reported in previous work on learning analytics [1], [4].

F. Advice Generation using LLM

The Advisor Agent makes use of a Large Language Model based on the LLaMA architecture to produce personalized academic advice.

The LLM functions only in an advisory capacity and does not get involved in prediction or decision-making. Structured inputs such as predicted performance category, risk probability and engagement level are given to the LLM via the orchestration layer of LangChain. LangChain handles agent coordination, prompt structuring and execution flow to provide controlled and reproducible advisory generation. This design enables the system to be able to translate analytical results into clear, actionable feedback alongside reducing the risks of ungrounded LLM outputs [9], [11], [12].

IV. Result And Discussion

The suggested multi-agent model proves to be a reliable prediction of academic risks in a way that is interpretable and transparent. The prediction model developed using regression gives coherent probabilistic results leading to easy identification of both low-risk and at-risk students, thus enabling early academic intervention. The model focuses on capturing both the behavioral and academic aspects, which affect the performance of students, by incorporating indicators associated with engagement, including the rate of attendance, class participation, and LMS activity, in addition to the academic attributes.

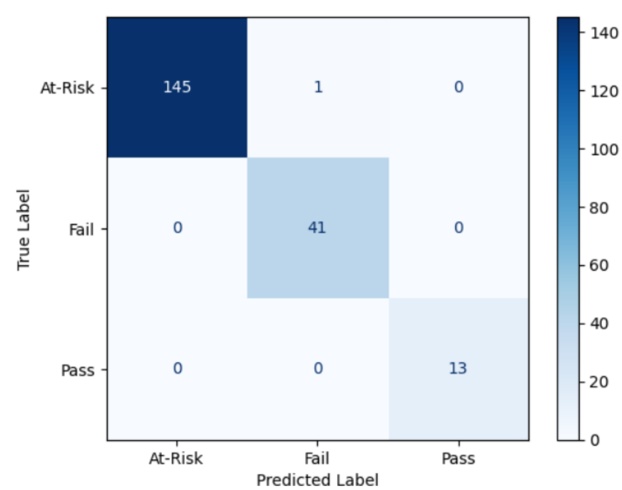

Fig.3.Confusion matrix of the Student Performance Prediction Model.

Fig. 3 indicates the distribution of the computed engagement scores and it can be seen that most students were in the moderate engagement ranges implying balanced participation patterns in the dataset and allows meaningful performance analysis.

The interpretability of the proposed framework is one of its main strengths. Regression-based approach, unlike complex black-box models, enables educators to comprehend the contribution made by individual features to the outcome of prediction. Such transparency allows teachers to examine the impacts of variables like attendance, participation and academic performance on the predicted levels of risk. Moreover, the decoupling of the analytical agents and the advisory agent makes the logic of prediction explainable and auditable, which is essential to be practically applicable to educational settings.

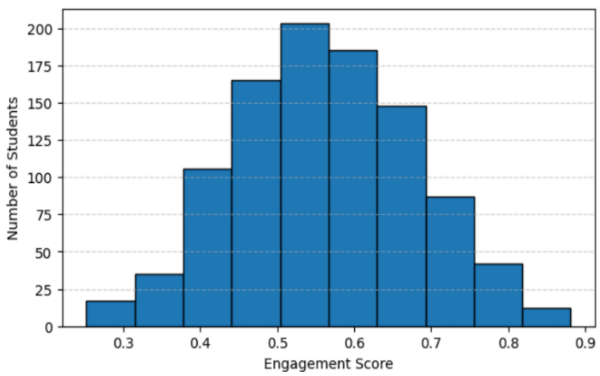

Fig.4.Distribution of normalized Student Engagement Scores.

Fig. 4 shows the confusion matrix used to classify the performance of the prediction model. The matrix will compare actual class labels with the predicted results based on the categories Pass, At-Risk and Find. The results show that most of the cases are properly grouped with a very few being misclassified in the neighbouring classes. This shows that this model is effective in capturing the relationship between engagement indicators and academic performance.

TABLE I
MODEL PERFORMANCE METRICS

Class	Precision	Recall	F1-score	Support
At-Risk	1.000	0.993	0.997	146.0
Fail	0.976	1.000	0.988	41.00
Pass	1.000	1.000	1.000	13.00
Macro Avg	0.992	0.998	0.995	200.0
Weighted Avg	0.995	0.995	0.995	200.0
Overall Accuracy	0.995	0.995	0.995	0.995

In order to examine the model performance further, Table I shows the metrics of evaluation accuracy, precision, recall, and F1 score of each category of performance. The presented prediction model demonstrates the overall accuracy of 99.5, and both precision and recall are rather steady throughout the classes. The findings also suggest that incorporating engagement functionalities boosts predictive performance at a significant rate without losing interpretability and reliability in academic risk assessment.

Additionally, to predictive accuracy, the suggested framework also offers some practical academic assistance via the advisory module that depends on a large language model. The advisory agent transforms the analytical outputs, which are structured as a prediction of the type of performance and engagement or category into individualized academic feedback and study suggestions. Notably, the LLM can only provide advice and has no effect on prediction outcomes to maintain consistency between the results of analytical work and the output in terms of advice.

System wise the modular multi-agent architecture increases scalability and flexibility. The agents carry out a particular task and may be extended separately, or updated, without influencing the general framework of the system. The modular structure helps hit the adjustment of the framework to various learning settings and educational institutions. Nevertheless, the success of the methodology will be determined by the quality and access of engagement data, and the consultative part will be based on well-thought prompts and well-structured inputs. The research ought to test the framework on different data sets of education in the future to further determine its generalization ability.

V. CONCLUSION AND FUTURE DIRECTIONS

The proposed multi-agent model for student performance prediction uses LLM based academic advising system. With a transparent logistic regression model, the system can identify academically at-risk drivers early through the use of engagement analytics, which is easy to interpret and accountable. The agent-based design, which is modular, generates analytical reasoning and advisory generation independently, both of which are necessary to provide reliable prediction and controlled use of large language models. The results of the experiments prove that the combination of academic and engagement measures contributes to the predictive stability, but without decreasing transparency.

Future direction will include the validation of the framework on a variety of educational large datasets with more interaction indicators, and the study of real-time academic support systems and human-in-the-loop supervision.

REFERENCES

[1] Siemens and R. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," in Proc. 2nd Int. Conf. Learning Analytics and Knowledge (LAK), Vancouver, BC, Canada, 2012, pp. 252–254.

[2] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris, "School engagement: Potential of the concept, state of the evidence," Review of Educational Research, vol. 74, no. 1, pp. 59–109, 2004.

[3] P. Cortez and A. Silva, "Using data mining to predict secondary school student performance," in Proc. Int. Conf. Educational Data Mining (EDM), Montreal, QC, Canada, 2008, pp. 5–12.

[4] Romero and S. Ventura, "Educational data mining: A survey," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 40, no. 6, pp. 601–618, Nov. 2010.

[5] S. Huang and N. Fang, "Predicting student academic performance using logistic regression," Educational Technology & Society, vol. 16, no. 2, pp. 160–173, 2013.

[6] Li, Y. Zhang, and Q. Yang, "Single-agent vs. multi-agent LLM strategies for automated student reflection assessment," in Proc. Pacific-Asia Conf. Knowledge Discovery and Data Mining (PAKDD), 2025.

[7] T. Wang, X. Liu, and Z. Chen, "LLM-powered multi-agent framework for goal-oriented learning in intelligent tutoring systems," in Proc. WWW Companion Conf., Austin, TX, USA, 2025.

[8] M. Tamascelli, O. Bunch, B. Fowler, M. Taeb, and A. Cohen, "Academic advising chatbot powered with AI agent," in Proc. ACM Southeast Conf. (ACMSE), Cape Girardeau, MO, USA, 2025.

[9] Z. Chu, Y. Li, and H. Zhao, "LLM agents for education: Advances and applications," arXiv preprint arXiv:2503.11733, 2025.

[10] J. P. López-Goyez, A. González-Briones, and Y. Demazeau, "An adaptive multi-agent architecture with reinforcement learning and generative AI for intelligent tutoring systems," Applied Sciences, vol. 16, no. 3, p. 1323, 2026

[11] S. Kasneci et al., "ChatGPT for education: Opportunities and challenges," arXiv preprint, 2023.

[12] M. Rasul et al., "The role of ChatGPT in higher education," arXiv preprint, 2023.

[13] IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2, IEEE, 2017.

[14] A. Abdelhamid, M. Hussein, and S. El-Ramly, "Adaptive multi-agent smart academic advising framework," IET Software, vol. 15, no. 3, pp. 199–210, 2021.