Design of a Smart MOOC Trust Model: Towards a Dynamic Peer Recommendation to Foster Collaboration and Learner's Engagement

https://doi.org/10.3991/ijet.v17i05.27705

Khadija Elghomary[1(✉)], Driss Bouzidi[1], Najima Daoudi[2]
[1] Mohammed V University in Rabat, National School of Computer Science and Systems Analysis (ENSIAS), Rabat, Morocco
[2] Information Science School (ESI), Rabat, Morocco
khadija.elghomary@um5r.ac.ma

Abstract—Recent evolutions in the Internet of Things (IoT) and Social IoT (SIoT) are facilitating collaboration as well as social interactions between entities in various environments, especially Smart Learning Ecosystems (SLEs). However, in these contexts, trust issues become more intense, learners feel suspicious and avoid collaborating with their peers, leading to their demotivation and disengagement. Hence, a Trust Management System (TMS) has become a crucial challenge to promote qualified collaboration and stimulate learners' engagement. In the literature, several trust models were proposed in various domains, but rarely those that address trust issues in SLEs, especially in MOOCs. While these models exclusively rank the best nodes and fail to detect the untrustworthy ones. Therefore, in this paper, we propose Machine Learning-based trust evaluation model that considers social and dynamic trust parameters to quantify entities' behaviors. It can distinguish trustworthy and untrustworthy behaviors in MOOCs to recommend benign peers while blocking malicious ones to build a dynamic trust-based peer recommendation in the future phase. Our model prevents learners from wasting their time in unprofitable interactions, protects them from malicious actions, and boosts their engagement. A simulation experiment using real-world SIoT datasets and encouraging results show the performance of our trust model.

Keywords—Trust Management System (TMS), Social Internet of Things (SIoT), Machine Learning (ML), smart education, Massive Open Online Course (MOOC), peer recommendation

1 Introduction

In the most recent decade, due to the prominent evolution of ICT and the advent of the Internet of Things (IoT) paradigm, the physical and virtual worlds will In the most recent decade, due to the prominent evolution of ICT and the advent of the Internet of Things (IoT) paradigm, the physical and virtual worlds will increasingly be distanced from each other [1]. Cyberspace becomes a part of real space while constituting a

pervasive space called ubiquitous computing [2]. Portable computing devices like smartphones, tablets, and wearables have become an integral part of our daily lives. In addition, various researchers have explored the possibilities of incorporating the concept of social networks in the IoT ecosystem. This integration has led to a new paradigm of the SIoT that represents a suitable platform for better interactions between people and things [3], [4]. Figure 1 shows the general evolution of connected things.

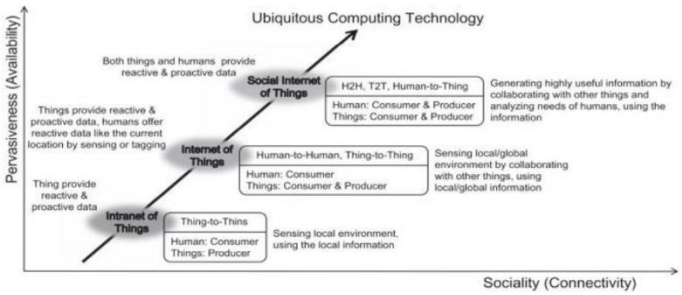

Fig. 1. Evolutionary history of Ubiquitous Computing Technology (Antonio et al., 2014)

Moreover, the accelerated evolution of computing technologies led to a considerable increase in the number of applications and services expected to exceed 75.44 billion by 2025 as shown in Figure 2:

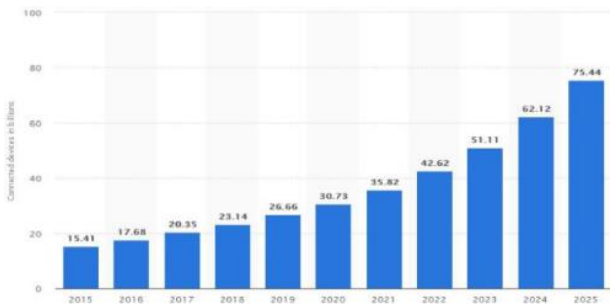

Fig. 2. Internet of Things – number of connected devices worldwide 2015-2025 (Statista, 2021)

Therefore, this impressive progress of ICT has strongly affected several areas and sectors including the education field. Education is greatly reconstructed in the most recent decades by the integration of IoT technologies. We refer to a new concept which is 'Smart education' or "Smart Learning" that describes learning in an intelligent era and that provides a facility to the learner for learning at any place and any time by using smart devices to learn knowledge, acquire skills and connect with their peers [5]. Indeed, the rapidly expanding possibilities of ICT in the education area have enabled the emergence of novel collaborative systems like MOOCs. These ecosystems revolutionize traditional education methods and attract attention in academic and industrial areas. They represent the famous category of Smart Learning [6]. How-

ever, in these contexts characterized by a big number of participants, with intensive interactions, heterogeneous communications, and various devices, learner engagement and completion are problematic [5],[7]. Thus, trust issues arise from the search for a trustworthy peer that can provide the desired service. This situation leads to learner demotivation and disengagement. Trust models could be adopted successfully in this context to help learners by selecting the most appropriate peer to overcome their learning difficulties and maintain their motivation. In general, trust has been widely used in diverse areas to improve the quality of social networking by fighting malicious peers, selecting appropriate partners or service providers, and enhancing the decision-making process. The definition of trust that we derive for our research and in the context of a pervasive world and the ubiquitous computing (IoT and the SIoT paradigm) is: "a qualitative or quantitative property of a trustee, evaluated by a Trustor as a measurable belief, subjectively or objectively, for a given task, in a specific context, for a specific period" [6]. Whereas Trust management is, "mechanism used to ensure trust in various types of systems, his role consists of computing a trust score, which will help nodes to decide on invoking or not, services provided by other nodes" [8]. Trust is a relationship including at least two entities: a "Trustor" entity and a "trustee" entity [9]. The former represents an entity that is supposed to initiate an interaction with another entity, while the latter is the second entity that provides the necessary information (knowledge, content, service) to the Trustor at its request [6]. Moreover, trust has several characteristics and properties, it is asymmetric transitive, propagative, and very dynamic [10], [11]. Ultimately, the trust evaluation process is dynamic in research our context. It involves the Trustor, Trustee, and the underlying context. Thus, the Smart learning Environment (SLE) network comprises users (learners) and devices owned by users. Rules are set by the owner (learner) to create relationships and to provide or obtain services from other objects. Figure 3 describes the idea of our Smart Learning context.

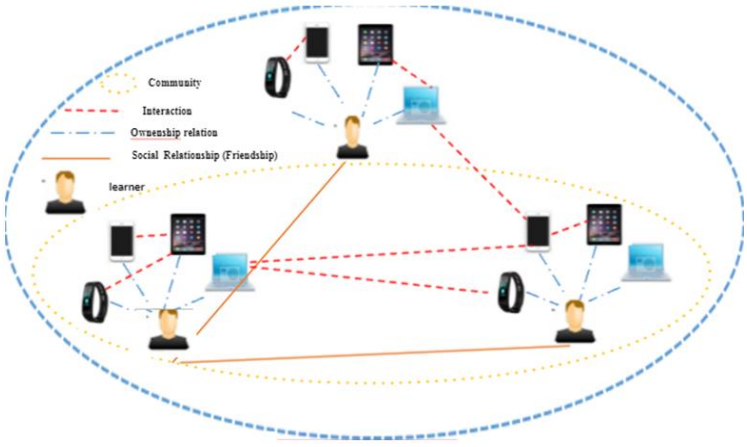

Fig. 3. Illustration of Smart Learning context

To the best of our knowledge, our work is the first that addresses trust evaluation issues among entities (learners and devices) in pervasive learning environments particularly MOOCs. Some works have addressed trust in MOOCs focusing on trust in platforms and MOOCs providers [12], [13]. However, our research handles social trust and is interested in trust among learners to ensure efficient collaboration. In addition, our work is the first that suggest a dynamic Peer Recommender Framework based on the proposed MOOC trust model that due to space constraints it will be presented in future work. In this work, we propose a new trust model based on new trust features derived from OSN and SIoT ecosystems since MOOCs resemble these contexts and have shared characteristics like openness, mobility, and dynamicity, a massive number of participants, and heterogeneity of the components.

In general, the main contribution of this research to the existing literature is that it produced results related to the concept of Trust and Trust models related to learners in MOOCs, an area in which there is currently limited research [12]. Then, we can present other scientific contributions that are summarized as follows:

- Analyzing recent works of trust evaluation in OSNs and SIoT ecosystems with a focus on trust models based ML methods, considering that trust models in these contexts are very advanced and the research on these models is in notable evolution.
- Design of a smart trust evaluation based on classification algorithms to predict the trustworthiness of each partner in future transactions.
- The proposed trust model will be the basis for dynamic peer recommendation. It is flexible and can be used in different application scenarios such as ubiquitous systems and large-scale collaborative systems.

The remainder of the paper is organized as follows. Section 2 reviews and analyzes the recent works of OSN and SIoT trust evaluation based on Machine Learning in the literature. Section 3 introduces and explains the proposed trust evaluation model. Section 4 covers and describes the methodology and material adopted in simulation setup, results comparison, and discussions. Finally, Section 5 concludes this paper and discusses the future works.

2 Literature review

In the literature, several trust models are proposed. So, to choose the most appropriate Machine Learning algorithm for handling trust evaluation concerns, some of the OSN and SIoT trust evaluation models based on ML suggested over the last few years were examined. Moreover, considering that in our previous works [14], [15], [16], we have given an examination and a study of the relevant OSN and SIoT trust models used traditional methods like weighted sum, fuzzy logic, and Bayesian belief [17], [18]. In this section, we have reviewed relevant as recent trust management schemes based on ML approaches.

2.1 Comparative study

In [19], researchers proposed a trust model between users on Facebook. Features are extracted from user interaction information and profile information. KNN, SVM, and MLP are used to predict trust levels. MLP provides the highest accuracy rate. In [20], the authors realized trust evaluation as a classification problem based on the SVM technique. The work of [21] presented trust model-based MLP based on the node's Packet Delivery Ratio (PDR) and set a threshold to distinguish them. In [22], the authors used the trust values calculated by a traditional method and some additional information as training features. They employ an LR method to classify nodes. The results showed that trust evaluation-based ML has higher accuracy by comparing it with other traditional methods. In [23], Eight ML methods were tested. Results showed that the performance of trust evaluation using LR and Neural Network (NN) was the best. The paper of [6] proposed a trust assessment model for IoT services. They used unsupervised ML (k-means) and supervised ML (SVM classifier) to combine six trust factors and classify trustworthy and untrustworthy nodes. In [24], authors expose a trust model that used SVM to aggregate trust features and compute trust among entities. In [25], researchers utilized ML techniques instead of traditional methods to classify vehicles into trustworthy and untrustworthy. They use real IoT data set to perform ML classifiers precisely SVM and KNN. Researchers in [8] proposed a trust model based on attributes derived from the description of the principal trust-related attacks cited in the literature. Their trust model can detect malicious nodes and isolate them for a resilient network. Recently, the previously mentioned researchers in [26] proposed TMS-based MLP able to detect malicious nodes and the types of attacks they have made. Table 1 presents the comparison of previous works according to four criteria:

Table 1. Comparison of OSN and SIoT Trust model based machine learning

Work	Dataset used	Trust Features	ML Technique	Effectiveness indicators
[19]	Facebook	User interaction Information and profile information	MLP/KNN/ SVM	MLP accuracy: 0.83
[20]	Weibo OSN	9 user features and their relationships	SVM	Precision:0.83/ Recall: 0.97
[21]	Ad hoc Networks	Packet Delivery Ratio (PDR)	MLP	Accuracy:0.98
[22]	Weibo OSN	Traditional trust value and auxiliary information	LR	Accuracy: 0.90
[8]	Sigcomm2009	7 features: Honesty/Reputation/Similarity/Direct Experience/Rating /frequency/ Quality of Provider/ Rating trend	Naives Bayes, Random Tree, MLP	MLP Precision: 0.9253/Recal: 0.92
[23]	Twitter OSN	user features (profile, behavior and interaction information)	8 models including SVM and LR	SVM accuracy: 0.993 LR accuracy :0.996
[24]	Sensor Networks	3 features: Communication trust, Packet trust, Energy Trust	SVM	Accuracy: 0.97
[6]	Sigcomm2009	5 features: Co-location relationship (CLR)/Co-	SVM	Precision: 0.89

Work	Dataset used	Trust Features	ML Technique	Effectiveness indicators
		work relationship (CWR)/Mutuality and Centrality (MC)/Cooperativeness Frequency-Duration (CFD)/ Reward		
[25]	Sigcomm2009	3 features: similarity/familiarity/Packet Delivery Ratio (PDR)	KNN, SVM	KNN accuracy: 0.90 SVM accuracy: 0.65
[26]	Sigcomm2009	7 features: Honesty/ Reputation/ Similarity/ Direct Experience/ Rating frequency/ Quality of Provider/ Rating trend	MLP	Precision:0.95 /Recall: 0.94
This study	Sigcomm2009	5 features: Direct Trust Value or PDR/ Packets forwarded of the current interaction (Pkts_f) / Packets dropped of the current interaction (Pkts_d) / Mutual Fiends (MF) / Common Interest Groups (CIG)	KNN, SVM, LG, MLP	SVM accuracy: 0.9978 KNN accuracy: 0.9673 LG accuracy: 0.9970 MLP accuracy: 0.9940

The comparative study highlight that most of the analyzed works handle the trust assessment issue as a classification problem. Then, this analysis conducted us choose the suitable ML model to elaborate our trust model.

2.2 Trust classifier selection

The ML models selected consist of SVM, KNN, LR, and MLP. We briefly describe each of them in Table 2:

Table 2. Description of the Machine Learning models

Machine Learning Model	Description
SVM	Involves the idea of a "margin" that separates two data classes [27].
KNN	Proposed by Cover and Hart [28], based on the principle that instances of a dataset usually exist near other instances with similar properties. Simple with high accuracy [19], [25].
LR	A binary classification method. Fast training speed [22], [29].
MLP	Based on the use of Artificial Neural Networks (ANN). Most used for numerical data [25], [30]. Composed of several perceptrons which are simple algorithm that performs binary classification [31].

3 Design of smart trust model

3.1 TMS life cycle

In this section, we present for the first time the fundamental components of TMS commonly known in the literature [32], [33], [34]. Thus, TMS is composed of five phases as follows: Gathering information, Trust calculation, Trust Decision, Trust up

Date and Reward and Punish. In this paper, we focused on the former three steps that are the basis of the proposed Trust model. Hence, we aim to develop the two last ones in future work especially the Reward and Punish phase related to the Peer Recommender Framework.

- ***Gathering information***: The TMS gathers information from all the nodes of the system. It comprises two functions explained subsequently:

  - Trust Composition: it consists of the extraction of trust parameters essential to trust value creation. These features can represent the Quality of Service (QoS) that an entity provides or represent the social behavior of an entity and its social relationships with other entities on the system (Social Trust) [32].
  - Trust Formation: it linked to building trust value on single or multiple parameters. The majority of TMS consider multiple parameters [32].

- ***Trust Calculation***: After gathering trust information, trust values are computed. it includes two major phases:

  - Trust Aggregation: its objective is to arrive at a final and an overall trust value that can be binary, (trustworthy/untrustworthy) or numerical to the ranking of the trustees. The most known technique is the weighted mean [16], [17]. Recently, to overcome the shortcoming presented by this latest mentioned, ML algorithms were applied [35].
  - Trust Propagation deals with how the trust information propagates through the network. They are two kinds: Centralized when there is a unique and central entity in charge of gathering, calculating, storing, and propagating trust information around the network. In a Decentralized scheme, information gathering and trust calculation are performed by all entities of the system.

- ***Trust Decision:*** this step permits the Trustor to decide to trust or not the trustee. They are two types of TMS:

  - Policy-based TMS: based on storing and sharing policies and credentials.
  - Reputation-based TMS is based on the trust evaluation process of a service provider by the service requester or other entities.

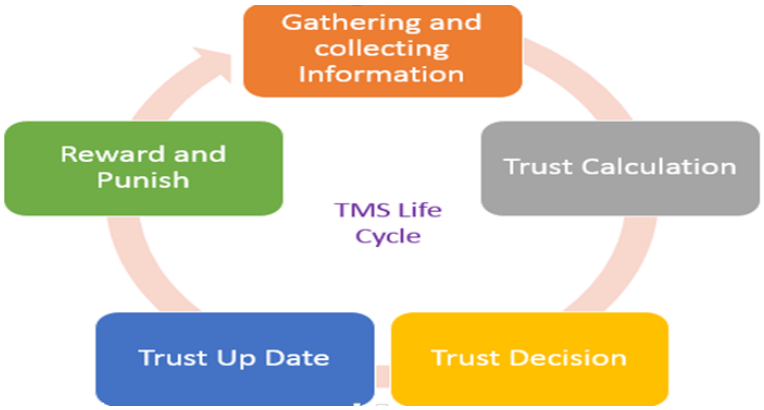

Fig. 4. Design components of TMS

3.2 Key phases of the proposed trust model

Our trust model steps are based on the TMS components already explained.

Step 1: Preprocessing the raw SIoT dataset. The preprocessing phase includes dealing with inappropriate values to convert data into a more suitable form for the selected ML algorithms. Nevertheless, finding an appropriate dataset is a challenging task. Thus, MOOCs contain personal data about learners, data related to a course, and data about learners' interaction with learning resources. This kind of data is insufficient to classify learners as trustworthy and untrustworthy. Therefore, there is a need for external information, especially, social information related to learners' social behavior such as their relationships, preferences, and interests to ensure a better assessment of trust. The learner is a human being, defined by different characteristics including his interests, preferences, and social context that represent an important factor in his choices and decision making [36]. Therefore, for these reasons, we used a raw SIoT dataset derived from MobiClique, a Mobile Social Network (MSN) used during SIGCOMM 2009 conference in Spain [37]. This MSN lets the availability of dynamic data. It helps users to explore and join various interest groups and to create new ones at any time. Likewise, it enables users to meet and find new friends. Hence, the list of interests and the friendship graph dynamically change. This dynamic data is desirable in the trust evaluation because trust is very dynamic.

The raw dataset is divided into several Comma-Separated Values (CSV) files. Table 3 exposes each file used in this search with a brief description of its contents. For more information on these files, visit the CRAWDAD platform (https://crawdad.org/thlab/sigcomm2009/20120715/index.html).

Table 3. Dataset files and content

CSV files	content
Participants	Includes a basic social profile: home city, country, and affiliation.
Friends 1	Contains a list of friends of the participants based on their Facebook friends
Friends 2	The new list of friends that the application enables participants to discover
Interests 1	List of initial interest groups of the participants based on their Facebook groups and networks
Interests 2	List of new interest groups created and joined by participants
Messages	List of messages stored, carried, and forwarded by participants during the experiment.
Transmission	Message transmission logs Data is transmitted between two devices using Bluetooth RFCOMM protocol.
Reception	List of messages receipted

Step 2: Dynamic Trust Features engineering. In general, the raw data is inoperable. Feature engineering is a crucial step since it impacts strongly Machine Learning's performance and consequently the decision-making process [38]. In a formal way, the problem is directed towards designing a set of features extracted and used to build a binary classification model **y** using a given training set such that it takes features X as an input and predicts the class label of a learner as an output. The label of each training sample i is denoted by $y\ (i)$: {untrustworthy, trustworthy}. In our context, a device is untrustworthy because its owner (learner) is untrustworthy. In the following, we present the eight trust features extracted from the dataset and we give the calculation formulas of the three calculated trust parameters. These trust attributes are inspired from the works of [39], [40] and described in Table 4:

Table 4. Description of the extracted Trust features

Trust feature	Description
Time instance	Time instance of an interaction between learner trustor and learner trustee
Learner Trustor	ID of an object (learner owner of devices) initializing an interaction
Learner Trustee	ID of an object (learner owner of devices) as a destination
Pkts_f	Packets forwarded of the current interaction among trustor and trustee
Pkts_d	Packets dropped of the current interaction among trustor and trustee
PDR / Direct Trust	Packet delivery ratio in the range [0,1]
Mutual friends	Similarity with learners with respect to the mutual friends among them in the range [0,1]
CIG	Similarity of nodes (learners) with respect to the social interest communities (the nodes share common interest groups) in the range [0,1]

We have used MATLAB R2018a to merge the different CSV files and to compute all trust features.

Packet delivery ratio. (PDR) or Direct Trust Value (DTV): It is related to the current direct trust observation. Also, it is linked to the ratio of the number of the packets successfully forwarded to the total number of the packets at any given time as:

$$PDR = \frac{Nbre\ of\ packets\ forwarded}{Nbre\ of\ packets\ Forwarded + Nbre\ of\ Packets\ Dropped} \quad (1)$$

In the literature, the PDR is considered the primary parameter for calculating direct trust to a trustee and a key criterion for designing trust models and for identifying malicious behaviors [6].

Mutuality or Mutual Friends (MF). if two users have mutual friends, these friends can close the trust gap between them. This feature is computed as the ratio of common friends between a Trustor and a trustee to the total number of friends between the two as:

$$MF = \left|\frac{F_i \cap Fj}{F_i \cup Fj}\right| \quad (2)$$

Where F_i and Fj represent the number of friends of a Trustor and a trustee respectively and $|.|$ shows the cardinality of a set which gives the count on the number of elements in the set.

Common Interest Groups (CIG). Two nodes with a degree of high community-interest, have more chances in interacting with each other, trust each other, and thus can result in better network performance. It represents the ratio of common interest groups to the total number of interest groups where both the Trustor and trustee are involved, and his calculation formula is as follow:

$$CIG = \frac{C_i \cap C_j}{C_i \cup C_j} \quad (3)$$

Where C_i depicts the communities of a trustor and C_j represents the count on communities of a trustee.

The dataset in Figure 5 shows a representative example of trust attributes and samples captured over the simulation scenario.

	Time	trustor	trustee	pkts_f	pkts_d	Direct Trust Value	MF	CIG
0	1	6	1	5	2	0.714286	0.00000	0.00000
1	2	6	14	6	1	0.857143	0.00000	0.75000
2	3	13	4	4	6	0.400000	0.00000	0.75000

Fig. 5. Representative samples of our dataset

We notice that in this paper, the small number of trust features is an advantage, which leads to a higher speed computation.

Data Labelling. In our model, we are supposed to perform labeling of the data to identify two different labels, namely those that are trustworthy and those that are not. Using k-means, the data set is simply divided into two clusters 1 and 0 arbitrarily. In our case, we have continuous trust values that are converted to binary values by comparing their value to the threshold that can be adapted to meet different requirements [29]. So, we used a conditional function and fixed a threshold which is used to decide when an entity is trustworthy or not). It is 0.5 for our study. Hence, a node is consid-

ered trustworthy if its PDR value is greater than 0.5 and the node has an MF value or CIG values greater than 0.5. If the PDR value is below 0.5 and MF or CIG values are below the threshold, the node is untrustworthy. Figure 6 shows the dataset after data labeling.

	Time	trustor	trustee	pkts_f	pkts_d	Direct Trust Value	MF	CIG	label
0	1	6	1	5	2	0.714286	0.00000	0.00000	0
1	2	6	14	6	1	0.857143	0.00000	0.75000	1
2	3	13	4	4	6	0.400000	0.00000	0.75000	0

Fig. 6. Samples of our dataset after data labeling

In our case, we have the sizes of the two classes that differ; from 5776 we have 4089 of class 1 and 1687 of class 0. Then, we chose a subset of 1678 nodes from these samples by using simple random sampling to shuffle the dataset for not having the same values consecutively for the label's feature to ensure a reliable ground truth. In addition, to avoid overfitting data and to obtain the maximum accuracy of the learning algorithm, 80% of samples were used for training purposes, whereas, 20% of them were used to evaluate the accuracy of the proposed model.

Classifiers Hyper parameter optimization. It is prominent that ML models cannot achieve the best performance without considering optimization techniques [29].

In the case of SVM, we tested the Linear Kernel (LK) and the Radial Basis Function Kernel (RFBK) [27]. The former shows high accuracy with 0.9978 and the RBKL gives 0.9940. Then, the choice of appropriate parameters is a crucial step for achieving reasonable results [41]. The settings of these parameters are based on a so-called "grid search" [27]. The goal is to identify two parameters: "C" and "gamma." that are conventionally used to avoid data overfitting [41]. For that, we have utilized part of the training samples as the cross-validation to find the best parameter set and the results obtained via the trained model are enhanced. For KNN, the appropriate value for K can be configured experimentally. Therefore, we can reach the optimal value of K by using 10-fold cross-validation on our dataset using a generated list of odd numbers ranging from (1-10). In our case, the optimal number of neighbors is K= 3 with accuracy of 0.9673. Concerning the MLP model, Table 5 reports the set of parameters used in MLP model and Figure 7 depicts the trust evaluation based on MLP:

Table 5. Summary of MLP configuration

Designation		**Value**
Input layer	Number of neurons	8
Hidden layer sizes	Number of hidden layers	(10, 10, 10) units
Output layer	Number of neurons	2
Optimizer		Adam
Activation function		Rectified Linear Units (ReLu)
Error calculation function		Back propagation by cross entropy

Designation	Value
Batch-size	Auto
Epoch	10

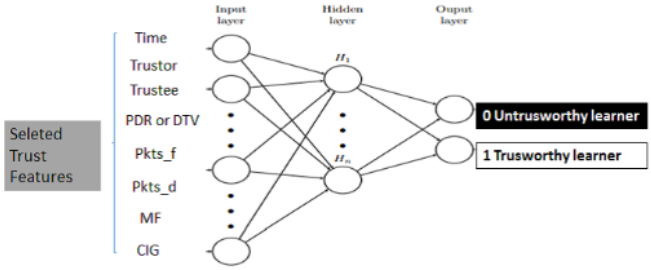

Fig. 7. Trust evaluation based on MLP model

Step 3: Training and classification using Machine Learning Classifiers. The four ML Models are tested, namely SVM, LR, KNN, and MLP to select the most appropriate.

Step 4: Performance Evaluation of the proposed model. In this step, we applied evaluation measures commonly used for trust prediction and classification issues that are reported and explained in detail in the subsequent section. Finally, a generic structure of the proposed MOOC trust model is shown in Figure 8.

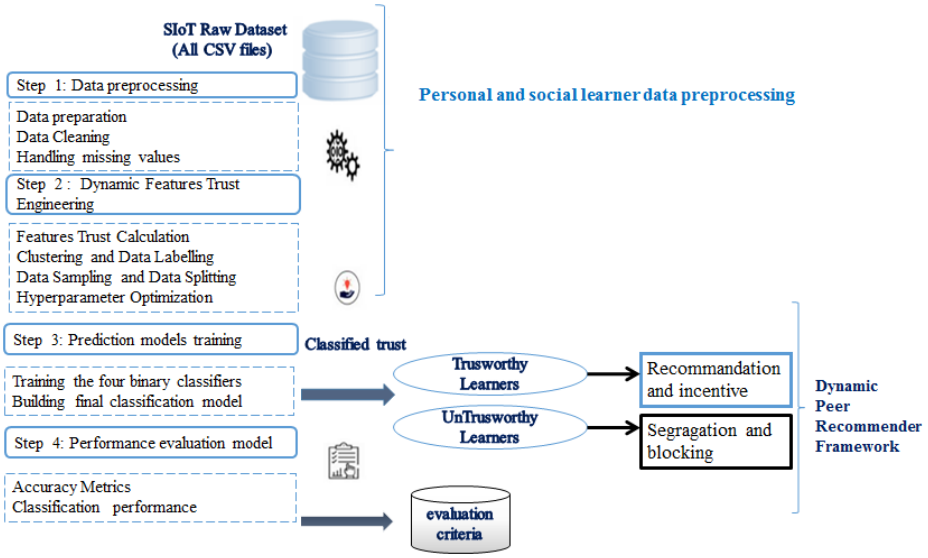

Fig. 8. Design of the smart MOOC TMS

The following section outlines the simulation environment and gives details related to experiment outcomes.

4 Material and methods

The subsequent section describes the experimental tools used. It gives information about the metrics used for evaluating the results. Finally, it provides a comparative analysis of the obtained results.

4.1 Experiment tools

The following experiments were all done under a personal computer which is configured as a win13 system, Intel(R) Core (TM) i5-3427U, 8Go RAM, 64-bit operating system. Concerning the Data preprocessing and the training ML models are performed in "Google Colab" which is a python notebook. It allows writing and running Python scripts in an internet browser with zero configurations required free access to GPUs, and easy sharing. The fact that “Google Colab” is based on Python makes the proposed model easy to integrate into MOOCs.

4.2 Performance analysis

The performance metrics used are: Accuracy, Recall, Precision, Receiver operating Characteristic (ROC) and, Area Under the Curve (AUC). Next, these metrics are explained in more details with their formulas as follows:

- Accuracy is the ratio of the number of correct predictions to the total number of input samples. The formula for calculating the Accuracy is:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Where:

TP = True Positive / TN = True Negative/ FP = False Positive / FN = False Negative

- Precision: This is the most well-known and general measure for evaluating the performance of classifiers. It reports the ratio of the correctly classified instances to all the instances. His formula is:

$$Precision = \frac{TP}{(TP+FP)} \quad (5)$$

- Recall or True Positive Rate (TPR): It provides important insight into classification performance relative to the number of incorrect predictions. It be calculated as follow:

$$Recall = \frac{TP}{(TP+FN)} \quad (6)$$

- Receiver Operator Character or ROC curve: it summarizes the prediction performance of a classification model by plotting the False Positive Rate (FPR) on the x-axis and True Positive Rate (TPR) or Recall on the y-axis. It can evaluate a classi-

fier's ability to predict both positive and negative classes. Thus, the ROC is a probability curve and the AUC represents the measure of class separability, it indicates how well the probabilities from the positive classes are separated from the negative.

We notice that both Precision and Recall can reflect the strength of the classifiers in predicting trust correctly since they are calculated based on the true positive/ negative and false positive/negative values, accordingly, as shown in Equations (5) and (6). Positive and negative represent trustworthy entities and untrustworthy entities respectively. Next, we present their interpretations in our real-life context that is a MOOC.

TP: It means the number of examples correctly classified as trustworthy. Correctly classifying learners can help to improve the quality of collaboration among them in the SLE because interactions will occur between reliable elements.

FP: it is the number of untrustworthy learners that are incorrectly labeled as trustworthy. Thus, co-learners that are supposed to be blocked for a learner are displayed and recommended for them. Undesirable actions of untrustworthy co-learners can result in the learner to drop out.

TN: it depicts the number of instances with distrust relationships that are correctly predicted as distrust. It is crucial for our context, to block untrustworthy learners and ensure MOOC network performance.

FN: it shows the number of trustworthy entities that are incorrectly classified as untrustworthy. This can result in poor quality services and prevent learners from collaborating with participants they truly trust.

4.3 Results comparison and discussion

This section outlines the comparison of the classification results obtained by the four classifiers. To demonstrate the effectiveness and the performance of the four classification methods, we generate a confusion matrix shown in Table 6:

Table 6. The confusion matrices of the four classifiers

	Untrustworthy	**Trustworthy**	**Untrustworthy**	**Trustworthy**
Untrustworthiness predictions	710	0	129	14
Trustworthiness prediction	1	1022	1	316
	711	1022	130	330
SVM (LK) accuracy: 0.9978			KNN (K=3) accuracy: 0.9673	
	Untrustworthy	Trustworthy	Untrustworthy	Trustworthy
Untrustworthiness predictions	91	0	89	1
Trustworthiness prediction	1	246	0	248
	92	246	89	249
LR accuracy: 0.9970			MLP accuracy: 0.9940	

Table 6 shows the proposed model can to obtain good accuracy scores for each classifier. Additionally, Figure 9 and Table 7 demonstrate that the SVM (LK) achieving the highest accuracy with 99, 78% proved to hold higher efficiency. LR and MLP were 99, 70% and 99, 40 % respectively. Finally, the accuracy of 3NN was 96, 73%, which presents the highest error rate compared to the other classifiers.

Table 7. Accuracy rate obtained with the four classifiers

Name of classifier	Accuracy rate	Error rate
SVM (LK)	0.9978	0.22
LR	0.9970	0.3
MLP	0.9940	0.6
3NN	0.9673	3.27

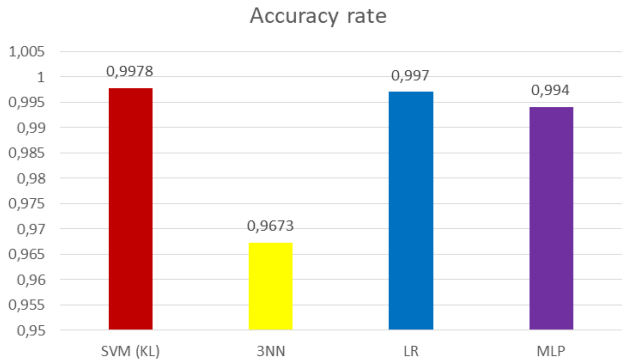

Fig. 9. The prediction accuracies rates obtained with the four classifiers

Subsequently, the obtained averaged Recall, Precision, and AUC for each model are reported in Figure10.

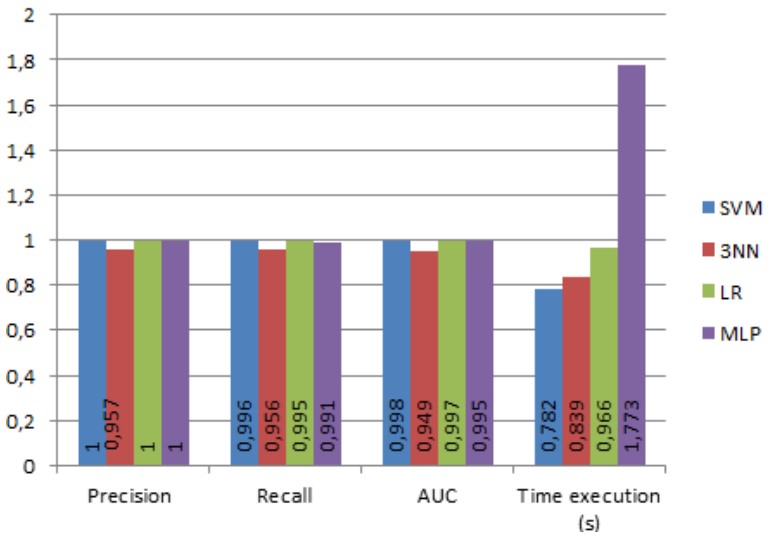

Fig. 10. Comparison of ML classifiers

Among the four tested methods, SVM (LK) obtained the best performance indicators with considerable values in the three evaluation criteria (Precision, Recall, and AUC) and a high-speed execution (0.782 seconds). Additionally, to have an evaluation metric for the correctness of the results, we draw the ROC diagram for the four classifiers, SVM, KNN, LR, and MLP. The corresponding diagrams are shown in Figure 11:

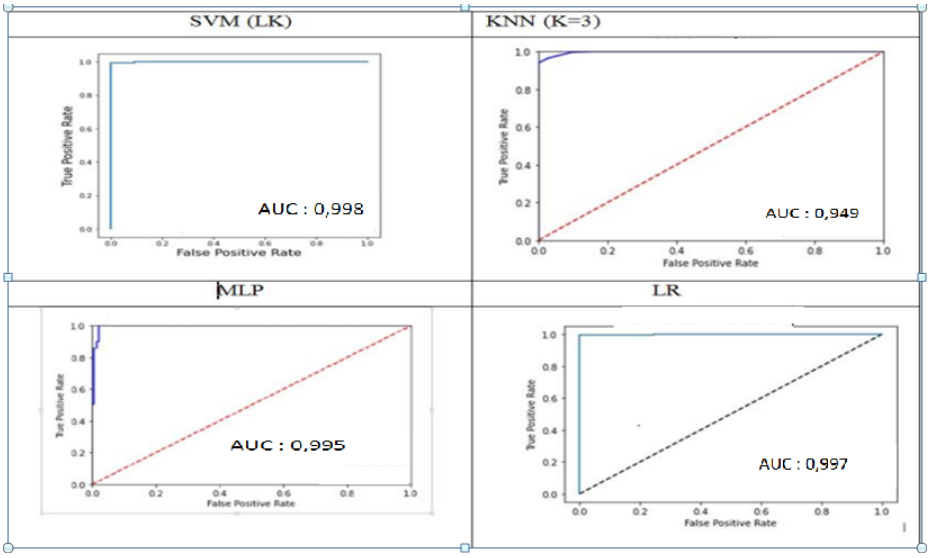

Fig. 11. Comparison of ROC Curve of the four classifiers models

Generally, all classifiers give good AUC values ranging from 0.95 to 0.99, demonstrating that the tested classifiers are better at distinguishing between untrustworthy and untrustworthy classes (0 and 1 respectively). However, SVM (LK) gives the highest values in the other evaluation metrics. An overview of the results shows that our proposed approach is achieving encouraging outcomes. We can observe that our selected trust features yield better results of a classification in all the evaluation metrics (Accuracy, Precision, Recall, and AUC) and surpasses 99%, which indicates the proposed trust features perform well to recognize the untrustworthy entities and identify the trustworthy ones as well. Moreover, two reasons lead to the improvement of our proposed trust evaluation method. Firstly, the trust feature extraction approach help to find out the optimal set of features for the trust evaluation process-based ML. In addition, the datasets which comprise dynamic and social data are essential for better trust evaluation performance. Therefore, learners in MOOCs use various tools outside the MOOC [42]. However, they prefer mobile devices like smartphones, tablets, etc., and the MSN apps have become an integral part of their lives [43]. It is suitable to incorporate in the MOOCs an adapted MSN that becomes a part of the MOOC design and which learners use inside the platform. This application will enable learners to collaborate, share content, create interest communities, and also make social and contextual data key of trust computation available and obtainable.

5 Conclusion and future works

In this paper, we designed an intelligent TMS based on ML techniques in MOOC ecosystems that can dynamically assess trust among learners allowing not only their classification but also the prediction of their future behaviors. Hence, we have conducted experiments on the real-world dataset from the MobiClique MSN that provides data from real mobile users and devices. Ultimately, our research can be expanded in two directions. First, we intend to build a dynamic Peer recommender Framework based on the proposed MOOC Trust model and to implement it in a MOOC platform to recommend trustworthy learning peers and to block untrustworthy ones. Second, we intend to propose a design of an MSN adapted to MOOCs platforms to boost social interaction and make available social data and trust information to guarantee an efficient trust evaluation. Our aim is to maximize the MOOC network performance, boost collaboration sustainability and ensure a better learning experience.

6 References

[1] C. Marche, L. Atzori and M. Nitti, "A Dataset for Performance Analysis of the Social Internet of Things," *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2018, pp. 1-5, 2018. https://doi.org/10.1109/PIMRC.2018.8580830

[2] M. Sharif, and A. Sadeghi-N, "Ubiquitous Sensor Network Simulation and Emulation Environments: A Survey". *Journal of Network and Computer Applications*. 2017. https://doi.org/10.1016/j.jnca.2017.05.009

[3] L. Atzori, A. Iera, G. Morabito, and M. Nitti," The social internet of things (SIoT) – when social networks meet the internet of things: concept, architecture and network characterization", *Computer Networks*, Vol. 56. 2012. https://doi.org/10.1016/j.comnet.2012.07.010

[4] M. Qingquan, J. Jiyou, and Z. Zhiyong, "A framework of smart pedagogy based on the facilitating of high order thinking skills". *Interactive Technology and Smart Education*, ahead-of-print (ahead-of print), 2020. https://doi.org/10.1108/ITSE-11-2019-0076

[5] L. Rai, and D. Chunrao, "Influencing factors of success and failure in MOOC and general analysis of learner behavior", *International Journal of Information and Education Technology*, Vol.6, 262, 2016. https://doi.org/10.7763/IJIET.2016.V6.697

[6] U. Jayasinghe, G. M. Lee, T-W. Um, and Q. Shi, "Machine Learning Based Trust Computational Model for IoT Services". *IEEE Transactions on Sustainable Computing*, Vol.4, pp.39–52, 2019. https://doi.org/10.1109/TSUSC.2018.2839623

[7] S., El Emrani, S., A., El Merzouqi and M., Khaldi, "An Intelligent Adaptive cMOOC "IACM" for Improving Learner's Engagement. *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, pp.82-94, 2021. https://doi.org/10.3991/ijet.v16i13.22261

[8] W. Abdelghani, C. Zayani, A. Amous, I., F. Sèdes ," Trust Evaluation Model for Attack Detection in Social Internet of Things". CRISIS 2018 - *13th International Conference on Risks and Security of Internet and Systems*, Arcachon, France. pp. 48-64, 2018. https://doi.org/10.1007/978-3-030-12143-3_5

[9] W. Abdelghani, C. Zayani, I. Amous, and Sedes, F, "Trust Management in Social Internet of Things: A Survey". 2016, 9844. 430-441. https://doi.org/10.1007/978-3-319-45234-0_39

[10] J. Golbeck "Computing and applying trust in Web–based social networks". Ph.D. dissertation, University of Maryland, College Park, 2005.

[11] N.B.Truong, T.W. Um, G.M. Lee and G. Lee, "Toward a Trust Evaluation Mechanism in the Social Internet of Things. Sensors". 2017. https://doi.org/10.3390/s17061346

[12] E., Costello, J., Brunton, M., Brown and L., Daly, "In MOOCs we Trust: Learner Perceptions of MOOC Quality via Trust and Credibility. *International Journal of Emerging Technologies in Learning (iJET)*, vol. 13, pp. 214-222, 2018. https://doi.org/10.3991/ijet.v13i06.8447

[13] A. Ahmad, R. Jabeur N., and M. K. Ijaz, "Machine Learning Prediction and Recommendation Framework to Support Introductory Programming Course". *International Journal of Emerging Technologies in Learning (iJET)*, Vol.16, no. 17, pp. 42-59, 2021. https://doi.org/10.3991/ijet.v16i17.18995

[14] K. Elghomary, D. Bouzidi, and N. Daoudi, "Study and review of OSN and SIoT trust models: towards a dynamic MOOC trust model", *Int. J. Multimedia Intelligence and Security*, Vol. 3, no. 4, pp.407–431, 2020. https://doi.org/10.1504/IJMIS.2020.114796

[15] K. Elghomary, D. Bouzidi and N. Daoudi, "A Comparative Analysis of OSN and SIoT Trust Models for a trust model adapted to MOOCs platforms", In Proc. 2nd International Conference on Networking, Information Systems & Security (NISS19). Association for Computing Machinery, New York, NY, USA, Article 9, 2019, pp.1–8. https://doi.org/10.1145/3320326.3320335

[16] K. Elghomary and D. Bouzidi, "Dynamic Peer Recommendation System based on Trust Model for sustainable social tutoring in MOOCs". In Proc. 1st International Conference on Smart Systems and Data Science (ICSSD), pp. 1-9, 2019. https://doi.org/10.1109/ICSSD47982.2019.9003154

[17] M. Nitti, R. Girau and L. Atzori, "Trustworthiness management in the social internet of things", *IEEE Transactions on Knowledge and Data Management*, Vol. 26, pp.1–11. 2014. https://doi.org/10.1109/TKDE.2013.105

[18] R. Chen, F. Bao and J. Guo, "Trust-based service management for social internet of things systems", IEEE *Trans Dependable SystComput*, Vol. 13, pp.684–96, 2015. https://doi.org/10.1109/TDSC.2015.2420552

[19] E. Khadangi and A. Bagheri, "Comparing MLP, SVM and KNN for predicting trust between users in Facebook", In Proc. ICCKE 2013, 2013, pp.466–470. https://doi.org/10.1109/ICCKE.2013.6682864

[20] K. Zhao and L. Pan, "A machine learning based trust evaluation framework for online social networks". In Proc. IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications, Beijing, 2014, pp.24-26. https://doi.org/10.1109/TrustCom.2014.13

[21] T. Yelena, M. Alexandru and T. Pavel, "Application of neural networks for decision making and evaluation of trust in ad-hoc networks", In Proc. IEEE 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC) – Valencia, Spain, 2017, pp.371–377. https://doi.org/10.1109/IWCMC.2017.7986315

[22] W. Yuji, "The trust value calculating for social network based on machine learning", In Proc. 9thInternational Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Vol. 2, 2017, pp.133-136. https://doi.org/10.1109/IHMSC.2017.145

[23] X. Chen, Y. Yuan, L. Lu and J. Yang, "A multidimensional trust evaluation framework for online social networks based on machine learning". *IEEE Access*, pp.175499–175513, 2019. https://doi.org/10.1109/ACCESS.2019.2957779

[24] G. Han, Y. He, J. Jiang, N. Wang, M. Guizani,and J. Adu Ansere, "A synergetic trust model based on SVM in underwater acoustic sensor networks"., *IEEE Transactions on Vehicular Technology*, Vol. 68, pp.11239–11247, 2019. https://doi.org/10.1109/TVT.2019.2939179

[25] S. A. Siddiqui, M.Adnan, Z. Wei Emma, and S. Quan, "Machine Learning Based Trust Model for Misbehaviour Detection in Internet-of-Vehicles", in Proc. International Conference on Neural Information Processing, ICONIP, Neural Information Processing, 2019,pp. 512-520. https://doi.org/10.1007/978-3-030-36808-1_56

[26] M. Masmoudi, W. Abdelghani, I. Amous and F. Sèdes, "Deep Learning for Trust-Related Attacks Detection in Social Internet of Things", *Lecture Notes on Data Engineering and Communications Technologies*, pp.389–404. 2020. https://doi.org/10.1007/978-3-030-34986-8_28

[27] C. Hsu, C. Chang and C. Lin, "A Practical Guide to Support Vector Classification" 2008.

[28] T. Cover and P. Hart,"Nearest neighbor pattern classification", *IEEE Trans*. Inf. Theory, Vol.13, pp.21-27, 1967. https://doi.org/10.1109/TIT.1967.1053964

[29] L. Shushu and Z. Lifang, "Predict Pairwise Trust Based on Machine Learning in Online Social Networks: A Survey", *IEEE Access*. 1-1. 2018. https://doi.org/10.1109/ACCESS.2018.2869699

[30] Z. Lin, L. Dong, "Clarifying trust in social internet of things. IEEE Trans". *Knowl. Data Eng*. Vol.30, pp.234–248, 2018. https://doi.org/10.1109/TKDE.2017.2762678

[31] I. Bouteraa, M. Derdour, A. Ahmim, "Intrusion detection using classification techniques: a comparative study". *International Journal of Data Mining, Modelling and Management*, Vol.12, pp.65-86. https://doi.org/10.1504/IJDMMM.2020.105596

[32] J. Guo, & I. Chen, "A Classification of Trust Computation Models for Service-Oriented Internet of Things Systems". *2015 IEEE International Conference on Services Computing*, pp.324-331. 2015. https://doi.org/10.1109/SCC.2015.52

[33] Chahal, R.K., Kumar, N., & Batra, S. (2020). "Trust management in social Internet of Things: A taxonomy, open issues, and challenges". *Comput. Commun., 150*, pp.13-46,2020. https://doi.org/10.1016/j.comcom.2019.10.034

[34] F. M. Gomez, and P. G. Martinez, "Towards pre-standardization of trust and reputation models for distributed and heterogeneous systems". *Computer Standards & Interfaces*. 32. pp.185-196. 2010. https://doi.org/10.1016/j.csi.2010.01.003

[35] J. Wang, X. Jing, Z. Yan, Y. Fu, W. Pedrycz and L. T. Yang," A Survey on Trust Evaluation Based on Machine Learning", *ACM Computing Surveys*, Vol. 53, 107. 2020. https://doi.org/10.1145/3408292

[36] W. Abdelghani, C., A. Zayani, I. Amous, and F. Sedes, "User-centric IoT: Challenges and Perspectives". 2018.

[37] A. Pietiläinen, C. Diot, CRAWDAD data set thlab/sigcomm2009, 2012, (Downloaded from https://crawdad.org/thlab/sigcomm2009/20120715/index.html).

[38] M. Youssef, S. Mohammed, E.K. Hamada, and B.F. Wafaa. "A predictive approach based on efficient feature selection and learning algorithms competition: Case of learners' dropout in MOOCs". *Education and Information Technologies*, Vol. 24, pp.3591–3618 .2019. https://doi.org/10.1007/s10639-019-09934-y

[39] S. Subhash, M. Adnan, S. Quan and S. A. Siddiqui, "SCaRT-SIoT: Towards a Scalable and Robust Trust Platform for Social Internet of Things", 2020. https://doi.org/10.1145/3384419.3430434

[40] S. Subhash, M. Adnan, S. Quan, Z. Munazza and Z. Wei Emma, "Towards a Machine Learning-driven Trust Evaluation Model for Social Internet of Things: A Time-aware Approach", *arXiv - CS - Social and Information Networks*, 2021. *arxiv-2102.10998*

[41] T. Eitrich, B. Lang, "Efficient optimization of support vector machine learning parameters for unbalanced datasets", *Journal of Computational and Applied Mathematics*,Vol. 196, pp.425-436. 2006. https://doi.org/10.1016/j.cam.2005.09.009

[42] M. Harju, T. Leppänen, & I. Virtanen, "Interaction and Student Dropout in Massive Open Online Courses". 2018. *ArXiv, abs/1810.08043*

[43] J-É. Pelet, M. Pratt and S. Fauvy, "MOOCs and the Integration of Social Media and Curation Tools in e-Learning". In Proc. International Workshop on Learning Technology for Education in Cloudpp conference.43-53, 2015. https://doi.org/10.1007/978-3-319-22629-3_4

7 Authors

Khadija Elghomary is a currently a Ph.D. Student from the National Superior School of Computer Science and Systems Analysis (ENSIAS), received her master's degree in information sciences in Information Sciences School (ESI). She began her Ph.D. studies since 2016 and is working on many research areas related to Technology-Enhanced Learning such as MOOC platforms, tutoring in Virtual and Smart Learning Communities, Machine Learning, Recommender Systems, Social Networks Analysis and Trust Models development.

Driss Bouzidi is an associate professor in Computer Sciences at ENSIAS, University Mohammed V, Rabat, Morocco. His research interests are mainly in the areas of distributed systems and security services. He has made many contributions to several chapters in some international books related to e-learning. He was vice-chair of the international conference NGNS'09, TCP chair of NGNS10, NGNS12, and ICEER13

and chair of JDSIRT'2017. He is a founding member of two research associations APRIMT and e-NGN (email: driss.bouzidi@um5.ac.ma).

Najima Daoudi is a Professor at the School of Information Sciences, Rabat, Morocco. She is an Engineer of the National Institute of Statistics and Applied Economics and has a Ph.D. in Computer Science from ENSIAS. She has produced several articles in E-learning, M-learning and Ontology development since 2005. She was chair of the international conference ICSSD'19 (email: ndaoudi@esi.ac.ma).

Article submitted 2021-10-20. Resubmitted 2021-12-15. Final acceptance 2021-12-18. Final version published as submitted by the authors.