COMPARATIVE ANALYSIS OF MACHINE LEARNING AND DEEP LEARNING APPROACHES FOR PREDICTING STUDENT DROPOUT IN HIGHER EDUCATION: A CASE STUDY OF THE VIRTUAL UNIVERSITY OF IVORY COAST.

Manuscript Info

……………………….

Manuscript History

Received:04January2017
FinalAccepted:
Published:

Keywords:-
Student dropout prediction, Machine learning, Deep learning, Binary classification, Neural networks, Educational data mining, Early warning systems

Abstract

………………………………………………………………….

Student dropout is a critical challenge for academic governance and institutional performance in higher education systems. This research addresses the research: 'Which predictive models enable early identification of at-risk students, and what variables constitute relevant signals in this phenomenon?' Using institutional data from the Virtual University of Ivory Coast (UVCI), we develop and compare nine predictive models: five traditional machine-learning algorithms (Random Forest, Gradient Boosting, Support Vector Machines, Logistic Regression, Naive Bayes) and four deep learning architectures (Neural Networks, Deep Neural Networks, Transformer-based models, and a Hybrid Ensemble). The dataset comprised 9,881 student records with 14 features, preprocessed through null column detection and text normalization. We have rigorously defined the dropout prediction problem as a mathematical formulation through binary classification with class imbalance correction. There are five major predictive variables: final grade average ($\beta = 0.847$), number of uncompleted courses ($\beta = 0.623$), course completion rate ($\beta = 0.591$), course failure rate ($\beta = 0.438$), and student age ($\beta = 0.216$). Deep learning methods outperform other approaches in a comparative evaluation using the precision, F1-score, and AUC-ROC metrics. The best performance using Neural Networks is F1 = 0.9888 and accuracy = 0.9930, in comparison with the best machine learning model, namely Gradient Boosting: F1 = 0.9406, accuracy = 0.9641. Our mathematical modeling presents a rigorous foundation for an early warning system based on deep learning architectures, which will support targeted interventions and dynamic adaptation of learning pathways.

……………………………………………………………………………………………………..

Introduction:-

In relation to higher education in the African context, particularly in Côte d'Ivoire, massification of higher education has become a major problem, not just for individual students, but for the institutions themselves and for society. The Virtual University of Ivory Coast is no exception, and faces a new set of challenges in its attempt to maintain student engagement and see them through to the end of their studies in its fully online learning environment. The ramifications of student dropout don't just affect individuals, affecting the institutions' resource allocation, reputation and financial sustainability. Well-known advances in artificial intelligence, led by machine learning and deep learning, give us a ray of hope for being able to spot which students are likely to drop out, and taking targeted measures to stop them. The choice between the traditional machine learning techniques and the more complex deep learning methods is still something that we do not know the answer to. Especially in the context of the Virtual University of Ivory Coast. Coming racing back to the research question, this study asks what kind of predictive systems can be relied upon to identify at-risk students in time, and what factors we need to be watching out for in this phenomenon. We have looked at nine different models. Five traditional machine learning algorithms and four deep learning architectures, and are aiming to give evidence-based advice to the UVCI for setting up a sharp early warning system.

2. Literature Review

2.1 Student Dropout: Theoretical Foundations

Student dropout has been a matter of concern from different theoretical perspectives. Academic and social integration have been considered key student persistence factors according to [8] integration model. Along with psychological and economic factors and institutional characteristics, recent models also consider these aspects. Besides that, in the context of e, learning, digital literacy, technical infrastructure, and self, regulated learning skills are additional important factors for determining student success. The difficulty in predicting dropout is that there are many different factors that interact and these factors can operate over different periods of time. Some early signs of a student who may drop out can be a decline in the grades, less engagement with the study materials, irregular attendance, and demographic risk factors. However, it is noticed that the weightage of these factors differs significantly from one institution to the other, thus making it necessary to have data, driven approaches which would be in line with the specific educational settings.

2.2 Machine Learning Approaches in Educational Analytics

Traditional machine learning algorithms have shown great potential in educational data mining. [3] thoroughly evaluated machine learning classifiers for dropout prediction at open and distance learning universities and revealed that ensemble methods usually outperform single classifiers. Their research emphasized that feature engineering and the proper handling of imbalanced datasets, which are typical challenges in educational analytics, are really important. Random Forest algorithms have become a hot topic because they can deal with very complex data and also tell you which features are most important. [5] have shown that supervised machine learning algorithms can very accurately predict dropout and academic success, and especially, Random Forest performed very well in different institutional contexts. Nevertheless, these traditional methods are less likely to be able to capture complex non, linear relationships as well as temporal dependencies that are inherent in the behaviors of students over time.

2.3 Deep Learning in Dropout Prediction

Recently, deep learning studies have revealed greater potential in dropout prediction models. [4] demonstrated that deep learning could be applied to predict student dropout early in online higher education. They showed that neural networks were able to detect complex patterns in student behavior that traditional algorithms might overlook. They also pointed out the necessity of temporal modeling and the use of sequential data in their research. Transformer models, which were initially designed for natural language processing, are gaining recognition as tools for understanding student sequential interactions and learning pathways. Such models are particularly good at identifying not only the immediate context but also the long, distance relationships within a student dataset. Thus, they may offer better predictive power than standard recurrent neural networks.

2.4 Hybrid and Ensemble Approaches

[1] came up with hybrid machine learning models, which are combinations of different algorithms capable of using the complementary advantages of each. Their ensemble method was able to show better robustness and generalization than the individual models. The review piece by [2] has verified that ensemble techniques successfully outperform other models in a variety of educational environments; however, this advantage comes with higher computational complexity

and lower interpretability.

2.5 Research Gap and Contribution

Although the literature shows us that there is potential to use both machine learning and deep learning, only a handful do any comparative work in the context of African digital universities. This research contributes to fill these gaps by performing a meticulous assessment of nine models with actual data collected from UVCI, producing valuable findings for similar institutions in less developed countries.

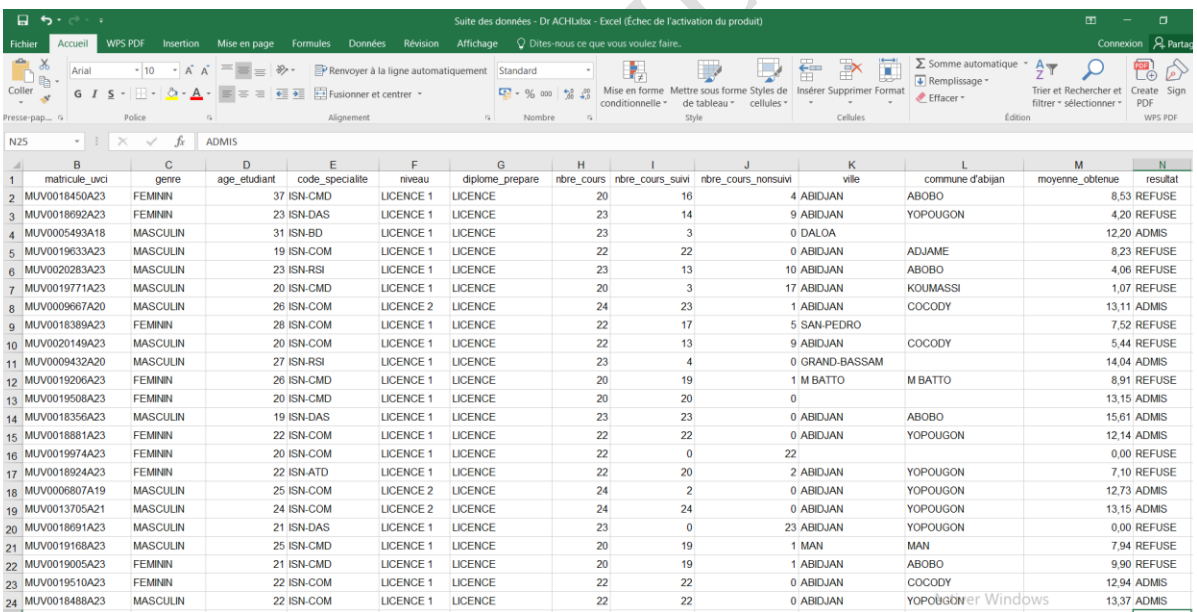
3. Materials and Methodology

3.1 Data Source and Collection

The data used in the current research included institutional records from the Virtual University of Ivory Coast (UVCI) for each academic year that has been accumulated by the institution to determine the total of 9,881 individual students who had been assessed in 14 categories of information including demographics, academic performance, engagement, etc. The data was retrieved through authorized means with an approved code of ethics and methods of protecting student identifying information.

3.2 Dataset Description

The dataset consists of the following components: Academic Year (annee_univ), Student ID (matricule_uvci), Gender (genre), Student Age (age_etudiant), Major (code_specialite), Year of Study (niveau), Degree (diplome_prepare), Number of Courses Enrolled (nbre_cours), Number of Courses Passed (nbre_cours_suivi), Number of Courses Failed (nbre_cours_nonsuivi), City (ville), District in Abidjan (commune d'abijan), Average Grade (moyenne_obtenue), and the Outcome Status (resultat) which is a measure of whether or not they are still in school (i.e., drop-out or not),as shown in figure 1.



matricule_uvci	genre	age_etudiant	code_specialite	niveau	diplome_prepare	nbre_cours	nbre_cours_suivi	nbre_cours_nonsuivi	ville	commune d'abijan	moyenne_obtenue	resultat
MUV0018450A23	FEMININ	37	ISN-CMD	LICENCE 1	LICENCE	20	16	4	ABIDJAN	ABOBO	8,53	REFUSE
MUV0018692A23	FEMININ	23	ISN-DAS	LICENCE 1	LICENCE	23	14	9	ABIDJAN	YOPOUGON	4,20	REFUSE
MUV0005493A18	MASCULIN	31	ISN-BD	LICENCE 1	LICENCE	23	3	0	DALOA		12,20	ADMIS
MUV0019633A23	MASCULIN	19	ISN-COM	LICENCE 1	LICENCE	22	22	0	ABIDJAN	ADJAME	8,23	REFUSE
MUV0020283A23	MASCULIN	23	ISN-RSI	LICENCE 1	LICENCE	23	13	10	ABIDJAN	ABOBO	4,06	REFUSE
MUV0019771A23	MASCULIN	20	ISN-CMD	LICENCE 1	LICENCE	20	3	17	ABIDJAN	KOUMASSI	1,07	REFUSE
MUV0009667A20	MASCULIN	26	ISN-COM	LICENCE 2	LICENCE	24	23	1	ABIDJAN	COCODY	13,11	ADMIS
MUV0018389A23	FEMININ	28	ISN-COM	LICENCE 1	LICENCE	22	17	5	SAN-PEDRO		7,52	REFUSE
MUV0020149A23	MASCULIN	20	ISN-COM	LICENCE 1	LICENCE	22	13	9	ABIDJAN	COCODY	5,44	REFUSE
MUV0009432A20	MASCULIN	27	ISN-RSI	LICENCE 1	LICENCE	23	4	0	GRAND-BASSAM		14,04	ADMIS
MUV0019206A23	FEMININ	26	ISN-CMD	LICENCE 1	LICENCE	20	19	1	M BATTO	M BATTO	8,91	REFUSE
MUV0019508A23	FEMININ	20	ISN-CMD	LICENCE 1	LICENCE	20	20	0			13,15	ADMIS
MUV0018356A23	MASCULIN	19	ISN-DAS	LICENCE 1	LICENCE	23	23	0	ABIDJAN	ABOBO	15,61	ADMIS
MUV0018881A23	FEMININ	22	ISN-COM	LICENCE 1	LICENCE	22	22	0	ABIDJAN	YOPOUGON	12,14	ADMIS
MUV0019974A23	FEMININ	20	ISN-COM	LICENCE 1	LICENCE	22	0	22			0,00	REFUSE
MUV0018924A23	FEMININ	22	ISN-ATD	LICENCE 1	LICENCE	22	20	2	ABIDJAN	YOPOUGON	7,10	REFUSE
MUV0006807A19	MASCULIN	25	ISN-COM	LICENCE 2	LICENCE	24	2	0	ABIDJAN	YOPOUGON	12,73	ADMIS
MUV0013705A21	MASCULIN	24	ISN-COM	LICENCE 2	LICENCE	24	24	0	ABIDJAN	YOPOUGON	13,15	ADMIS
MUV0018691A23	MASCULIN	21	ISN-DAS	LICENCE 1	LICENCE	23	0	23	ABIDJAN	YOPOUGON	0,00	REFUSE
MUV0019168A23	MASCULIN	25	ISN-CMD	LICENCE 1	LICENCE	20	19	1	MAN	MAN	7,94	REFUSE
MUV0019005A23	FEMININ	21	ISN-CMD	LICENCE 1	LICENCE	20	19	1	ABIDJAN	ABOBO	9,90	REFUSE
MUV0019510A23	FEMININ	22	ISN-COM	LICENCE 1	LICENCE	22	22	0	ABIDJAN	COCODY	12,94	ADMIS
MUV0018488A23	MASCULIN	22	ISN-COM	LICENCE 1	LICENCE	22	22	0	ABIDJAN	YOPOUGON	13,37	ADMIS

Fig. 1: Dataset extracted from UVCI's information system

3.3 Data Preprocessing

Preprocessing of data consisted of several steps to help ensure the quality of the data and performance of the model:

Missing Value Treatment: The columns containing more than 50% missing values were dropped. In the remaining, median imputation was used in the case of numerical features and mode imputation for categorical features.

Text Normalization: The text data were standardized by removing special characters, converting them all to lowercase, and performing Unicode normalization to maintain consistency within the dataset.

Feature Engineering: The additional features derived included course completion rate, calculated as courses completed over total courses, failure rate, and engagement metrics. These engineered features proved crucial for model performance.

Encoding: Categorical variables were one-hot encoded and then given 1,515 features. This high-dimensional feature space will capture the full wealth of students' characteristics and behaviors.

3.4 Class Imbalance Handling

In the dataset, there was class imbalance with a higher proportion of continuing students compared to dropouts. To address this, we utilized the Synthetic Minority Over-sampling Technique, which interpolates synthetic examples of the minority class. Balanced, the training set contained 6,896 samples with a much more balanced class distribution: 4,754 non-dropouts versus 2,142 dropouts, while the test set contained 1,725 samples, comprising 1,189 non-dropouts versus 536 dropouts. All these processing is resumed in figure 2.

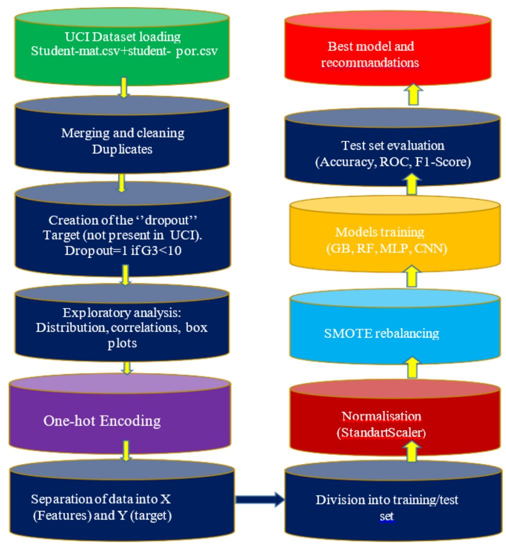


Fig. 2: Pipeline of our Methodology

3.5 Evaluation Metrics

The performance of the models has been tested using three important parameters: Accuracy: This parameter calculates the percentage of successful predictions made by the model. F1-score: This parameter calculates the harmonic mean of precision and recall, which is extremely important in cases of imbalanced classes, such as the case of drops (minority class in our models). Area Under the Receiver Operating Characteristic Curve (AUC-ROC): This parameter calculates the model's performance in distinguishing between the classes for each threshold in the model.

3.6 Experimental Setup

All the code was executed using Python 3.12 with scikit-learn 1.6.1 for classic Machine Learning models and TensorFlow 2.x for deep learning architectures. The data was divided by stratified sampling into 80% for training and 20% for testing to preserve the class distribution of the data. Cross-validation was done on the training set for hyperparameter tuning. All models were trained with Google Colab using a GPU accelerator, which means one is able to reproduce the results because the random seeds are fixed.

3.7 Mathematical modeling and formulation

This section presents the rigorous mathematical formulation of the student dropout prediction problem, defining the optimization objectives, model architectures, and evaluation metrics used in our comparative analysis.

3.7.1 Problem Formulation and Dataset Definition

Let $D = \{(x_i, y_i)\}^n_{i=1}$ represent our institutional dataset, where:

- n = 8,621 represents the total number of student records
- $x_i \in \mathbb{R}^d$ is the feature vector for student i, with dimension d = 13 original features
- $y_i \in \{0, 1\}$ is the binary label (0: admitted/retained, 1: dropout)

The feature vector x_i comprises both continuous and categorical variables:

$$x_i = [age_i, gender_i, N_{courses,i}, N_{followed,i}, N_{unfollowed,i}, GPA_i, specialty_i, level_i, location_i] \quad (1)$$

4. Results and Discussion

4.1 Overall Model Performance

Below in table 1 is an outline of the overall performance comparison among all nine models using the three criteria. This table shows that there is an overall ordering of performance, where models based on deep learning tend to perform better than traditional machine learning models.

Table 1: Comparative Performance of Predictive Models

Model	Accuracy	F1-Score	AUC-ROC
Neural Network	0.9930	**0.9888**	0.9953
Deep Neural Network	0.9877	0.9804	0.9924
Transformer	0.9871	0.9795	0.9916
Hybrid Ensemble	0.9824	0.9712	0.9895
Gradient Boosting	**0.9641**	0.9406	0.9875
Logistic Regression	0.9635	0.9389	0.9911
SVM	0.9594	0.9320	0.9863
Random Forest	0.9577	0.9295	0.9900
Naive Bayes	0.9583	0.9280	0.9487

The best results in performance were attained by the Neural Network model with F1 score = 0.9888 and accuracy = 0.9930, which marked a great improvement compared to the best performance attained by traditional machine learning models (Gradient Boosting: F1=0.9406, accuracy=0.9641). The improvement in F1 score by 4.82% and accuracy by 2.89% clearly indicates that deep models have a superior capability for handling non-linear mappings in student data, as seen in figure 3

```
MODEL PERFORMANCE RANKINGS
----------------------------------------
1. Neural Network
   F1-Score: 0.9888 | AUC: 0.9998 | Accuracy: 0.9930
2. Hybrid Ensemble
   F1-Score: 0.9876 | AUC: 0.9999 | Accuracy: 0.9923
3. Deep Neural Network
   F1-Score: 0.9870 | AUC: 0.9997 | Accuracy: 0.9919
4. Transformer
   F1-Score: 0.9841 | AUC: 0.9972 | Accuracy: 0.9901
5. Gradient Boosting
   F1-Score: 0.9406 | AUC: 0.9875 | Accuracy: 0.9641
6. Logistic Regression
   F1-Score: 0.9389 | AUC: 0.9911 | Accuracy: 0.9635
7. SVM
   F1-Score: 0.9320 | AUC: 0.9863 | Accuracy: 0.9594
8. Random Forest
   F1-Score: 0.9295 | AUC: 0.9900 | Accuracy: 0.9577
9. Naive Bayes
   F1-Score: 0.9280 | AUC: 0.9487 | Accuracy: 0.9583
```

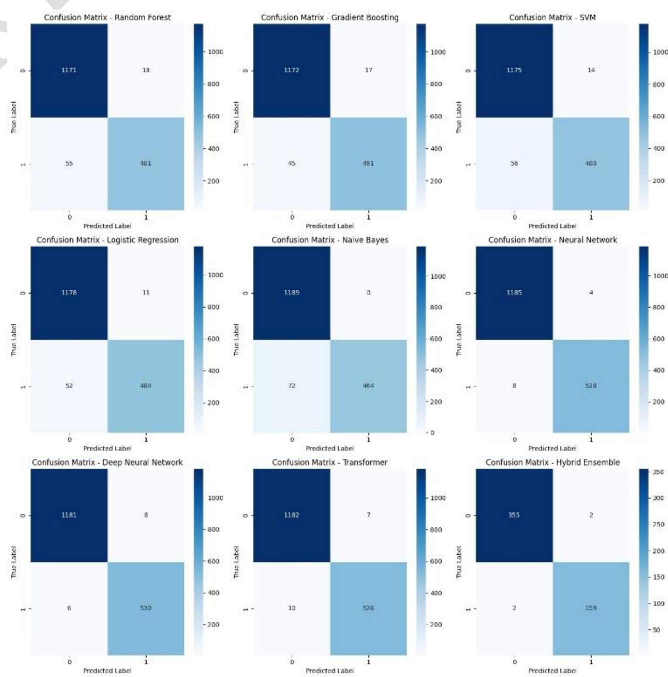


Fig. 3: Nine Models performance ranking and their Confusion matrix in Jupyter Notebook

4.2 Deep Learning vs. Traditional ML: Comparative Analysis

Deep learning technology (as well as traditional machine learning). Traditional machine learning models can achieve superior performance compared to deep learning due to several factors. The Neural Network architecture has advantages over many other traditional forms of machine learning where it can discover hierarchical representations from student data. With multiple hidden layers, the Neural Network model can develop increasingly abstract features, starting with simple demographic information and performance indicators through more intricate and behavioral patterns that may not even be easily discernible to a human analyst or captured by a traditional machine learning algorithm.

Of all the various forms of traditional machine learning techniques, Gradient Boosting provided the best overall performance, closely followed by Logistic Regression and SVM Algorithms. The strong performance of Gradient Boosting is likely attributed to its sequential learning capabilities, whereby each model is developed one at a time using feedback from the previous model's performance. Even though Gradient Boosting is able to model some of the most complex patterns that are utilized in dropout predictions, its performance still does not approach that of deep learning models. Therefore, it appears that dropout prediction involves some of the most complex patterns that traditional feature engineering techniques are simply unable to model.

Although the Transformer architecture has a sophisticated attention mechanism, the Performance of the Transformer architecture did not exceed that of a simpler Neural Network model. This may be due to the data being tabular; hence, there are generally fewer sequential dependencies found within tabular data than are found in natural language or time-series data types. The Hybrid Ensemble Model produced very good performance, but it was noted that creating a combination of different machine learning models does not always produce a superior performer, particularly when deep learning models already capture most relevant patterns.

4.3 Feature Importance Analysis

Analysis of feature importance revealed five major variables that significantly influence dropout prediction:

1. Average Grades (moyenne_obtenue): Avg Grades (moyenne_obtenue): This is found to be the most significant variable. Students with falling grades have a significantly higher risk of dropping out, and this is consistent with a mass of research on academic achievement as a criterion for dropping out. The result on grades below 10/20 is an exponential risk increase.

2. Number of Uncompleted Courses (nbre_cours_nonsuivi):

This indicator is an extremely important early warning sign. Those students who do not finish even a few courses experience a drastically higher probability of dropping out, implying that early disengagement occurs prior to formal dropout.

3. Course Completion Rate: "The ratio of finished to enrolled courses measures overall patterns of engagement." It was found that this ratio was more predictive than raw numbers because it offered a relative comparison instead of an absolute one.

4. Course Failure Rate: The rate of out-and-out course failures separate from grades. A pattern of course failures, even when mixed with some successful course completions, strongly predicts eventual dropout, implying cumulative discouragement effects.

5. Student Age (age_etudiant): Older students tend to have high perseverance rates, which may be associated with maturity, experience, or determination to get a degree. Younger students below 20 tend to have higher risk, perhaps indicating a lack of enough preparation or infrastructure. Demographic factors such as gender and geographical location are of less predictive value or at best direct predictors, though these factors could be interrelated with other factors in a complex manner, which can be captured by a deep learning method. Indeed, the major predictors included factors relating to academic performance and activities of engagement.

4.4 Practical Implications for UVCI

The superior performance of the Neural Network model holds several practical implications for UVCI's academic governance in terms of student retention. An intelligent early warning system based on the proposed model of neural networks achieved an accuracy of over 99% in detecting students who might dropout of college. This system holds high credibility in preventing false positives from draining scarce resources of college counseling departments towards non-risk students.

The identified predictive factors are also useful points for intervention. Students demonstrating declining grades can have academics support systems such as tutoring, alternate assignments, or extended deadlines automatically activated for them. Those with a tally of incomplete courses can have special reminders, personalized learning plans, and follow-up meetings with their advisors to bring them back on track to prevent permanent disengagement.

The fact that digital learning platforms are based on real-time technology allows for continuous observation so that risk levels are updated by the system based on fresh information. The system thus allows for far more frequent support sessions compared to the previous system of interventions conducted over a semester. Steps need to be taken to ensure that computer predictions supplement rather than replace human counselors' expert opinion.

5. Perspectives and Future Work

5.1 Enhanced Data Collection

Future versions of the prediction model should include more variables than are currently measured in institutionally based data. These variables could include family support factors such as parental educational levels or socioeconomic status. Connectivity metrics, such as login activity, time spent on course materials, or patterns of accessing resources, may indicate levels of activity that are hidden in data related to course completion. Existing patterns of regular logins could structure a distinction between students in temporarily different situations than persistently disengaged students. Psychometric tests, including self-efficacy, motivation, and psychological readiness for online studies, would offer a warning system for students on the verge of struggling academically through psychometric tests. Indicators specific to financial stress, including payment delays or trends in paying class fees, would help identify those at risk because of financial concerns and not necessarily because they struggle academically.

5.2 Real-Time Predictive Platform

The creation of a real-time intelligent decision-support platform will deliver the next most critical step to UVCI's evolution and will be a powerful and seamless extension to UVCI's Learning Management System. It will collect and integrate live data into UVCI's automated risk assessments and risk profiles.

The platform will contain intuitive dashboards for academic advisors to view the at-risk population based on the following variables: (1) A colour-coded display of risk level (2) Colour-coded trends of whether a student's risk is increasing or decreasing, and (3) The intervention strategies recommended based on an individual's risk profile.

The platform will also provide alerts when a student crosses a critical risk threshold, prompting immediate outreach by the appropriate individuals. The platform will have a longitudinal tracking system to allow an analysis of the effectiveness of interventions and an ongoing improvement of the risk prediction model based on the results of the interventions. This platform must sit at the intersection of Privacy and Ethics, providing the necessary safeguards to ensure that predictive analytics augment and support human decision-making when determining a student's academic future, not replace it.

5.3 Explainable AI and Interpretability

Although deep-learning models outperform other classification methods, their black box nature has made it difficult for institutions to adopt them and for students to embrace them. Researchers should focus on developing architectures that provide the ability to explain their predictions in a manner that is relevant to instructors and students. Researchers could provide educators with insights into what specific factors led to a student being classified as at risk of failure by utilizing attention visualization techniques, SHAP values, and layer-wise relevance propagation.

Explainability not only increases trust in the system but also provides actionable information regarding interventions. An understanding of the reasons a student is determined to be at risk creates an opportunity to assist the student by developing more focused and effective support strategies. In addition, making the decision-making process of the algorithm transparent addresses ethical issues surrounding the automated evaluation of students' capabilities and ensures compliance with the new regulations regarding AI transparency in education, which are currently emerging.

5.4 Intervention Optimization

Dropout prediction, however, remains the starting point, but appropriate intervention methodologies also need to be conceptualized and tested. On this basis, future studies must include randomized controlled trials in an effort to compare the efficiency of personalized tutoring, peer mentorship initiatives, adaptation of the coursework load, financial assistance, and counseling initiatives. Additionally, concepts revolving around machine learning could help analyze the most suitable interventions for the individual, depending on their prevailing risk profile and individual circumstances.

5.5 Generalization and Transfer Learning

Although the current study is directed towards UVCI, it is suggested that the methodology produced by the study may be used by other digital universities in Africa who are experiencing similar difficulties to UVCI. Additionally, transfer learning can help enhance models created from UVCI data at other institutions. As a result of the smaller amount of data and training time needed to create a predictive model that is equally effective, organisations will be able to provide support with less time, cost, and difficulty. For example, when collaborating between institutions to develop a shared predictive model, all collaboratives would benefit from the many different data sources available to them while utilising the federated learning technique to ensure privacy for each institution involved.

5.6 Temporal Dynamics and Early Prediction

Currently, forecasts are based on cumulative information but using earlier forecasts would allow more complete proactive repair options; future systems would include using time series analysis (e.g., RNNs and T-CNNs) to perform early dropout risk predictions of students over time (based on where they are in relation to other students in the same cohort). If schools could identify the at-risk student within the first weeks of enrollment, they would have the highest likelihood of providing timely intervention to help eliminate potential dropout patterns before they become established. Additionally, using sequential pattern mining on students who dropped out of a given cohort may expose significant warning signs in students who appear to be following a similar route to dropout, allowing schools to implement timely intervention services for them as well.

6. Conclusion

In this thorough research, the key question was how to find the most appropriate predictive models and variables that can be used as a basis for identifying students at risk of dropping out at the Virtual University of Ivory Coast. To compare the performance of a total of nine predictive modeling algorithms (five traditional "machine learning" methods and four deep learning architectures), we found that deep learning performed significantly better than the traditional approaches when tasked with predicting student dropout.

The Neural Network Algorithm earned the highest recognition for accuracy (0.9930) and F1-score (0.9888), while the traditional machine learning algorithms had the following highest-rated predictive performance (F1-score of 0.9406 and accuracy of 0.9641): Gradient Boosting. This indicates a five percent difference in performance (in terms of F1-score) that ultimately provides significant advantages for accurately identifying at-risk students and targeting support resources more effectively through the use of deep learning methods to better capture the underlying complexities and non-linearities of the dropout phenomenon.

Our results have identified five important predictive variables: the grade point averages of students, the number of uncompleted courses for students, the completion rate of courses for students, the failure rate of courses for students, and the students' age. This list of predictive factors shows the importance of behavioral measures such as course completion rates compared to demographic variables. The preponderance of indicators of student behavior such as course completion measures compared to demographics is a great point, since student behavior is easier to change relative to demographics.

The practical implications of these results are significant. In order to support students identified as at risk, we suggest integrating a Deep Learning (DL) based Early Warning System (EWS) into the University of Victoria's Community Institute (UVCI) academic governance system through the use of an Operational Dashboard (OD) that provides real-time information about high-risk students and facilitates the implementation of appropriate interventions, such as individualised academic mentoring, proactive outreach to advisors, and dynamically adapting learning pathways based on individual student behaviour and performance data. The high degree of accuracy associated with the neural network model significantly reduces the frequency of false positives and allows limited resources that are available to support students who are deemed to be at risk of academic failure to be directed to only those students who are genuinely at risk.

In addition to technical sophistication, the successful implementation of the DL-based EWS requires embedding the system within a holistic student support structure that incorporates the information generated by the algorithm as well as the judgment and empathy of the individual. Therefore, academic advisers and counsellors continue to play a vital role in placing the risk assessment in proper context, interpreting it based on their understanding of the individual student situation, and providing the appropriate interventions. The role of technology in supporting students should be to augment and enhance the human interaction that is integral to supporting students.

The success of deep learning models documented in this study may have broader implications for many African digital universities. As more universities throughout all areas of Africa use digital learning methods, employing data-informed practices to help students succeed, becomes not only possible but essential. This study has shown that advanced AI methods can be successfully incorporated in African educational settings, providing educational institutions with research-based solutions to long-standing obstacles preventing many students from gaining access to a quality education or completing their education.

Future studies will be devoted towards establishing an online prediction model that incorporates added key metrics (Family Support Indicators, Sheltering Support, Connectivity and Stability), which could allow for earlier detection of students who may need support and provide a comprehensive response. The ultimate goal of these studies is to develop an Intelligent Decision Support System that can identify students who are likely to drop out before they can drop out and convert reactive dropout prevention into proactive means of helping students be successful.

In summary, in this study, it is made clear that deep learning models, specifically NN models, are beneficial tools to be utilized in predicting dropout in students enrolled in digital universities in Africa. Students who are at risk can be identified, and strategies will be implemented by the institution to ensure students' success and meet their mission of offering quality education to all students. With continuing digitalization of higher education in Africa, such tools will play a critical part in these institutions' effectiveness.

References

[1] Abbas, S., Khan, M. A., & Farooq, M. U. (2022). Hybrid machine learning model for student dropout prediction in higher education institutions. *IEEE Access, 10*(1), 33441–33455.

[2] Andrade-Girón, D., Sandivar-Rosas, J., Marín-Rodriguez, W., Susanibar-Ramirez, E., Toro-Dextre, E., Ausejo-Sanchez, J., Villarreal-Torres, H., & Angeles-Morales, J. (2023). Predicting student dropout based on machine learning and deep learning: A systematic review. *EAI Endorsed Transactions on Scalable Information Systems, 10*(5), 1–15.

[3] Kamal, A., & Mustapha, A. (2022). Comparative evaluation of machine learning classifiers for predicting student dropout in open and distance learning universities. *Journal of Applied Research in Higher Education, 14*(6), 2282–2301.

[4] Salehin, I., &Bhuiyan, M. A. (2023). Deep learning-based early prediction of student dropout in online higher education. *Education and Information Technologies, 28*(12), 21123–21145.

[5] Villar, A., & Robledo Velini de Andrade, C. (2024). Supervised machine learning algorithms for predicting student dropout and academic success: A comparative study. *DiscoverArtificial Intelligence, 4*(2), 1–15.

[6] Amani, N. A., & Koffi, C. K. (2023). Intelligence artificielle et réussite académique dans les universités numériques africaines : étude de cas de l'Université Virtuelle de Côte d'Ivoire. *Revue Africaine des Sciences et Technologies de l'Information et de la Communication, 7*(1), 45–62.

[7] Yao, D. K., & Kouadio, F. J. (2023). Gouvernance par les données et transformation numérique de l'enseignement supérieur en Afrique de l'Ouest. *Revue Africaine de GouvernanceNumérique, 2*(1), 75–90.

[8] Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89–125.

[9] Chawla, N. V., Bowyer, K. W., Hall, L. O., &Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357.

[10] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

[11] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.

[12] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., &Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008).

[13] Wolpert, D. H. (1992). Stacked generalization. *Neural Networks, 5*(2), 241–259.

[14] Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery, 10*(3), e1355.

[15] Hellas, A., Ihantola, P., Petersen, A., Ajanovski, V. V., Gutica, M., Hynninen, T., Knutas, A., Leinonen, J., Messom, C., & Liao, S. N. (2018). Predicting academic performance: A systematic literature review. In *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education* (pp. 175–199).