ijar
ISSN NO. 2320-5407

Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI:10.21474/IJAR01/17839
DOI URL: http://dx.doi.org/10.21474/IJAR01/17839

RESEARCH ARTICLE

MODEL OF A PERSONALISED E-LEARNING PROCESS BASED ON A DECISION TREE ALGORITHM

Petey Kragbi Olivier[1], Kone Tiémoman[2] and Kanga Koffi[3]

1. Virtual University of Cote d'Ivoire (UVCI), Digital Research and Expertise Unit - Abidjan, Côte d'Ivoire.
2. Virtual University of Cote d'Ivoire (UVCI), Digital Research and Expertise Unit - Abidjan, Côte d'Ivoire.
3. African Higher School of ICT (ESATIC), LASTIC (Laboratory of Information and Communication Sciences and Technologies) - Abidjan, Côte d'Ivoire.

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

Manuscript Info

……………………….

Manuscript History
Received: 10 September 2023
Final Accepted: 14 October 2023
Published: November 2023

Key words:
Personalized And Adaptive E-Learning, Artificial Intelligence, The Decision Tree Algorithm, Input Component, Learning Component, Output Component

Abstract

………………………………………………………………

The supply and demand for training, having undergone a revolution, involves information and communication technologies including artificial intelligence. However, following training adapted to the evolution of the learner's skills remains a challenge. Our study aims to provide a solution aimed at promoting a personalized online learning process. Our approach consisted of choosing a decision tree algorithm following a comparative study and developing an architecture based upstream on the evaluation of the learner's knowledge. This architecture directs the learner, according to their performance, towards educational resources (documents, courses, sections, videos, etc.) or learning devices, in an iterative and incremental manner until the end of the learning process. The results obtained reside in the proposed model based on the gradient boosting algorithm adapted to the personalization of human learning. This model takes into account three essential components driven by artificial intelligence and covers an entire personalized learning process from checking prerequisites to the end of successful learning.

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

Introduction:-

The constant evolution of digital technology and training needs, towards greater efficiency, greater flexibility and cost reduction, has led to the emergence of educational and IT tools. The supply and demand for training, having undergone a revolution, involves information and communication technologies. Educational and training activity is inevitably called upon to be redefined, in other words to be the subject of innovation. Therefore, training without spatio-temporal constraints becomes an ideal opportunity to exploit. The constant evolution of digital technology and training needs, towards greater efficiency, greater flexibility and cost reduction, has led to the emergence of educational and IT tools. The supply and demand for training, having undergone a revolution, involves information and communication technologies. Educational and training activity is inevitably called upon to be redefined, in other words to be the subject of innovation. Therefore, training without spatio-temporal constraints becomes an ideal opportunity to exploit.However, following training adapted to the evolution of the learner's skills remains a challenge [1]. It should be noted that the quality of service provided during training is the result of the capacity of new approaches to provide learners with appropriate educational resources [2]. Our approach aims to design a

Corresponding Author:- Petey Kragbi Olivier
Address**:-** Virtual University of Cote d'Ivoire (UVCI), Digital Research and Expertise Unit - Abidjan, Côte d'Ivoire.

personalized e-learning system. We opted for the use of a decision tree algorithm. This choice is motivated by the structuring of course content in tree form which easily lends itself to classification.

This paper is organized as follows:
- section 2 deals with the state of the art,
- section 3 highlights the problem,
- section 4 presents the methodology
- section 5 describes our contribution,
- the last section includes the conclusion and perspectives.

State of the art

The digital revolution has led to a proliferation of distance learning systems. From then on, certain difficulties emerged, notably the personalization of training. Related work has been carried out with the aim of finding solutions.

Personalized Resources

The digital revolution has led to a proliferation of distance learning systems. From then on, certain difficulties emerged, notably the personalization of training. Related work has been carried out with the aim of finding solutions.

A “Clavertesting” plug-in architecture was proposed by Mohammed Boussakuk and his collaborators [3]. This assessment architecture is based on five components: the domain model, the learner model, the test editor, the item bank and the assessment engine. The actors involved being the evaluator, the tutor and the learner. A semi-automatic generator connects to the bank of exercises, selects and presents to each learner tailor-made items based on the learner's mastery of the subject, their profile and the answers to the questions previously administered.

According to our apprehension, these studies would not take into account an entry system whose role is to test the prerequisites. These prerequisites allow for good understanding and facilitate learning.

Personalized educational activities

Marouane Birjali et al [4] proposed an online learning model based on Big Data using knowledge based on social learning skills and activities. This study aimed to optimize personalized online learning through big data. A model and an associated algorithm have been developed. This ant colony optimization algorithm was based on MapReduce, generating an adaptive learning path for each learner. This model is certainly interesting but has limitations linked to insufficient information if the learner is not on social networks. Also it does not take into account the entire learning process.

A multi-layer multi-agent architecture for an educational adaptation of distance learning has been the subject of investigation by Zouhair et al[5]. This focused on the design and creation of a computer system capable of initiating learning and managing individualized teaching and monitoring. Predicting and reducing the number of dropouts was the objective. One approach was based on MultiAgent Systems (MAS), capable of cooperating and coordinating their actions to provide educational adaptation by learner profile based on traces. The architecture was based on four layers of agents resulting in a pyramidal relationship. The interesting results showed that the entire architecture makes it possible to evaluate the scenarios found as changes occur in the observed situation. However, based on traces, to our knowledge this architecture would not sufficiently take into account the performance of the learner in terms of evaluating his renderings.

Work takes into account the elements for piloting personalized courses. This work carried out by Azziz Anghour et al [6] proposed a methodology for constructing and piloting training courses. Educational grain selection criteria constituted the course based on the profile of the learner and the appreciation of the users. This system architecture was coupled with two designed ontologies: The profile ontology and the domain ontology governed by a management system based on a controller, monitoring agents and a scripting engine.

It should be noted that in view of the information available to us, the various works as a whole have once again omitted the prerequisites which represent determining aspects in learning, so this work would not sufficiently highlight the evolution of performance of the learners.

Personalized assessments

A model focused on adaptive, individualized and equitable assessment was the subject of a study conducted by Mohammed Boussakuk et al [3]. The objective was to promote the learner's acquired knowledge, with a view to creating a new intelligent environment which guarantees an adaptive evaluation of the learner according to his cognitive state, by taking advantage of artificial intelligence techniques, theories in cognitive psychology, educational sciences, pedagogy and didactics.

This study resulted in the implementation of a methodology for creating personalized courses and their management. However, it would not take into account multi-peer piloting for learning in a collaborative situation, and the question of the learning process is lightly addressed and even less that of prerequisites.

A personalized assessment system using Web services and allowing the selection and presentation of learning and assessment resources adapted to the learner's level of knowledge was designed by Lilia ChenitiBelcadhi et al[7]. This work aimed to personalize assessments through web and mobile platforms with a view to more adapted learning. To this end, the researchers implemented two architectures: an E-evaluation architecture based on a web platform and an M-evaluation architecture based on a mobile platform. According to our understanding, the assessments would essentially consist of choice questions without specifying its compliance with standards or benchmarks of good practice in this area. Which could pose a reliability concern.

In the same vein, authors D. Bañeres et al[8] and Sara Ouald Chaib et al[9] respectively propose a model that takes into account several pieces of evidence in order to generate the next evaluation and a platform based on adaptive learning for student self-study in scientific French.

These systems as a whole would be focused on the adaptive generation of assessment content. They would not emphasize learning resources and the diversification of the evaluation method. Also they would not take into account the entire learning process.

Problematic

From this literary review, it appears that interesting work has been carried out with the aim of providing a response to the personalization of e-learning training. This work focused on educational resources, educational activities, and assessments.

However, these works and tools produced have limitations because they do not cover the entire training process. For example, prerequisites which represent a requirement in terms of prior knowledge are not taken into account. Also these works do not personalize resources, activities and evaluations. The question that arises in these conditions is the following: is there a reference model for personalized e-learning based on artificial intelligence? A response to this concern will be useful in dealing with the diverse and dynamic nature of learners' abilities.

Methodology:-

Our approach is as follows:

- We opted for a decision tree algorithm to meet a classification need. It should be noted that each learner represents a unique reality. This singularity is the result of belonging according to a given tree structure to categories and subcategories.
- A comparative study is developed with a view to choosing the decision tree algorithm to use.

The criteria taken into account are the data type, ease of implementation, robustness to noise, robustness to outliers, binary and multiple branching, maximization of precision.

- A functional architecture[10] aligned with the learning system and its subsystems. The learning system[11] for a training module is composed of three subsystems including the input system, the learning system and the output system.

Contributions

Comparative study of decision tree algorithms

We compare decision tree algorithms[12] in order to make an optimal choice.

Table 1:- Comparison of decision tree algorithms.

Criteria / Algorithme	Data type	Ease of implementation	Robustness to noise	Robustness to outliers	Binary and multiple connection	Maximizingprecision
ID3 Algorithm	Discreet and homogeneous	yes	yes	yes	no	no
C4.5 Algorithm	Heterogeneous data	yes	no	no	yes	no
CART Algorithm	Heterogeneous data	yes	no	no	yes	no
CHAID Algorithm	Categorical	yes	yes	yes	yes	no
QUEST Algorithm	Continuous and heterogeneous	yes	yes	yes	yes	no
MARS Algorithm	Continue	yes	no	no	no	no
Randomforestalgorithm	Categorical	no	no	no	no	yes
Gradient Boosting algorithm	Heterogeneous	no	yes	yes	yes	yes

Looking at this table, we see that the Gradient Boosting algorithm[13] has more advantages because it satisfies a greater number of performance criteria.

The proposed model is based on recursive decision tree algorithms, suitable for the personalization of human learning.
The model takes into account three essential components:
-the input system component (ISC),
-the learning system component (LSC),
-the output system component (OSC).
This is an architecture involving gradient boosting. Gradient boosting after classification will allow redirection to the appropriate component. This redirection may involve the following component:
- the overall learning component which combines the three components mentioned above. It is a training system comprising a broader training program.
- the learning unit (LU) which is a section or subsection of the learning system component (LSC).
- the elementary learning unit (ELU) which represents a smaller learning unit. We could speak of an atomic unit.

The Input System Component (ISC)
This component is responsible for intelligent management of the entry system.

Description of the constituent elements of the input system component
The constituent elements of the entry system component architecture are defined as follows:
- Gradient boosting: machine learning algorithm allowing classification within the framework of our study.
- A transition engine between components

- An engine for defining objectives
- An objective verification engine.
- A prerequisite check engine.

Architecture of the input system component

The architecture designated in Figure 1 shows our proposed solution at the input system level.

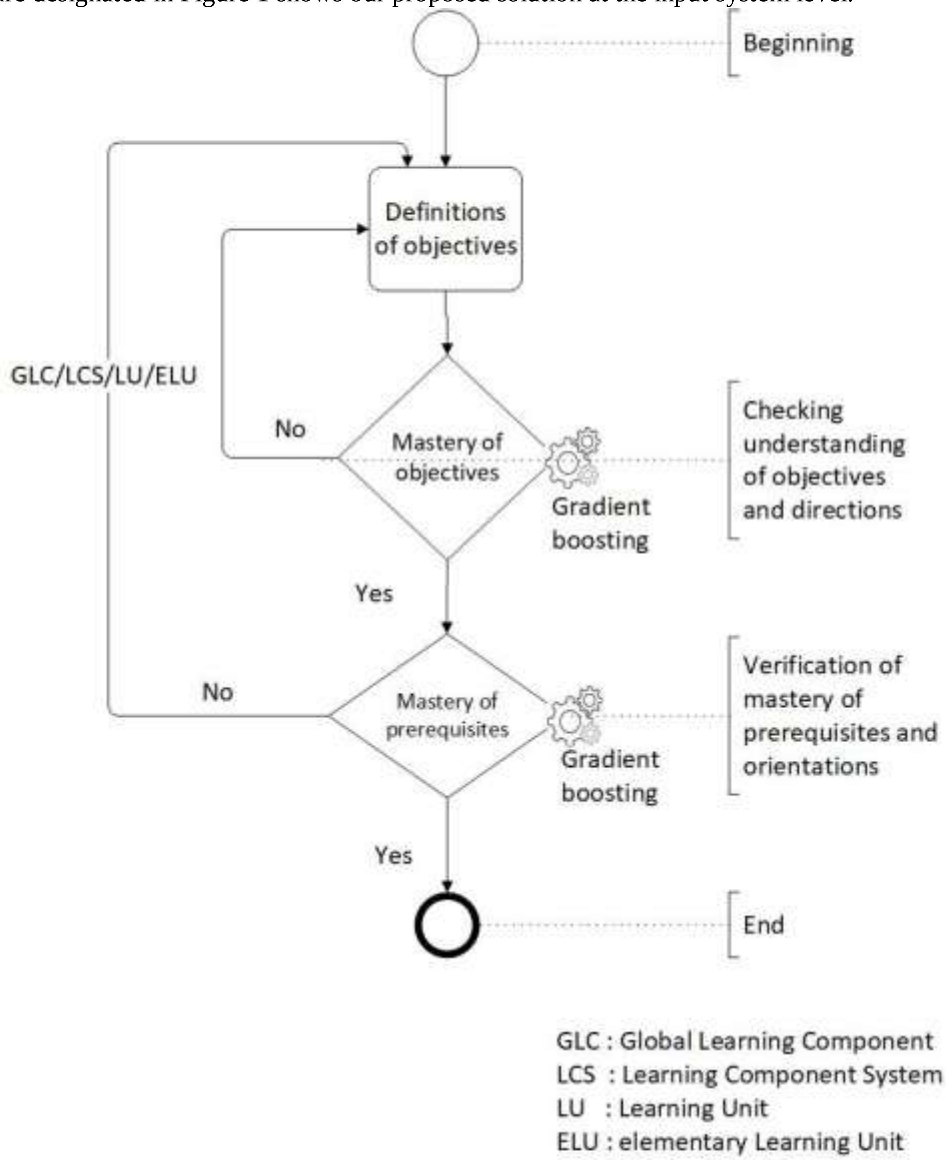

Figure 1:- Input Component System (ICS).

Operation of the entry system component.

Step 1: The system must define the training objectives

Step 2: Through its objective verification engine, ensure that the objectives are well perceived. If this is the case, we move on to checking the prerequisites. Otherwise, the objectives will be reformulated and verified iteratively until a satisfactory result is obtained. The Gradient boosting algorithm will be responsible for classifying the learners based on the results obtained.

Step 3: Proceed to check the prerequisites. If the test proves satisfactory, we move on to the learning system component (LSC), otherwise an appropriate training system will be recommended to the learner to enable him or her to have the necessary prerequisites. The Gradient boosting algorithm will be responsible for classifying learners based on the results obtained.

The learning system component

This component is responsible for controlling the learning system. The architecture designated in Figure 2 presents our proposed solution at the learning system level. This architecture is also based on elements cited in the entry system component.

Description of the elements of the learning system component

The learning system component requires the same building blocks as the input system component. This component also has module-based learning processes. Each learning process includes educational activities and resources.
Educational activities can include, among others:
- web conferencing allowing real-time sharing of audio, video, slides and screen, as well as chat, a multi-user whiteboard, breakout rooms, polls and emojis.
-chats that allow participants to have a synchronous discussion in real time, in text mode.
- forums allowing participants to hold asynchronous discussions, that is to say not requiring their participation at the same time.
- workshops helping to collect and examine the work of participants, and to have them evaluated by peers.
- lessons that help the teacher to offer content and/or exercise activities in an interesting and flexible way.
- homework which allows a teacher to communicate tasks to participants, collect work and provide them with feedback and grades.
- tests allowing the teacher to create tests with questions of various types, including multiple choice, true-false, matching, short answer or calculated questions.

Resources include the following:
- files
- videos
- hypertext links

The learning system component also has:
- A transition engine between components
- A learning verification engine

Architecture of the learning system component

The architecture in Figure 2 presents our proposed solution at the learning system level.

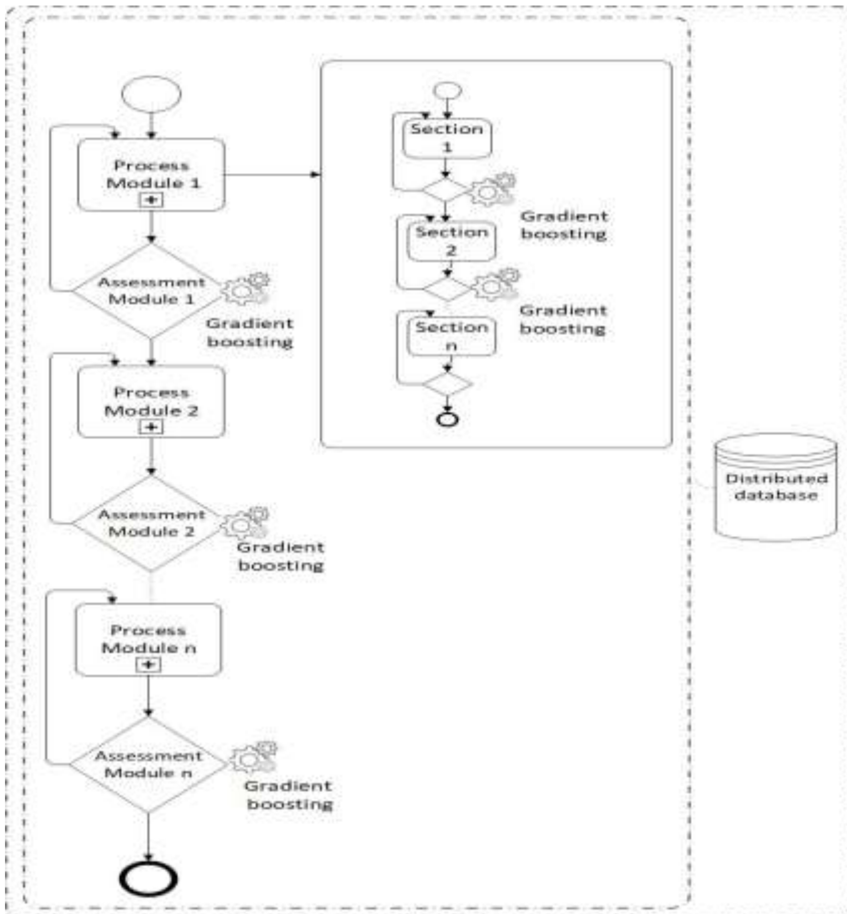

Figure 2:- Learning System Component (LSC).

How the learning system component works.

A training module is established in a process.

A training module can be broken down into two or more sections. We have different scales.

- At the scale of a section (or subsection):

Each section (or subsection) contains activities and educational resources. At the end of each section, a partial evaluation can be carried out. The results of the evaluation will be collected and reinjected into the Gradient boosting algorithm which after carrying out a classification which will serve as a basis for decision-making. On the one hand, the decision can be to direct the learner towards a learning sub-unit, a learning unit, a learning component, or an overall learning component. On the other hand, the decision may consist of moving on to the next section.

- At the scale of a module:

At the end of each module an evaluation is carried out. The results collected from the evaluation will be used in the Gradient boosting algorithm which after classification will give orientations in accordance with the scale diagram of the section (or section). But this time, success in the module evaluation will result in the next module.

On the scale of a training program (comprising several modules)

Here too we will see almost the same type of operation. It should be noted that this is a recursive and incremental process at different scales.

The output system component

Just like the components described previously, this component is responsible for intelligently controlling the output system.

Description of the constituent elements of the output system component (OSC)

The output system component is composed of:

- a bank of items [3] from the final test activities
- a learning verification engine[3].
- a transition engine between components

Architecture of the output system component.

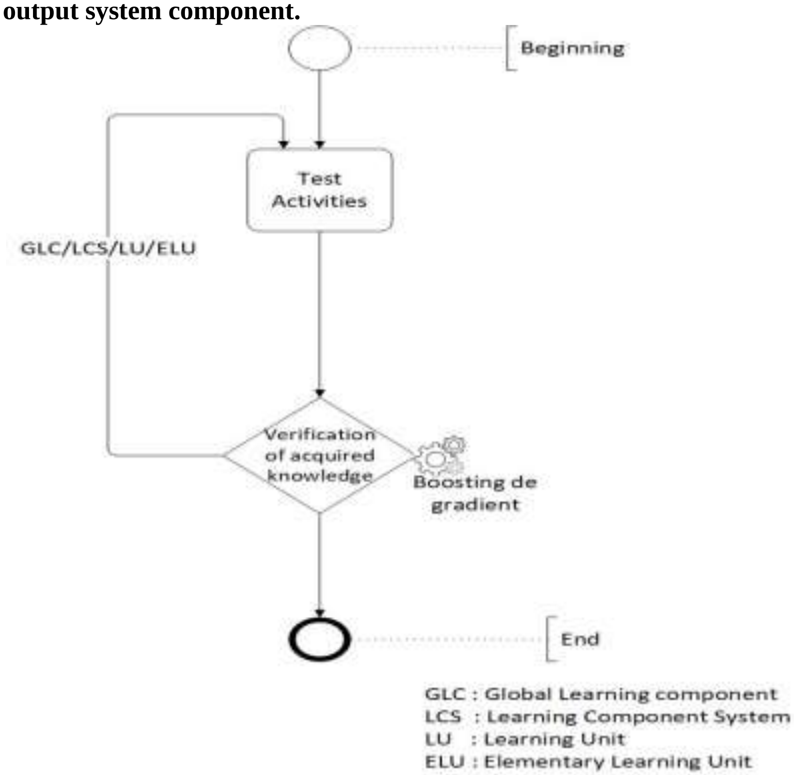

Figure 3:- Output System Component (OSC).

Operation of the output system component.

At the end of the training program, a summative evaluation is carried out. Depending on the results obtained, Gradient boosting will be responsible for classifying learners based on their performance during the assessments. This Classification will be used to guide the educational path.

If the training program is successful, otherwise the learner will be redirected to an appropriate learning system.

Modeling of gradient boost for personalized e-learning.

Gradient Boosting modeling for adaptive learning in e-learning in the context of learner classification involves the creation of a predictive model capable of classifying learners into different categories based on specific characteristics. Here are the general steps you can follow:

Modeling Steps for Classification

Learner data collection:

- Collects learner data, including characteristics such as time spent on learning materials, quiz results, interactions with the learning platform, etc. But the framework of our study we emphasize the results of the evaluations.
- Classification labels can be based on criteria such as successful completion of a module, skill level achieved, etc.

Data Preprocessing:

- Treatment of missing values and outliers.
- Encoding categorical variables.
- Normalize or scale the features if necessary.

Problem definition :
- Definition of classification categories (e.g. high, medium, low).
- Choosing an appropriate loss function for classification (e.g. cross-entropy for multi-class classification).

Choice of Gradient Boosting Algorithm:
- Choice of a Gradient Boosting algorithm (XGBoost, LightGBM, CatBoost) adapted to the classification.
- Configure hyperparameters[14], such as learning rate, tree depth, number of iterations, etc.

Model Training:
- Divide the data into training and testing sets.
- Train the model on the training set using Gradient Boosting.
- Optimize hyperparameters if necessary using cross-validation.

Model Evaluation:
- Evaluation of model performance on test set using metrics such as precision, recall, F1-score, confusion matrix, etc.
- Analyze the results to understand the model's ability to classify learners.

Dynamic Adaptation to learner:
- Use the trained model to predict the learner category.
- Use these predictions to personalize the learning experience based on the assigned category.

Reiteration and Improvement:
- Collecting new data periodically to update the model.
- Retrain the model to reflect changes in learning behaviors.
An adjustment of the model parameters must be taken into account to avoid overfitting.

Dynamic adaptation requires a system interface that can respond to model predictions to personalize the learning experience.

This modeling approach can be refined based on the specific characteristics of your data and the particular classification goals for your e-learning adaptive learning system.

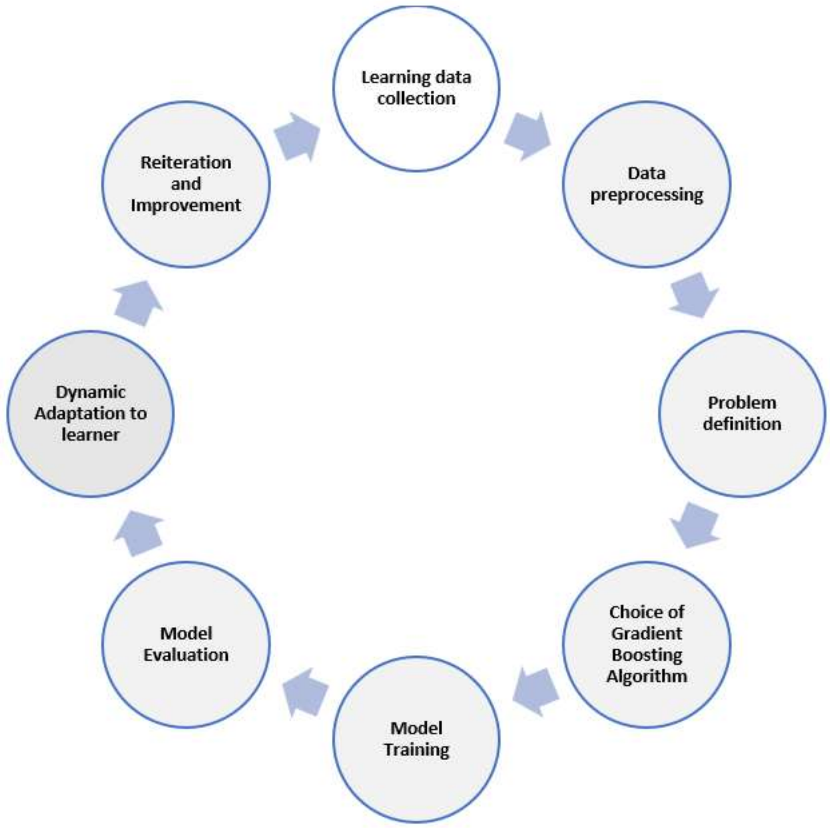

Figure 4:- Simplified diagram of the steps of gradient boosting modeling for classification.

These sequences implicitly represent a process of continuous improvement[15]. This continuous improvement will allow the model to be more and more efficient. This good performance will lead to better adaptation of e-learning according to the learner's profile[16].

Mathematical formulation of gradient boost for personalized e-learning.
Notations et symbole

- i: Index of the learner, $i=1,2,....,N$
- xi: Features of learner i.
- y_i : True class label of learner i.
- $F_t(x_i)$: Model at iteration t for learner i.
- $L(y_i,F_t(x_i))$: Loss function for the prediction of learner $i$at iteration t.
- α: Learning rate, a hyperparameter controlling the contribution of each weak model.
- T : Total number of iterations or weak models.

Mathematical Formulation
The mathematical formulation of the gradient boost for personalized e-learning in the case of a classification is described in the figure below:

1. **Initialisation :**

$$F_0(x_i) = \arg\min_\gamma \sum_{i=1}^{N} L(y_i, \gamma)$$

2. **Itérations pour $t = 1$ à T :**

a. **Calcul du résidu :**

$$r_{it} = -\frac{\partial L(y_i, F_{t-1}(x_i))}{\partial F_{t-1}(x_i)}$$

b. **Ajustement du modèle faible (arbre de décision) :**

$$h_t(x_i) = \arg\min_h \sum_{i=1}^{N} L(y_i, F_{t-1}(x_i) + \alpha h(x_i))$$

c. **Mise à jour du modèle :**

$$F_t(x_i) = F_{t-1}(x_i) + \alpha h_t(x_i)$$

3. **Prédiction finale :**

$$\hat{y}_i = \mathrm{argmax}_y F_T(x_i)$$

Figure 5:- Mathématical formulation.

The choice of the loss function $L(y_i,F_t(xi))$ depends on the specific nature of your classification problem (e.g., cross-entropy for multi-class classification).

Partial derivatives $\partial L(y_i,F_t(x_i))/ \partial F_t(x_i)$ are computed based on the chosen loss function.

The depth of trees and other decision tree parameters can be adjusted to control model complexity.

This mathematical formulation represents the iterative process of Gradient Boosting for the classification of learners in the context of adaptive e-learning.

Conclusion:-

The objective of this work is the implementation of an e-learning personalization model to improve the online training offer. To do this, we reviewed the most representative works to our knowledge. From this work, we identified the limits and made our contribution which consists of personalizing an e-learning system based on a decision tree algorithm. The model ensures a tailor-made learning experience by taking into account the specific needs of each learner. By offering relevant and adapted content, the model promotes learner engagement and motivation. By providing personalized educational pathways, the model aims to improve learning outcomes and knowledge retention. Future work is part of a perspective of improvement which could take into account a large number of parameters and involve neural networks given the complexity of the data.

Bibliography:-

[1] Elghibari, Fatiha, Elouahbi, Rachid, and El Khoukhi, Fatima. model of a learning process adapted to the cognitive developments of the learner profile. In: from the International Workshop on Pedagogical Approaches & E-Learning, Fez, Morocco. 2015.
[2] Anghour, Azziz, Lamolle, Myriam, Belhadj, Farès, et al. Real-time adaptive learning by multi-agent system. ISI, 2018, vol. 23, no. 2.
[3] Boussakuk, Mohammed, Bouchboua, Ahmed, Elghazi, Mohammed, et al. Towards an adaptive, individualized and equitable assessment. Mediterranean Telecommunications Review, 2018, vol. 8, no. 2.
[4] Birjali, Marouane, Beni-Hssane, Abderrahim, and Erritali, Mohammed. A novel adaptive e-learning model based on Big Data by using competence-based knowledge and social learner activities. Applied Soft Computing, 2018, vol. 69, p. 14-32.
[5] Zouhair, Abdelhamid, Boukachour, Hadhoum, Person, Patrick, et al. A multi-layer multi-agent architecture for educational adaptation of distance learning. 2010.

[6] Anghour, Azziz, Delmas, Guylain, and Lamolle, Myriam. Elements for managing personalized routes. In: 7th Conference on Computing Environments for Human Learning (EIAH 2015). 2015. p. 423-425.
[7] Cheniti-Belcadhi, Lilia, Henze, Nicola, and Braham, Rafik. Implementation of a personalized assessment web service. In: Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06). IEEE, 2006. p. 586-590.
[8] BañeresBesora, David, BaróSolé, Xavier, Guerrero Roldán, Ana Elena, et al. Adaptive e-assessment system: A general approach. International Journal of Emerging Technologies in Learning, 2016, 11 (7), 2016.
[9] Chaib, Sara Ouald, Joti, Imane, and Khoulji, Samira. An adaptive evaluation platform for learning in scientific French: case of the Faculty of Sciences of Tétouan.
[10] Ramazani, Trésor Poyo, Lwango, Jormy Kabwe, Malonga, David Benda, et al. Development and deployment of an online training platform (E-LEARNING) based on service-oriented architecture within academic establishments, University of Likasi. Africa SCIENCE, 2023, vol. 22, no. 1, p. 113-125.
[11] Depover, Christian, De Lièvre, Bruno, Quintin, Jean-Jacques, et al. Pedagogical structuring of an EAD course. Training and Learning Technologies (TECFA). University of Geneva, Switzerland, 2000.
[12] Rakotomalala, Ricco. Decision trees. Modulad Review, 2005, vol. 33, p. 163-187..
[13] Ferraris, Laure, Liautaud, Paul, and BOREL, Émile. Gradient Boosting method. 2022.
[14]Heutte, Laurent, Bernard, Simon, Adam, Sébastien, et al. On the selection of decision trees in random forests. In: 10th International Francophone Conference on Writing and Documents (CIFED). Written Communication Research Group, 2008. p. 163-168.
[15] Lorino, Philippe. On the importance of pragmatism for organization theory. 2019.
[16]Zaoudi, Mohammed et Belhadaoui, Hicham. Adaptive E-learning: Adaptation of content according to the continuous evolution of the learner during his training. In : Proceedings of the 3rd International Conference on Networking, Information Systems & Security. 2020. p. 1-6.