Advancing Models Portability of Students' Failure Prediction Using Ontology Modeling Approach

Mohamed Daoudi[1,*], El Miloud Smaili[2], Ilyas Alloug[1], Ilham Oumaira[1], and Moulay El Hassan Charaf[2]

[1]Engineering Sciences laboratory, Ibn Tofail University, Kénitra, Morocco
[2]LaRI Laboratory, Ibn Tofail University, Kénitra, Morocco
Email: mohamed.daoudi@uit.ac.ma (M.D.); miloud.smaili@uit.ac.ma (E.M.S.); ilyas.alloug@uit.ac.ma (I.A.); ilham.oumaira@uit.ac.ma (I.O.); my.charaf@uit.ac.ma (M.E.H.C.)
*Corresponding author

Manuscript received November 23, 2023; revised December 12, 2023; accepted April 22, 2024; published August 23, 2024

Abstract—**Predicting student failure has recently become a major topic of interest for academic researchers. Research in the field of e-learning focuses mainly on the performance and accuracy of machine learning models designed for specific courses. Within the existing literature, not much attention has been given to the application of a prediction model, initially developed for a specific course, and across various other courses. By "model portability", we mean the capacity of a machine learning model to be applied in many different courses and platforms while maintaining consistent or minimal loss in accuracy. Several factors can significantly affect the portability of a model. The purpose of this study is to evaluate the portability of models developed from data extracted from Moodle logs, augmented with an ontology layer. The adopted approach aims to determine how the quantity of data extracted from Moodle activity traces, the type of attributes chosen (numerical or discretized), and particularly the integration of an ontological structure, affect the portability of models to predict student performance. We applied the K-NN classification algorithm to a set of courses at a similar level to build a model. Then, we evaluated the transferability to other courses by assessing accuracy. The results show that the portability of machine learning models and their implementation with different courses is possible in some cases with an accepted decrease in accuracy. Moreover, the findings demonstrate that the use of an ontology allows a notable improvement in terms of portability.**

Keywords—**failure prediction, portability of models, machine Learning, K-NN, ontologies**

I. Introduction

In recent years, e-learning has become a key tool for Moroccan students. This online approach offers benefits such as flexibility in timing and location, facilitating a dynamic exchange of information between teachers and students through platforms known as Learning Management Systems (LMS). Moodle is a popular open-source Learning Management System (LMS) allowing users to build powerful, attractive, and flexible e-learning courses [1]. The interaction logs of students with resources in Moodle provide a fundamental dataset for building machine learning models to record student behavior and predict learners' performance [2].

Predicting student failure in the e-learning context holds immense significance as it allows for the early identification of students at risk. This, in turn, facilitates timely interventions, such as personalized support and additional resources to prevent academic setbacks. In environments with limited resources, predictive models facilitate efficient resource allocation, allowing institutions to focus support efforts on at-risk students and optimize interventions. Furthermore, predictive insights obtained from such models contribute to personalized learning experiences. Educators can tailor content, pace, and teaching methods based on individual student needs, improving the overall quality of the learning experience.

Using data from Moodle, we have successfully created predictive models specific to single courses. The remaining challenge is to develop models that can be generalized across multiple educational contexts, enabling them to perform efficiently on different courses and platforms. These models, often referred to as 'portable' or 'transferable' models, hold the key to a more adaptable and scalable approach to predicting student outcomes [3, 4]. The challenge is to design models based on a single course that are generalizable across multiple courses or future iterations of the same course. These are commonly referred to as transferable or portable models [5, 6].

In this study, we firstly emphasize early identification of students at-risk so that institutions can intervene proactively. The choice of the K-NN algorithm in this research is underpinned by its distinctive strengths and suitability for predicting student outcomes. K-NN, or K-nearest neighbors, stands out for its simplicity, flexibility, and capacity to handle diverse and high-dimensional educational data, aligning well with the multifaceted nature of student performance metrics in e-learning environments. Its ability to capture local interactions and patterns makes it particularly effective in scenarios where neighboring data points hold significant informational value.

Afterwards, the study explores predictive models' adaptability across different courses. This offers insights into how well a model developed for one course can be applied to others. For this purpose, we explore the portability of models generated by adopting an approach based on semantic web resources, especially ontologies, to analyze the logs of learner interaction with the content management system. Basically, the process incorporates an ontology of e-learning attributes based on Bloom's taxonomy to evaluate the portability of models. In this context, models were constructed using student interactions from Moodle logs, with the target attribute being binary, indicating if the participant has successfully passed or failed the course (Pass/Fail). Furthermore, courses were categorized according to the extent of activity, resource utilization, and interactions within a Moodle course.

doi: 10.18178/ijiet.2024.14.8.2145

The research question driving this study is: Is the generalization of a predictive failure model, initially developed for one online course (source), possible for application to other online courses (target) with a similar usage level and without prediction quality loss? Thus, the principal contributions of this research are as follows:

- Constructing a hierarchical class structure within the ontology model to describe Moodle student interactions.
- Developing portable predictive models for predicting student failure.
- Evaluating the transferability of models using ontology across courses with varying levels of usage.

To the best of our knowledge, we consider model portability as the retention of predictive effectiveness below specific thresholds of Area Under the Curve (AUC) reduction—15% for high activity courses and 20% for low activity courses—thus balancing accuracy and transferability across different educational contexts.

The paper is organized as follows: in section "Related Work," we review the literature on this study. The section "Materials and Methods" presents the tools and methodologies employed, including data extraction and ontology presentation. In the "Results" section, we present and discuss the findings of the research. Finally, we conclude with future orientation of this study.

II. Related Work

The Open Academic Analytics Initiative (OAAI) has acknowledged the issue of predictive model portability since 2011. OAAI's purpose is to advance the field of learning analytics by exploring the challenges of expanding learning analytics across higher education institutions [7]. The two primary sub-objectives of this initiative are the expansion of predictive models and the development of an open-source model for predicting student success [8]. Most research is focused on analyzing and predicting student behavior within e-learning platforms, with a collection of Educational Data Mining instruments to measure student performance [9] or detect student at-risk [10].

In addition, recent studies focus on maximizing the accuracy of the model for a specific class or course. In this context, the Educational Data Mining (EDM) community is inclined toward standardizing machine learning applications in the e-learning sector. The aim is to develop predictive models that are both generalizable and portable across varied classes and courses, without compromising prediction accuracy. This continues to be a significant challenge within the EDM discipline. In [11], the authors conduct an investigation into the portability of predictive models across courses from different universities. Utilizing the J48 classification algorithm in two distinct experiments involving twenty-four courses, the findings suggest that the feasibility of transferring models is still subject to specific conditions. According to Beltran [12], the portability of predictive models is highly dependent on certain high-level attributes, which guarantee better prediction accuracy through model interoperability. In contrast, this article focuses primarily on exploiting the portability of predictive models to mitigate student failure in various courses, by identifying features that ensure effective interoperability of educational data mining predictive models.

Furthermore, the authors [3] show how a generic predictive model can be developed to identify at-risk students in a wide variety of courses. The authors confirm that the portability of a model between several courses is useful because these generic models are less resource intensive, easier to maintain and less likely to be overfit under certain conditions. Another work investigates the portability of machine learning models for predicting 30-day readmission risk in healthcare [4]. It highlights the challenges of transferring models between institutions, pointing out that models developed at one site often perform less well when applied to another site. The study introduces a new transfer learning technique, demonstrating improved predictive accuracy between different sites, highlighting the potential for more effective and generalizable readmission risk prediction models in clinical decision support systems.

Moreover, the study [6] evaluates the portability of machine learning models for landslide detection using remote sensing images. It compares the performance of Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Random Forest (RF) models across different geographic areas. The results indicate that while SVM and ANN models show significant accuracy reduction in non-local areas, the RF model demonstrates stable portability and maintains consistent performance, suggesting its suitability for landslide detection in diverse geographical settings.

The authors [11] describe ontology as a formal representation of knowledge, that offers the possibility of conceptually and operationally describing aspects associated with pre-processing tools in the field of data mining. The authors suggest extracting student data from the Moodle database and remodel it in an ontological structure, but the usefulness of the ontology is not clearly demonstrated in this study. In the study of Beltran's [12], authors investigate the crucial question of the portability of predictive models for higher education by applying a transfer learning approach on Learning Management Systems logs (i.e., Moodle). The research aims to create a multi-faceted and adaptable framework that can effectively predict the performance of undergraduates in different university courses.

In another work [13], authors conducted an analysis of predictive models based on students' study habits and social relationships in blended courses, highlighting the significance of generalizability across different class offerings. The research incorporated data from two introductory Computer Science courses, highlighting the potential of social metrics in predicting student performance, even before the first test.

These findings underscored the models' generalizability, as they proved effective in identifying at-risk students across various course settings. However, the authors [14] found that the portability of predictive academic performance models poses significant challenges. Their research highlights that even when applied to closely related student groups, substantial differences in model performance emerge, showing the complexity of creating universally applicable models based solely on past evaluation results. This emphasizes the need for a mixed approach to enhance model portability, suggesting that internal group dynamics play a crucial role in this context. Another work [15] explores the use of an ontology-based recommender system to identify the degree of confusion related to a specific concept.

Comparing the results of this research to previous studies related to predictive modeling in e-learning environments reveals some similarities as well as differences. This provides insights into how various machine learning methods can be used in different ways to improve education outcomes in the future. The alignment between this study and previous research is evident in the recognition of the significance of predictive modeling for identifying at-risk students in e-learning environments. Like prior studies, the findings underscore the potential of machine learning algorithms, such as K-NN, in early identification, providing educators with a valuable tool for proactive intervention and support. However, the main differences may arise in the specific methodologies employed, the choice of algorithms, or the integration of additional layers, such as ontology, as demonstrated in this paper. The study's focus on assessing model portability and the impact of ontology on predictive performance adds a nuanced layer to the existing body of research, contributing insights into the generalizability of predictive models across different courses and platforms. This emphasis on model portability distinguishes this research from studies that might primarily focus on predictive accuracy within specific course contexts.

Moreover, trends in this research may include a growing awareness of the importance of considering ontological layers in predictive modeling. The study suggests that the integration of ontology improves the portability of machine learning models, offering a notable enhancement in terms of adaptability to diverse courses and platforms. This insight could potentially mark a trend towards more sophisticated modeling approaches that leverage structured knowledge representation for increased accuracy and applicability across varied educational settings.

III. Materials and Methods

A. Materials

1) Protégé

Protégé is a free and open-source software that offers a large and growing community of users a suite of tools for creating ontology-based domain models and knowledge-based applications [16].

In the field of ontology engineering and knowledge representation, the Protégé tool has attracted a great deal of interest due to its strong capabilities in facilitating ontology development, editing and management. Deployed in a variety of scientific disciplines, including healthcare and education, Protégé offers a comprehensive platform that enables researchers to build domain-specific ontologies with high degrees of sophistication and granularity [17]. In particular, the tool offers a range of customizable features, such as visual editors and reasoning capabilities, which facilitate the validation and application of ontologies [18].

2) K-NN

In the domain of e-learning, the K-NN algorithm has found a unique niche, especially in the optimization of personalized learning experiences and early identification of students at risk of underperforming or dropping out [19]. For example, a study conducted at a major university utilized K-NN to predict student grades based on their interactions with an online learning platform, achieving notable success in identifying at-risk students well before traditional midterm evaluations [20]. By leveraging the inherent strengths of K-NN in handling diverse and high-dimensional educational data, the study demonstrates the algorithm's potential as a powerful tool in developing more adaptive and responsive e-learning environments.

Comparing K-NN with other machine learning algorithms commonly used to predict learner performance provides important evidence of its effectiveness. For instance, decision trees are known for their interpretability, allowing educators to understand the decision-making process. However, they might struggle with capturing complex relationships in high-dimensional data. Support vector machines excel at handling high-dimensional data and managing noise but might be less interpretable. In contrast, K-NN offers a balance, providing adaptability to diverse learning patterns while maintaining interpretability to a certain extent.

In the specific context of predicting student outcomes in e-learning, where individual learning patterns and interactions are diverse, the K-NN's adaptability becomes an asset. Its effectiveness in early identification of at-risk students, as demonstrated in the study, highlights its practical utility. However, it is essential to acknowledge the computational demands of K-NN, which may pose challenges with scalability, particularly with larger datasets common in educational settings. Additionally, K-NN is sensitive to outliers and noise in the data, factors that need careful consideration in the preprocessing stages.

Our research highlights the use of the K-NN algorithm in educational data analysis, focusing on optimizing K-values for model portability and improving the algorithm's accuracy given the complexity of Moodle data. This involves considering missing values, noise, and high dimensionality to increase accuracy and ensure our approach improves model portability to predict student performance effectively.

Developing a model based on the K-NN (K-Nearest Neighbors) algorithm, we built on its unique capacity to make predictions based on the spatial relationships between observations, minimizing the problems associated with multicollinearity [21]. Contrary to models such as linear regression, where multicollinearity can have a significant impact on performance by increasing variance [22], the K-NN algorithm's dependence on distances rather than coefficient estimates allows for a more inclusive feature selection strategy. Although we identified strong correlations between features in some dataset, we chose to retain all features, to take advantage of K-NN's robustness to multicollinearity. This approach ensures that our model benefits from the full data available for more accurate predictions.

B. Data Collection and Data Preprocessing

This section presents the description and pre-processing of the data used in this research. We demonstrate the data collected, as well as the sources and methods of data extraction.

1) Data collection

The aim is to create a heterogeneous population to explore the impact of varying academic backgrounds on the portability of failure prediction models. Although this approach provides a diverse sample, it raises questions about generalizability, which future research will address. Data collection from Moodle ensures objectivity, minimizing bias. However, focusing on a single study area may affect the wider applicability of the results. Therefore, we extracted course data from various fields and institutions.

In the context of the study, the research data source is Moodle platform, specifically from three affiliated institutions of Ibn Tofail University: the National School of Business and Management, the National School of Applied Sciences, and the Faculty of Economics and Management.

Courses selected based on specific criteria and a Structured Query Language (SQL) query formulated to retrieve details of these courses and student activity in the platform. The period from October 2020 to February 2021, as specified in our SQL query, corresponds to a full academic semester affected by the COVID-19 pandemic, allowing us to capture a maximum number of online interactions. During this period, courses were delivered entirely online, including those available since 2013 and conducted by tutors with significant e-learning experience, considerably enriching our dataset. Although this selection may introduce some bias by focusing on a specific period, it has enabled us to build up a comprehensive dataset, which is essential for analysis in our predictive model.

The Moodle database archives all user activity in its log tables. More precisely, the "logstore_standard_log" table records the various actions taken by users in all activities such as: courses, quizzes, polls, forums, and others. In this context, to determine if a teacher has assigned activities in each course, the choice, quiz, modules, and course_modules tables are consulted. The user table stores learner data. Fig. 1 illustrates the physical data model and the relations between the various tables used in the context of this study.

Fig. 1. Physical relational data model.

Table 1. Courses information

Course code	No. of activities	Subscribed Students	Total actions	Usage
TDN	0	115	26,858	Heigh
ME	0	182	5,485	Heigh
FPVR	0	78	97	Low
RI	0	35	797	Low
EMF	4	472	33,568	Heigh
DC	0	102	1,264	Low
PR	0	453	19,474	Heigh
APP	15	227	127,597	Heigh
SP	68	32	111	Low
EL	0	131	18,218	Heigh
OP	0	68	414	Low
SD	0	370	24,172	Heigh
IB	3	46	44	Low
AN	1	65	160	Low
EM	0	32	5,350	Heigh

We analyzed log table data from 2,408 students enrolled in 15 different courses at Ibn Tofail University, each of which has been available on the platform since 2013. Table 1 provides a summary of these courses, including the course code, the quantity of activities planned by the instructor, the student count, the total number of actions executed within each course, and the degree of Moodle utilization (categorized as Low, Medium, or High). Due to ethical and confidentiality considerations, we obtained approval from the Moodle platform administrator and anonymized the personal details of both teachers and students.

In our research, we extracted learner characteristics by collecting data during interactions with the Moodle platform. The extracted features passed through a pre-treatment process for further analysis. Table 2 provides an overview of the pre-processed features, categorized with their respective descriptions, offering insight into various aspects of learner engagement and interaction within the Moodle platform for the selected courses.

Table 2. Features description

Feature	Description
Course_view	Number of visits to a course (numeric).
Page_view	Number of page accesses by a user on the Moodle platform (Numeric).
Resource_view	Number of consults a learning resource within a course (Numeric).
Url_view	Total of viewed external URLs linked within the Moodle platform (Numeric).
Forum_add_discussion	Number of discussions initiated by the user in a forum (Numeric).
Forum_add_post	Number of times a user posts a problem in an existing forum discussion (Numeric).
Forum_subscribe	Number of forums to which the user subscribed (Numeric).
Forum_user_report	User activity or performance in a forum is reported (Numeric).
Assignment_upload	Total number of uploaded assignments (Numeric).
Assignment_view	Total number of viewed assignments (Numeric).
Assignment_view_submission	Count of user views of their submitted assignment (Numeric).
Choice_choose_again	Count of user revisits and changes their choice in a Choice activity (Numeric).
Choice_view	Count of user views of a Choice activity (Numeric).
Quiz_attempt	Count of user attempts at a quiz (Numeric).
Quiz_close_attempt	Total number of attempts to complete or close a quiz (Numeric).
Quiz_continue_attempt	Total number of times a user resumes or continues a previously started quiz (Numeric).
Quiz_review	Count of user reviews of feedback or results from a quiz attempt (Numeric).
Total_interactions	Total number of interactions or activities performed by a user (Numeric).
Number_of_days_connected	Total number of days a user has been connected or engaged with the Moodle platform (Numeric).
final_grade	Final grade or assessment score achieved by a user in a course, boolean: (0 to 10 = False), (10 to 20 = True).

2) *Data preprocessing*

Data preprocessing is a fundamental step in the preparation of data. Typically, the native data includes outliers, null values or is in an inappropriate form that cannot be employed in machine learning models [1]. The cleaning of collected data is required. In this study, the collected data from the Moodle platform underwent a systematic pre-processing phase to check its quality, suitability, and adequacy for further use.

The first step in preprocessing is to ensure that learners' information has been appropriately anonymized, in line with ethical standards, confidentiality and Moroccan legislation. We used "Data Masking" as the main technique in our study. This technique consists of replacing the first and last names of students and teachers with numerical codes that do not allow unique identification of an individual. Similarly, we have chosen to use course codes instead of course titles to avoid any association between courses and teachers.

a) *Checking missing data*

After identifying the missing values in our dataset, we processed them using the mean imputation method. This method consists in replacing the missing values by the mean of the features concerned. By employing this strategy, we ensured a representative and complete dataset for our study, reducing the influence of missing values on integrity and reliability of our findings.

b) *Detects and eliminates outliers*

We employed the z-score to filter and remove outliers from our dataset. In statistical jargon, the z-score represents the number of standard deviations of a value compared to the mean.

Fig. 2 shows the presence of outliers for the 5 categories of our features (1: Evaluating_Examining, 2: Working, 3: Communicating, 4: Reading_Viewing, 5: Engagement) for the six datasets used in this study. After applying the z-score to filter outliers from our dataset, we proceeded to refine our data by removing identified outliers.

c) *Scaling numeric features*

In this paper, we have included the practice of scaling numerical features as a key step in data preprocessing. Scaling ensures that numerical features are transformed into a normalized range, preventing certain features from predominating over others in the model learning process. This is implemented using the StandardScaler from Scikit-learn library, by centering the data around its mean and scaling it to a standard deviation of 1.

We notice that the selection of StandardScaler as the preferred scaling technique is grounded in its appropriateness for features following a normal distribution. This aligns with the assumptions of certain machine learning algorithms.

Given the nature of the collected data from the Moodle platform, where features are expected to exhibit a Gaussian distribution, StandardScaler proves effective in centering the data around its mean and scaling it to a standard deviation of 1. This normalization ensures that each feature contributes fairly to the model learning process. This prevents any feature from disproportionately influencing outcomes due to differences in scale.

Furthermore, the use of StandardScaler aligns with the requirements of certain algorithms, particularly those that assume features are normally distributed. By transforming the data in this way, StandardScaler facilitates a more balanced and accurate learning process for these algorithms, enhancing their ability to discern patterns within the normalized features. While it is acknowledged that StandardScaler may be sensitive to outliers, the preprocessing phase likely includes measures to handle such outliers or alternative strategies to address this sensitivity.

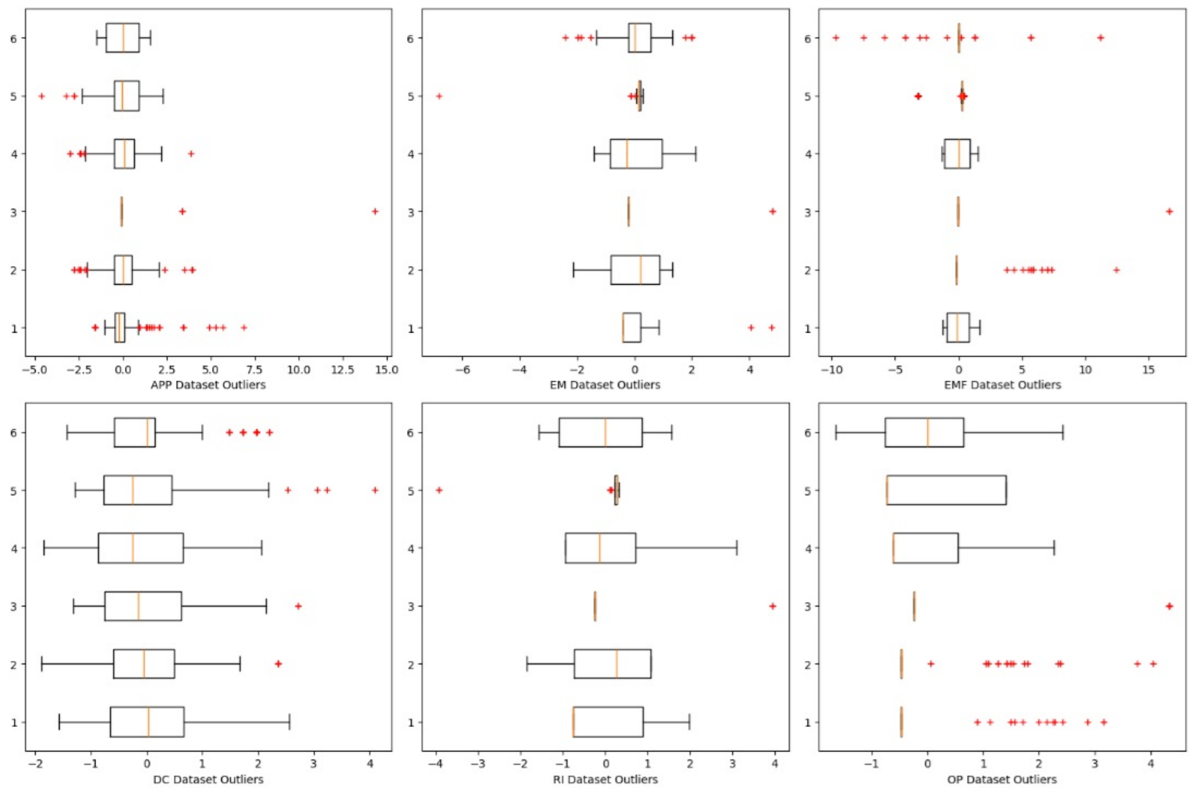

Fig. 2. Outlier detection by features category.

d) Correlation matrix

The correlation matrices measure the direction (positive/negative) and the intensity (low/medium/high) of variable interrelations [23]. As shown in Fig. 3, this information provides important context for understanding the behavior and performance dynamics of learners on the Moodle platform.

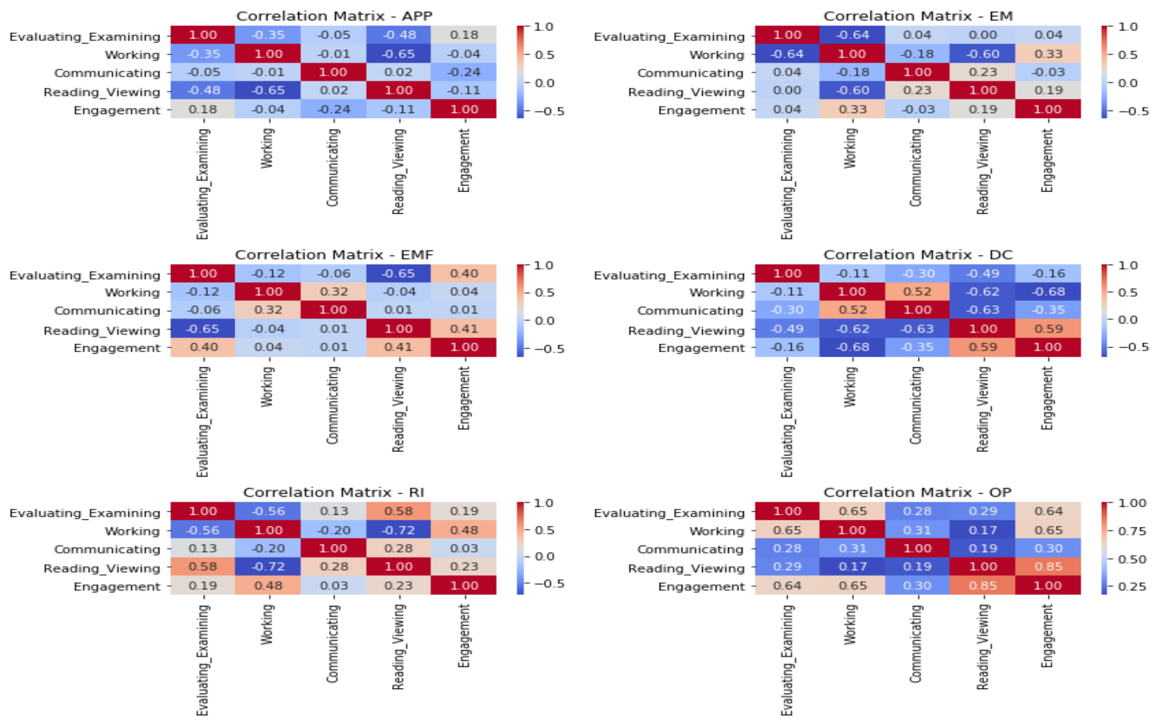

Fig. 3. Correlations between features across datasets.

The most notable results of our study are the negative correlations between certain aspects of engagement, such as "Evaluating_Examining" and "Working", suggesting potential trade-offs. Conversely, positive correlations highlight synergies, such as between "Working" and "Engagement".

The last step in our data pre-processing pipeline, is to transform the "Final grade" feature into a Boolean format. This transformation categorizes the scores, representing a binary result: "False" for scores between 0 and 10 and "True" for scores between 10 and 20. This binary representation of the "Final Score" feature enables a more targeted use of the variable in our machine learning models.

IV. Experimentation Methodology

A. Methodology

In our research, we propose a structured methodology (Fig. 4) for evaluating the portability of predictive models of failure in different educational environments. The initial phase focuses on data extraction and preparation. Then, we meticulously clean the relevant data after extracting it from the Moodle platform. We organize this data into prepared datasets and enrich each course with distinct ontological structures, available in both numerical and discretized forms.

In the next phases, we focus on model development and group classification. We build individual predictive failure models for each dataset using the K-NN algorithm. Next, we classify the courses based on their level of usage on the Moodle platform. Within these groups, we identify a model through a selection process that will serve for future testing.

In the final steps, the methodology includes an inter-course evaluation phase and the calculation of performance metrics. We test the selected model on datasets from other courses within the same group to validate its portability. Performance metrics result from calculating the Area Under the Curve (AUC) value for each test scenario. This measurement includes the determination of the AUC loss, which quantifies the effectiveness of the model's portability. The AUC has an important statistical property: the AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance [24].

This methodology aims to provide a powerful approach for assessing the adaptability of predictive failure models in a variety of educational environments.

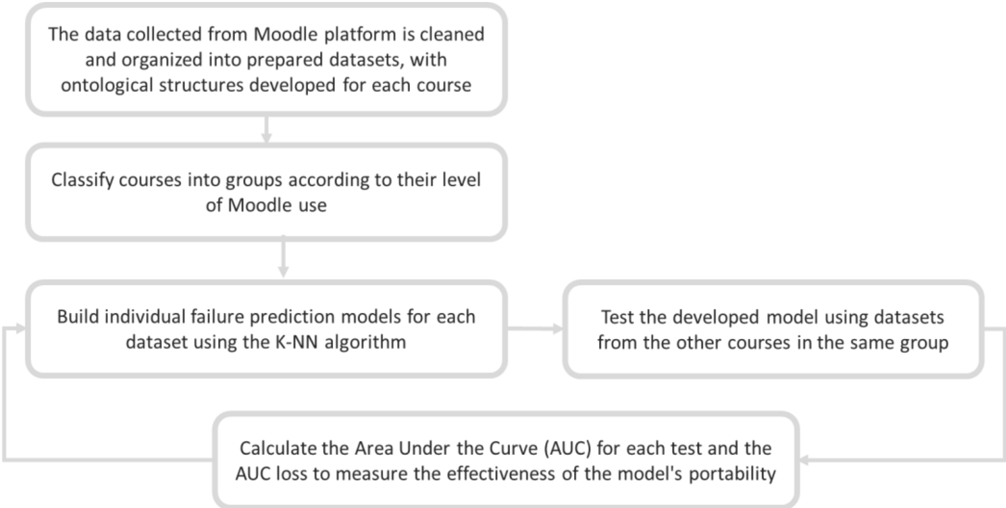

Fig. 4. Experimental methodology of our research.

B. Ontology Modeling Approach

As shown in Table 3, we grouped our 15 different courses into three different levels of Moodle activity usage. Moodle offers us a variety of resources (web pages, files, links to websites, and videos) and various types of activities (quizzes, assignments, polls, forums, wikis, etc.) [11]. We defined three usage levels depending on the activities used and the actions performed by participants as part of the course:

Table 3. Groups of courses

No	Groups	No. Of Courses
1	High level	8
2	Medium level	1
3	Low level	6

- Low level: The course did not contain any activities proposed by teachers, such as surveys, quizzes, forums, assignments, etc., and had fewer than 10,000 actions performed by the students.
- Medium level: the course has between 10,000 and 20,000 actions performed by students.
- High level: the course counts over 20,000 student actions.

In our ontology developed using Protégé, we start by building a structured taxonomy of classes to model the stages of a learning process, focusing on the different aspects of engagement and interaction. As shown in Table 4, the taxonomy is composed of five primary categories/classes from 19 attributes or actions provided by the Moodle logs (Data properties): "Viewing", "Communication", "Participating", "Assessment", and "Involvement".

Viewing: this class encompasses the initial stage where learners engage with learning materials. It represents passive consumption and observation of content, such as reading, watching videos, or reviewing resources.

Communication: the "Communication" class captures interactions initiated by learners to ask questions, inquiry clarification or discuss topics. This class focuses on the exchange of information, problem solving and ideas.

Participating: the "Participating" class signifies active engagement in collaborative activities. Students participate in-group discussions, submit homework assignments, and provide responses to surveys.

Assessment: the "Assessment" class reflects evaluation processes. This phase includes quizzes, exams, assignments, and other forms of assessment.

Involvement: the "Involvement" class represent the holistic engagement. It encompasses the count of connected days and the total number of interactions with learning content.

Table 4. Attributes of each category

Viewing	Communication	Participating	Assessment	Involvement
Course_view	Forum_add_discussion	Assignment_upload	Quiz_attempt	Total_user_interactions
Page_view	Forum_add_post	Assignment_view	Quiz_close_attempt	Number_ of_days_user_connected
Resource_view	Forum_subscribe	Assignment_view_submission	Quiz_continue_attempt	
Url_view	Forum_user_report	Choice_choose_again	Quiz_review	
		Choice_view		

This hierarchy includes seven other key classes (Fig. 5): "Course", "Platform", "Activity", "User", "Student", "Teacher" and "Resource". The "Course" class refers to teaching units, while the "Platform" refers to the digital environment. The "Activity" class is the parent class of the five classes above (Viewing, Communication, Participating, Assessment, and Involvement). The "User" class covers individuals, with two sub-classes: "Students" and "Teachers" represent learners and instructors. The "Resources" class includes learning materials.

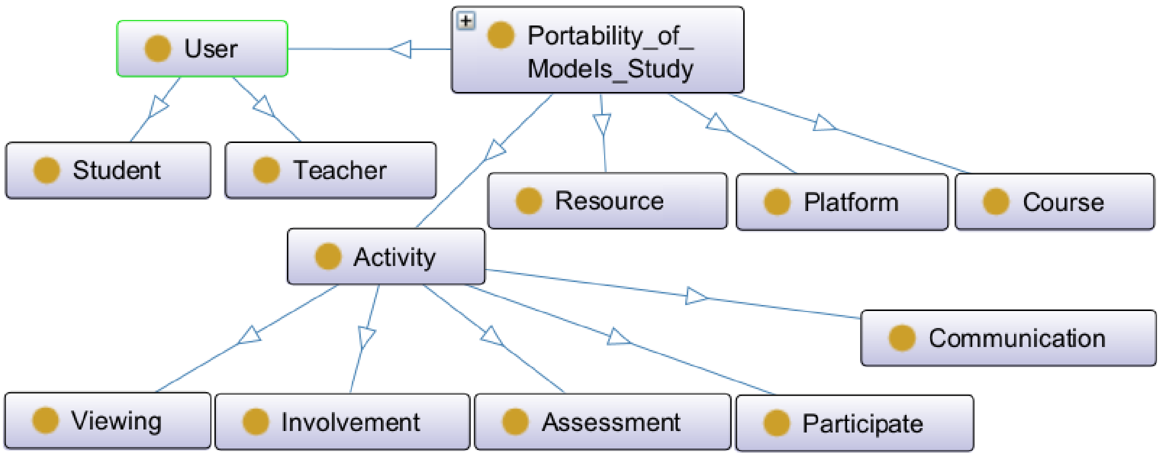

Fig. 5. Class hierarchy and OntoGraf visualization of our ontology.

Our taxonomy shows a structured progression of learning activities from passive to active engagement, encompassing communication, collaboration, evaluation, and overall involvement. This ontology enables analysis of learner behavior and engagement patterns.

To build our dataset, we give each category a value (between 0 and 100) representing the sum of the percentages of events each student has done in Moodle that belong to that category. To discretize the continuous numerical features into discrete bins, we used the KBinsDiscretizer pre-processing technique provided by the scikit-learn library in Python. Finally, we generated two distinct datasets or CSV files: one using the numerical data and the second with the attributes of the 5 categories previously described.

In both datasets, we have added an additional attribute named "Target". This attribute corresponds to the final grade obtained in the course concerned and serves as a predictive target for our machine learning model. The final grade, which lies between 0 and 20, is classified into two distinct categories: "FALSE" if the participant's final grade is less than 10 and "TRUE" if the participant's final grade is greater or equal to 10.

V. Results and Discussion

In this section, we compare the performance of models with and without ontologies. First, we present metrics for models with simple data (numerical and discretized), and then analyze the impact of ontologies on accuracy and performance. The courses comprise two distinct categories: high-level courses and low-level courses. The experiment aims to predict student failure in different courses using the K-NN algorithm, to assess the model's portability within the same category of courses. We evaluated the performance of the K-NN algorithm using numerical and discretized datasets.

A. *Model Performance Without Ontologies*

1) *Courses with a high level of use*

Using numerical or discretized datasets from courses with a high level of usage (Table 5), the K-NN algorithm shows superior performance on the APP course, achieving an AUC of 0.684. The EMF course follows with an AUC of 0.5, then the EM course with an AUC of 0.333. When using numerical or discretized data for high-level courses, the average AUC is 0.505.

Table 5. High-usage course: Portability metrics without ontologies

		APP	EMF	EM	AVG
numerical dataset	Course code /AUC	0.684	0.179	0.705	0.522
		0.5	0.5	0.5	0.5
		0.536	0.689	0.333	0.519
	Course code / AUC Loss	-	−0.505	+0.021	−0.242
		0	-	0	0
		+0.203	+0.356	-	+0.280
discretized dataset	Course code / AUC	0.684	0.454	0.586	0.575
		0.5	0.5	0.5	0.5
		0.408	0.501	0.333	0.414
	Course code / AUC Loss	-	−0.230	−0.098	−0.164
		0	-	0	0
		+0.075	+0.168	-	+0.122

Using numerical data, the model trained on APP course data exhibited a clear drop in performance only when applied to the EMF course. Conversely, for the other courses, we noted a relative performance improvement of approximately 4.4%. With discretized data, a loss of accuracy of 14% on average, which is still acceptable.

2) Courses with low level of use

As shown in Table 6, with numerical or discretized data, the K-NN algorithm is the most efficient for the RI course, achieving an AUC of 0.916. It is then followed by the DC course, with an AUC of 0.566, and the OP course, with an AUC of 0.423. The average AUC for low-level courses, when using either type of data, is 0.613.

In courses characterized by low activity levels, the models exhibit a significant decline in performance. An average accuracy loss of 34% is noted when utilizing numerical data, and a 36% reduction in accuracy is seen with discretized data. This raises the matter of the quantity of data.

Table 6. Low-usage course: portability metrics without ontologies

		RI	DC	OP	AVG
numerical dataset	Course code / AUC	0.916	0.470	0.505	0.630
		0.5	0.566	0.5	0.522
		0.478	0.541	0.423	0.480
	Course code / AUC Loss	-	−0.446	−0.411	−0.428
		−0.066	-	−0.066	−0.066
		+0,055	+0,118	-	+0,086
discretized dataset	Course code / AUC	0.916	0.470	0.472	0.619
		0.5	0.566	0.5	0.522
		0.472	0.541	0.423	0.478
	Course code / AUC Loss	-	−0.446	−0.444	−0.445
		−0.066	-	−0.066	−0.066
		+0.049	+0.118	-	+0.083

In summary, for models developed in the absence of ontologies, K-NN algorithm performs better on low-level courses with an average AUC of 0.613, In contrast, for high-level courses, the K-NN algorithm shows similar performance with both numerical and discretized datasets, with average AUC values of 0.505. The K-NN algorithm's performance on individual courses varies, with the best performance observed for the RI course in low-level courses (AUC = 0.916) and for the APP course in high-level courses (AUC = 0.684).

Furthermore, we discovered that models trained on high-activity courses retained their performance when used in courses with similar activity levels. In contrast, using models trained on low-activity courses in similar contexts resulted in a significant loss of accuracy.

Future research should examine the influence of different data pre-processing methodologies, such as discretization, on the algorithmic performance of K-NN. In addition, researchers should attempt to identify the variables behind the performance variations observed in different courses.

B. Model Performance Using Ontologies

1) Courses with a high level of use

In the case of high activity courses, several significant findings emerge (see Table 7). Firstly, the model developed from APP course data showed the best performance when operated with numerical attributes, particularly in the context of the largest dataset. This result underlines the significant impact of data volume on predictive effectiveness. Examination shows that the distinction between discrete and numeric data types have minimal influence on model portability. The performance drop results in a range conservatively between 0% and 15.2%, a range generally considered acceptable. In some cases, it is interesting to note that the application of a model developed for one course to another increases predictive performance, indicating the potential for knowledge transfer between courses.

The implementation of an ontological structure provides a more comprehensive perspective for our research. Analysis reveals that when applied to numerical data, AUC loss fluctuates between 0% and 15.2%, reflecting a delicate trade-off between ontology integration and preservation of predictive accuracy. In contrast, for discretized data, AUC loss ranged from 5.8% to 19%, suggesting a slightly greater impact on predictive performance. Remarkably, in the high-activity category, ontology integration has no substantial influence on performance results. This lack of pronounced effect can be attributed to the inherent stability achieved using processed data, further underlined by sporadic examples of improved performance in some courses, where the benefits of ontology integration are more evident.

Table 7. High-usage course: Portability metrics using ontologies

		APP	EMF	EM	AVG
numerical dataset	Course code /AUC	0.618	0.433	0.500	0.517
		0.500	0.500	0.500	0.500
		0.496	0.652	0.500	0.549
	Course code /AUC Loss	-	−0.185	−0.118	−0.152
		0	-	0	0
		−0.004	+0.152	-	−0.074
discretized dataset	Course code /AUC	0.500	0.384	0.500	0.461
		0.500	0.500	0.500	0.500
		0.541	0.839	0.500	0.627
	Course code /AUC Loss	-	−0.116	0	−0.058
		0	-	0	0
		+0.041	+0.339	-	+0.190

2) Courses with low level of use

The model developed using the RI dataset, incorporating discretized variables, achieved the best results. The model's portability was more pronounced when using continuous numerical data, with the discretization process introducing a detectable drop in portability performance.

As illustrated in Table 8, based on the developed ontology, the AUC loss range extends from 1.9% to 9.8% for digital data, and from 2.42% to 28% for discretized data. In the absence of ontology, performance loss shows a wider spectrum, ranging from 6.6% to 42.8%. Similarly, at equivalent usage levels, performance degradation ranges from 6.6% to 44.5%. The incorporation of an ontology resulted in a perceptible but manageable compromise in terms of performance loss, albeit accompanied by a reduction in model development performance.

Table 8. Low-usage course: portability metrics using ontologies

		RI	DC	OP	AVG
numerical dataset	Course code /AUC	0.500	0.460	0.500	0.487
		0.333	0.566	0.506	0.468
		0.333	0.500	0.426	0.420
	Course code /AUC Loss	-	−0.040	0	−0.013
		−0.233	-	−0.060	−0.098
		−0.093	+0.074	-	−0.019
discretized dataset	Course code /AUC	0.750	0.440	0.500	0.563
		0.333	0.500	0.506	0.446
		0.333	0.392	0.426	0.384
	Course code /AUC Loss	-	−0.310	−0.250	−0.280
		−0.167	-	+0.006	−0.057
		−0.093	−0.034	-	−0.042

In summary, the study evaluates the impact of various data types and ontology integration on predictive model performance across different educational courses. The incorporation of an ontology generally led to manageable performance loss while improving model portability, especially in courses with high activity levels. Differences in data types, such as discrete versus numerical, showed minimal impact on model portability, and in some cases, applying a model developed for one course to another even enhanced predictive performance.

Based on our results, and considering a foundation for future studies, we suggested portability thresholds of 15% for courses with high levels of activity and 20% for those with low levels. This decision was based on a thorough understanding of the balance between model accuracy and portability in various educational contexts. These thresholds, formulated from an examination of model effectiveness with and without ontology integration, provide a measurable frame for assessing the adaptability of machine learning models. The establishment of these thresholds marks a step towards improving the flexibility and effectiveness of predictive models, opening the way for future research aimed at refining and validating these thresholds in other scenarios.

VI. Conclusion

In this study, we introduced our approach to assess the portability of predictive failure models in various educational environments, using data from Ibn Tofail University's Moodle journal as a case study. As part of our work, we enhance predictive modeling in education by integrating ontological structures, thereby improving model portability across a range of educational contexts.

Using the K-NN algorithm, our results indicate that models trained on high activity course data retain their predictive effectiveness when deployed in similar contexts, reinforcing the robustness of these models, particularly when enriched with ontological features. Furthermore, models developed from low activity courses show a significant drop in accuracy when applied to different courses at the same level. This performance degradation demonstrates the influence of data volume and pre-processing techniques such as discretization.

In addition, we propose portability thresholds based on a comprehensive analysis of model accuracy and portability. These thresholds provide a measurable framework to evaluate the portability of machine learning models, serving as a step towards refining and enhancing them through the incorporation of additional features and machine learning techniques.

All these observations explain the need for more strategic approaches to data selection and pre-processing and highlight the importance of creating portable, ontology-enriched predictive models to help prevent student failure in a diverse set of courses. Accordingly, this work makes a significant contribution to the advancement of generalized predictive models in e-learning environments, facilitating then the creation of more robust and effective educational support systems for students.

In our forthcoming efforts, we plan to examine the portability of failure prediction models across various machine learning algorithms and diverse educational settings. Another area of interest is to investigate the influence of learner knowledge prerequisites on model accuracy. Thus, applying our models to heterogeneous datasets would assess their robustness and generalizability in predicting student failure across different student populations and academic disciplines.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

The research was conceptualized by Mohamed Daoudi who, in collaboration with Ilyas Alloug and El Miloud Smaili, developed the methodology, carried out the formal analysis and wrote the initial draft. Mohamed Daoudi provided the essential resources, while Ilham Oumaira supervised the research, ensuring its clarity and validity. Moulay El Hassan Charaf contributed to the revision and improvement of the article's linguistics and content. El Miloud Smaili oversaw editing and formatting the article. The team, including Ilham Oumaira, participated in the thorough revision and editing process, making significant contributions to the final manuscript.

Acknowledgment

The authors express their gratefulness to Mr. GAHI Youssef and Mr. Charaf Moulay Hassan for their precious guidance and constructive comments, which contributed

significantly to improving the quality and scientific rigor of this paper.

REFERENCES

[1] M. Daoudi, N. Lebkiri, Y. Ouali, and I. Oumaira, "Student involvement in mobile-learning: Case of Ibn Tofail university," *Stat. Optim. Inf. Comput.*, vol. 10, no. 1, pp. 59–74, 2022. doi: 10.19139/soic-2310-5070-1217

[2] M. M. Tamada, R. Giusti, and J. F. de Magalhães Netto, "Predicting student performance based on logs in moodle LMS," in *Proc. 2021 IEEE Frontiers In Education conference (FIE)*, 2021, pp. 1–8. doi: 10.1109/FIE49875.2021.9637274

[3] G. Ramaswami, T. Susnjak, and A. Mathrani, "On developing generic models for predicting student outcomes in educational data mining," *Big Data Cogn. Comput.*, vol. 6, no. 1, 2022. doi: 10.3390/bdcc6010006

[4] S. Nyemba *et al.*, "Distillation to enhance the portability of risk models across institutions with large patient claims database," arXiv, preprint arXiv: 2207.02445, 2022. doi: 10.48550/arXiv.2207.02445

[5] S. Boyer and K. Veeramachaneni, "Transfer learning for predictive models in massive open online courses," *Artificial Intelligence in Education*, Springer, Cham, 2015, pp. 54–63. doi: 10.1007/978-3-319-19773-9_6

[6] Q. Hu, Y. Zhou, S. Wang, F. Wang, and H. Wang, "Improving the accuracy of landslide detection in 'off-site' area by machine learning model portability comparison: A case study of Jiuzhaigou earthquake, China," *Remote Sens.*, vol. 11, 2530, Oct. 2019. doi: 10.3390/rs11212530

[7] S. M. Jayaprakash, E. W. Moody, E. J. M. Lauría, J. R. Regan, and J. D. Baron, "Early alert of academically at-risk students: An open source analytics initiative," *J. Learn. Anal.*, vol. 1, no. 1, 2014. doi: 10.18608/jla.2014.11.3

[8] E. J. M. Lauría, J. D. Baron, M. Devireddy, V. Sundararaju, and S. M. Jayaprakash, "Mining academic data to improve college student retention: an open source perspective," in *Proc. the 2nd International Conference on Learning Analytics and Knowledge*, Vancouver British Columbia Canada: ACM, Apr. 2012, pp. 139–142. doi: 10.1145/2330601.2330637

[9] M. Yagci, "Educational data mining: Prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, vol. 9, Mar. 2022. doi: 10.1186/s40561-022-00192-z

[10] H. Waheed, S.-U. Hassan, R. Nawaz, N. Aljohani, G. Chen, and D. Gasevic, "Early prediction of learners at risk in self-paced education: A neural network approach," *Expert Syst. Appl.*, vol. 213, 118868, Sep. 2022. doi: 10.1016/j.eswa.2022.118868

[11] J. López-Zambrano, J. A. Lara, and C. Romero, "Improving the portability of predicting students' performance models by using ontologies," *J. Comput. High. Educ.*, vol. 34, no. 1, 2022. doi: 10.1007/s12528-021-09273-3

[12] C. A. R. Beltran, J. C. X. Júnior, C. A. S. Barreto, A. C. Gorgônio, and S. J. M. S. da Costa, "The application of models portability to predict undergraduate students' performance by using transfer learning," *Res. Soc. Dev.*, vol. 11, no. 5, 2022. doi: 10.33448/rsd-v11i5.27938

[13] N. Gitinabard, Y. Xu, S. Heckman, T. Barnes, and C. F. Lynch, "How widely can prediction models be generalized? performance prediction in blended courses," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 184–197, Apr. 2019. doi: 10.1109/TLT.2019.2911832

[14] J. L. Arroyo-Barrigüete, S. Carabias-López, T. Curto-González, and A. Hernández, "Portability of predictive academic performance models: An empirical sensitivity analysis," *Mathematics*, vol. 9, no. 8, 2021. doi: 10.3390/math9080870

[15] K. Tazi, S. Azzouzi, and M. E. H. Charaf, *An Ontology-Based Recommender System for Identifying Learners' Confusion in MOOCs*, 2023, p. 496. doi: 10.1109/CiSt56084.2023.10409939

[16] Protégé (August 2023). [Online]. Available: https://protege.stanford.edu/software.php#desktop-protege

[17] M. A. Musen, "The protégé project: A look back and a look forward," *AI Matters*, vol. 1, no. 4, pp. 4–12, Jun. 2015. doi: 10.1145/2757001.2757003

[18] T. Tudorache, C. Nyulas, N. F. Noy, and M. A. Musen, "WebProtégé: A collaborative ontology editor and knowledge acquisition tool for the web," *Semantic Web*, vol. 4, no. 1, pp. 89–99, Jan. 2013. doi: 10.3233/SW-2012-0057

[19] T. Tanner and H. Toivonen, "Predicting and preventing student failure —Using the k-nearest neighbour method to predict student performance in an online course environment," *IJLT*, vol. 5, pp. 356–377, Jan. 2010. doi: 10.1504/IJLT.2010.038772

[20] E. Jones, M. Priestley, L. Brewster, S. Wilbraham, G. Hughes, and L. Spanner, "Student wellbeing and assessment in higher education: The balancing act," *Assess. Eval. High. Educ.*, vol. 46, pp. 1–13, Jun. 2020. doi: 10.1080/02602938.2020.1782344

[21] N. A. Ramli, M. T. Ismail, and C.-W. Hooy, "An analysis on two different data sets by using ensemble of k-nearest neighbor classifiers," *WSEAS Trans. Math.*, vol. 13, pp. 780–789, Jan. 2014.

[22] A. K. Bayo, A. B. Rafiu, A. T. Funmilayo, and O. I. Oluyemi, "Investigating the impact of multicollinearity on linear regression estimates," *Malays. J. Comput. (MJoC)*, vol. 6, no. 1, pp. 698–714, Apr. 2021.

[23] C. D. R. Urbano. (November 2023). Knowledge extraction from courses and online learning activities. Master thesis. [Online]. Available: https://run.unl.pt/handle/10362/150426

[24] T. Fawcett, "Introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, pp. 861–874, Jun. 2006. doi: 10.1016/j.patrec.2005.10.010