A LEARNING FUZZY COGNITIVE MAP (LFCM) APPROACH TO PREDICT STUDENT PERFORMANCE

Taha Mansouri  
University of Salford, Salford, UK  
T.Mansouri@salford.ac.uk

Ahad ZareRavasan*  
Masaryk University, Brno, Czech Republic  
Ahad.ZareRavasan@econ.muni.cz

Amir Ashrafi  
University of Manchester, Manchester, UK  
Amir.Ashrafi@postgrad.manchester.ac.uk

* Corresponding author

ABSTRACT

Aim/Purpose This research aims to present a brand-new approach for student performance prediction using the Learning Fuzzy Cognitive Map (LFCM) approach.

Background Predicting student academic performance has long been an important research topic in many academic disciplines. Different mathematical models have been employed to predict student performance. Although the available sets of common prediction approaches, such as Artificial Neural Networks (ANN) and regression, work well with large datasets, they face challenges dealing with small sample sizes, limiting their practical applications in real practices.

Methodology Six distinct categories of performance antecedents are adopted here as course characteristics, LMS characteristics, student characteristics, student engagement, student support, and institutional factors, along with measurement items within each category. Furthermore, we assessed the student's overall performance using three items of student satisfaction score, knowledge construction level, and student GPA. We have collected longitudinal data from 30 postgraduates in four subsequent semesters and analyzed data using the Learning Fuzzy Cognitive Map (LFCM) technique.

Contribution This research proposes a brand new approach, Learning Fuzzy Cognitive Map (LFCM), to predict student performance. Using this approach, we identified the most influential determinants of student performance, such as student engagement. Besides, this research depicts a model of interrelations among the student performance determinants.
Findings

The results suggest that the model reasonably predicts the incoming sequence when there is a limited sample size. The results also reveal that students’ total online time and the regularity of learning interval in LMS have the largest effect on overall performance. The student engagement category also has the highest direct effect on student’s overall performance.

Recommendations for Practitioners

Academic institutions can use the results and approach developed in this paper to identify students’ performance antecedents, predict the performance, and establish action plans to resolve the shortcomings in the long term. Instructors can adjust their learning methods based on the feedback from students in the short run on the operational level.

Recommendations for Researchers

Researchers can use the proposed approach in this research to deal with the problems in other domains, such as using LMS for organizational/institutional education. Besides, they can focus on specific dimensions of the proposed model, such as exploring ways to boost student engagement in the learning process.

Impact on Society

Our results revealed that students are at the center of the learning process. The degree to which they are dedicated to learning is the most crucial determinant of the learning outcome. Therefore, learners should consider this finding in order the gain value from the learning process.

Future Research

As a potential for future works, the proposed approach could be used in other contexts to test its applicability. Future studies could also improve the performance level of the proposed LFMC model by tuning the model’s elements.

Keywords

e-learning, Learning Analytics (LA), Learning Fuzzy Cognitive Map (LFMC), Learning Management System (LMS), Student Engagement, Student Performance

INTRODUCTION

The ever-increasing use of data and technology in the education environment has opened new research areas, including Educational Data Mining (EDM), Artificial Intelligence (AI), and Learning Analytics (LA) to effectively manage and use data throughout learning processes (Knight et al., 2020). E-learning refers to IT-based learning and has had a significant impact on educational institutions’ quality of learning outcomes (ZareRavasan & Ashrafi, 2019). Numerous institutions are now adopting different forms of e-learning to achieve their learning objectives (Sharma et al., 2017; ZareRavasan et al., 2018). A study by Open University found that producing and providing e-learning courses consumes 90% less energy and has 85% fewer CO₂ emissions per person than face-to-face training (Roy et al., 2004). The global eLearning market is expected to reach a total market value of $107 billion in 2015 to $325 billion by 2025, showing that it will have nearly tripled in a decade (Pappas, 2019). As a result, a huge amount of data related to educational processes are generated in every institution. Such data contain useful information that may be analyzed, simulated, and used in decision-making systems. Thus, a new research stream within the data analytics domain emerged called Learning Analysis (LA) (Papamitsiou & Economides, 2014). LA is the calculation, conservation, analysis, and reporting of learning subjects and their context to understand and optimize the very essence of learning and the conditions in which it occurs (Lang et al., 2017). In particular, LA is applying statistical analysis methods over the collected data to get useful insights.

One of LA’s main applications is predicting a student (learner) performance or success (Mwalumbwe & Mtebe, 2017; Salinas et al., 2019). Predicting student academic performance is vital because the instructor can take proactive measures (such as one-on-one tutoring and provide remedial lessons) to
improve student learning, especially for low-performed students (Huang & Fang, 2013). The prediction results might help students better understand how they would perform, pushing them to adjust their learning style. For prediction purposes, sophisticated mathematical techniques such as Artificial Neural Networks (ANN), decision trees, and regression are generally applied to a large dataset containing information about previous students on the course. This information could include: (a) things that are known at the start of the course (e.g., the students’ previous educational experience and demographic data such as age or gender); (b) things that become known during the course (e.g., login frequency, and post views); and (c) assessment data (e.g., final exam, and quizzes) (Clow, 2013). A model can be developed using this data and then applied to the information available for potential current students to predict how any specific student will perform for the course.

Indeed, prediction models are doing almost the same function as a traditional teacher, observing which students face difficulties in the class and giving them extra support, but in a different e-learning context. However, there are some practical variances. First, the prediction output is a sort of quantitative probabilities rather than the subjective and qualitative understanding of a teacher. Second, the quantitative output could be made available to all relevant stakeholders, not just the teacher. Third, supported with some intelligent agents, the output can be used to trigger proactive supporting actions even without engaging a teacher or a teaching assistant at all (Clow, 2013).

However, the available sets of common prediction approaches, such as ANN and regression, work well with large datasets, face challenges dealing with small sample sizes, limiting their practical applications in real practices (Yoon et al., 2019). Hence, to address this research and practice concern, this research aims to present a brand new approach for student performance prediction using the Learning Fuzzy Cognitive Map (LFCM) approach. FCM is a nonlinear model because a nonlinear transformation function (usually the sigmoid function) transforms the cumulative impact of causal concepts on the effect concept. Therefore, it can capture most of the nonlinearity in complex systems (Salmeron et al., 2019). An advantage of the FCM model over ANN is that FCMs can be interpreted easily by humans, and each FCM node and arc has a specific meaning known to the expert (Salmeron & Froelich, 2016). FCMs can be built over either experts’ knowledge extraction or a learning process. In this work, since we had several sequential data, we chose the Learning FCM (LFCM) proposed in Salmeron et al. (2019).

To build our LFCM model, we use a longitudinal quantitative and qualitative dataset of 30 masters students in an online program for three consecutive semesters, trying to predict every individual student’s performance for the fourth semester. Employing the LFCM approach, we also map the causal map of effects among observed variables and their impact on student performance.

The rest of the paper is organized as follows. We first present a literature review of the learning analytics and student performance prediction and then FCM applications in an e-learning context. The next section presents the methodology. Next, the results are presented and, finally, we present a discussion and conclude the paper.

**LITERATURE REVIEW**

**LEARNING ANALYTICS AND STUDENT PERFORMANCE PREDICTION**

Learning analytics (LA) is defined as “the measurement, collection, analysis, and reporting of data about learners and their contexts, for understanding and optimizing learning and the environments in which it occurs” (Siemens & Long, 2011, p. 34). Regarding the literature, LA is overlapped to some degree with academic analytics and EDM. Academic analytics means the use of business intelligence in the education context, focusing on the institutional and national level, rather than on individual students and courses (Siemens & Long, 2011). The EDM seeks to develop methods for analyzing educational data focusing on the technical challenges rather than on the pedagogical aspects. According to Clow (2013), LA is the first and foremost concern in learning. This can be applied for different
purposes such as student behavior modeling, prediction of performance, prediction of dropout and retention, and resources recommendation (Papamitsiou & Economides, 2014). The task of predicting student performance involves approximating students’ future status given a record of the past sequence of behaviors exhibited or activities engaged by a student (Raga & Raga, 2019). A brief overview of the main related previous research on student performance prediction is presented in the following.

One of the main research streams predicts student performance (and clustering them in some cases) based on different quantitative and qualitative data sets. Castellano and Martinez (2008) use collaborative filtering techniques to exploit students’ grades to generate group profiles that could facilitate academic orientation. Kumar and Uma (2009) use the classification process to examine various attributes affecting student performance. Lykourentzou et al. (2009) develop multiple feed-forward ANN to predict students’ final achievement dynamically and to cluster them in two virtual groups, according to their performance. Moridis and Economides (2009) demonstrate how various pieces of evidence could be combined to optimize inferences about affective states during an online self-assessment test. A method has been developed to predict students’ mood and was tested using data emulated from experiments made with 153 high school students. Thai-Nghe et al. (2009) propose improving student performance prediction by dealing with the class imbalance problem, using support vector machines. Yu et al. (2010) use linear support vector machines together with feature engineering and ensembling techniques for predicting student performance. These methods work well in case we have enough meta-data about students and tasks. Abdous et al. (2012) used ordinal logistic regression analysis to predict relationships between online question theme and final grade. Pardos et al. (2013) build a unified model that predicts student standardized examination scores from a combination of student effect, disengaged behavior, and performance within the learning system. More recently, Meghji et al. (2018), using Association Rule Mining and Pattern Discovery in the context of a higher educational institute, observe that student class performance is directly influenced by the attention given to a lecture, proper note-taking and the tendency to self-solve assignments.

Considering the literature, scholars also tried to develop tools to support learners and facilitate the learning process. For instance, Romero-Zaldivar et al. (2012) develop some virtual appliances to monitor students’ behavior and progress and then correlate them to the final grade. In this line, researchers have recently proposed using recommender system techniques (e.g., matrix factorization) to predict student performance (Thai-Nghe et al., 2012; Toscher & Jahrer, 2010). Extended from these works, Thai-Nghe et al. (2011) propose tensor factorization models to consider the sequential effect (for modeling how student knowledge changes over time). Moreover, Thai-Nghe et al. (2012) extend their work by introducing a new method of “tensor factorization forecasting” for predicting student performance. Macfadyen and Dawson (2010) analyze the Learning Management System (LMS) tracking data from a course that identified 15 variables demonstrating a significant correlation with students’ final grades. They also develop an early warning system for educators. Along the same lines, Vialardi et al. (2011) used data mining techniques that employed the students’ academic performance records to design a recommender system in support of the enrolment process. Raga and Raga (2019) develop a prediction model for student performance in the early stages of blended learning courses using deep neural network architecture and utilizing online activity attributes as input patterns.

Researchers recently focused mainly on a comparative analysis of the accuracy of different machine learning methods in student performance prediction applications. For instance, Huang and Fang (2013) develop and compare four types of mathematical models (i.e., multiple linear regression model, the multilayer perception network model, the radial basis function network model, and the support vector machine model) to predict student academic performance. Abdulwahhab and Abdulwahhab (2017) also used the Prediction Tree (CPT+) algorithm to predict the next grade for upcoming courses or the registered course(s) and find that it outperforms other prediction models. Umer et al. (2018) conduct a comparative analysis of four predictive models; Random Forest, Naïve Bayes, K-Nearest Neighbor, and Linear discriminant analysis, to predict student performances. Al Breiki et al. 
Mansouri, ZareRavasan, & Ashrafi (2019) compared the accuracy of different machine learning methods (i.e., SMOReg, Random Forest, Linear Regression, Multilayer Perceptron, K-Nearest Neighbor, Gaussian, Processes Random Tree, Decision Table, and Simple Log Regres) in predicting student performance at the United Arab Emirates University (UAEU). Sheshadri et al. (2018) use Logistic Regression, Decision Tree, and K-Nearest Neighbor classifiers to predict student performance using data generated before the test and over the course during the semester. Using the model, they classify students into distinction/non-distinction groups. Turabieh (2019) applies a hybrid feature selection algorithm with different machine learning classifiers (i.e., K-Nearest Neighbor, Convolutional Neural Network (CNN), Naive Bayes, and Decision Tree) to predict the student’s performance. They also apply a binary genetic algorithm as a wrapper feature selection. In addition, Mohammadi et al. (2019) apply three supervised learning algorithms (K-Nearest Neighbor, Naive Bayes, and Decision Tree) for student performance prediction. Masood et al. (2019) implemented 11 machine learning models to predict student performance using two public student databases. According to the results, Decision Tree and Random Forest are the machine learning models with the highest accuracy.

This short review shows that prior research applied different variables and techniques to model and predict student learning performance. Linear models such as linear and logistic regression were the mainstream applications in early research because of their simplicity and ability to account for certain linear patterns between available input variables and student achievement. More advanced methods, such as ANN, CNN, Naive Bayes, and Decision Tree, have been applied later. Due to different levels of complexity of these methods, recent studies attempted to compare different models’ accuracy in predicting student performance. However, as most of the adopted models contain many parameters to adjust and they pass data through a number of transforming layers, they only work well with large datasets and have serious shortcomings with small sample sizes (Yoon et al., 2019).

**FCM in the E-Learning Context**

Nowadays, FCMs play an important role in different scientific fields, including e-learning and the LMS context. FCM has been applied in this field to address different research aims. For instance, some research adapted FCM to model interrelations among factors affecting LMS or learning. In this line, Salmeron (2009) employs an Augmented FCM for modeling Critical Success Factors (CSFs) in LMS. Nownaisin et al. (2012) identify the causalities of the education management of the Thailand science-based technology school. Yesil et al. (2013) model the control engineering educational CSFs. Takacs et al. (2014) introduce a novel FCM algorithm to calculate the values of the interrelation levels between the factors in a system for the students’ grade evaluation.

Others used FCM to model the design, adoption, or assessment of LMS. For instance, Tsadiras and Stamatis (2008) examine the use of FCMs for planning network learning and enhancing the success of learning programs. Hostein and Brooks (2008) model educational software adoption at three UK secondary schools based on stakeholders’ perceptions. Georgiou and Botsios (2008) identify learning styles in adaptive educational hypermedia systems. Baron et al. (2015) present a learning assessment system that uses multivariate analysis based on Structural Equation Modeling (SEM), facilitating the assessment of learning in interactive environments.


Other research streams in the application of FCM in the learning discipline applied FCM in the decision-making process. For instance, Georgopoulos et al. (2014) expand the utilization of FCM based on medical decision support systems for learning and educational purposes using a scenario-based learning approach. Baykasoglu and Golcuk (2015) propose a new model that integrates Fuzzy TOPSIS and FCMs to model and solve decision problems related to industrial engineering. Pandey and Singh (2015) propose an FCM based recommender system for e-learning system with a multi-agent framework. Aguilar (2016) proposes a recommender learning resource in a smart classroom using FCM.

This review shows that FCM has the potential to address different research and practice problems in the learning discipline. Despite this, no attempt has been made to leverage the FCM’s potentials to predict student performance.

**RESEARCH METHOD**

The purpose of this study is to present a model to predict student performance using the LFCM approach. To achieve this goal, we utilized a research structure based on a literature review to identify factors as antecedents of student performance, longitudinal data, and LFCM as the data analysis procedure. The main research steps are illustrated in Figure 1.

![Figure 1. The research steps](image)

**LEARNING FCM**

FCMs are a combination of ANNs and Fuzzy Logic. They are fuzzy weighted digraphs, including several nodes and arcs (Salmeron & Froelich, 2016). FCMs define relationships using fuzzy terms such as very low or high while describing complex systems, including interrelated concepts (Salmeron et al., 2019). The nodes or concepts (\(C_i\)) and related weighted arcs (\(W_{ij}\)) display the modeled problem and causal dependencies between them, respectively (Salmeron & Froelich, 2016). Therefore, each \(W_{ij}\) sign shows the direction of causality between concepts \(C_i\) and \(C_j\), whereas its amount shows the intensity of that effect. The values of concepts are in the range of \([0,+1]\), and the weights are in the range of \([-1,+1]\).
range of \([-1, +1]\). Moreover, a combination of concepts summarizes a snapshot of the modeled system at any time, as a state vector \(A = \{A_1, \ldots, A_N\}\), in which \(N\) is the number of concepts, and \(A_i\) is the value of \(C_i\). In order to update \(A_i\) at each time, the following rule is proposed (Salmeron et al., 2019):

\[
A_i^t = f(A_i^{t-1} + \sum_{j \neq i}^{N} A_j^{t-1} W_{ji})
\]

Eq(1)

where \(A_i^t\) is the value of concept \(C_i\) at time \(t\), \(A_i^{t-1}\) the value of concept \(C_i\) at time \(t - 1\), and \(W_{ji}\) the connection from \(C_j\) to \(C_i\). The function \(f\) is an activation function to map and bound the result in an interval; and for this work, we use the unipolar sigmoid as proposed in (Bueno & Salmeron, 2009) as follows:

\[
f(x) = \frac{1}{1+e^{-ax}}
\]

Eq(2)

where \(x\) is described as the value of \(A_i\) and \(a\) is a real positive number modeling the function’s slope.

There are two types of analyses in FCMs as static and dynamic. Static analysis displays the concepts’ causal effects using the maximum amount among different paths from input concepts ending to target concepts (ZareRavasan & Mansouri, 2014). In this analysis, at first, a causal path from a concept node \(C_i\) to another concept node \(C_j\), through other nodes, say \(C_i \sim C_{k1}, C_{k1}\sim \ldots \sim C_{kn}, C_{kn}\sim C_j\) can be indicated by sequence \((i, k1, \ldots, kn, j)\). Then the indirect effect of \(C_i\) on \(C_j\) is the causality \(C_i \sim I\) impart to \(C_j\) via the path \((i, k1, \ldots, kn, j)\). The total effect of \(C_i\) on \(C_j\) is the composite of all indirect effect causalities \(C_i \sim I\) imparts to \(C_j\).

A simple fuzzy causal algebra is created by interpreting the indirect effect of the operator \(I\) as the minimum operator (or \(t\)-norm) and the total effect of operator \(T\) as the maximum operator (or \(s\)-norm) on the partially ordered set \(P\) of causal values (so we simply call it as the maximum of minimums). Formally let \(\sim\) be a causal concept space and let \(e: \sim \times \sim \rightarrow P\) be a fuzzy causal edge function, and assume that there are \(m\)-many causal paths from \(C_i\) to \(C_j\): \((i, k\sim \ldots \sim k, j)\) for \(1 \sim< r \sim< m\). Then let \(Ir(C_i, C_j)\) denote the indirect effect of concept \(C_i\) on concept \(C_j\) via the \(r\)th causal path, and let \(T(i, C_j)\) denote the total effect of \(C_i\) on \(C_j\) over all \(m\) causal paths. These operators are shown in Eq(3) and Eq(4):

\[
I_{\sim}(C_i, C_j) = \min\{w(C_p, C_{p+1})): (p, p+1) \sim (i, k\sim \ldots \sim k, j)\}
\]

Eq(3)

\[
T(C_i, C_j) = \max(Ir(C_i, C_j)), \text{ where } 1 \sim< r \sim< m
\]

Eq(4)

where \(p\) and \(p+1\) are contiguous left to right path indices (Kosko, 1986).

The dynamic analysis starts with an initial state vector such as \(A = \{A_1, \ldots, A_N\}\); and keep updating with Eq (1) and Eq (2) until reaching an equilibrium point in which the amounts of the resulted vector have plateau (Hobbs et al., 2002).

FCMs can be constructed either by experts or historical data. However, there are some limitations in relying on experts’ subjective opinions (Kahvandi et al., 2018; ZareRavasan & Mansouri, 2016). The learning approaches for FCMs are concentrated on learning the weighted matrix \(W\), based either on expert intervention or available historical data (Papageorgioua & Kannappan, 2012). There are various learning approaches to train the FCM’s weight matrix, mostly from the ANN discipline (Salmeron et al., 2019). In this paper, we use Modified Asexual Reproduction Optimization (MARO) proposed by Salmeron et al. (2019).

The main objective function of FCM-MARO is minimizing the following error:

\[
in - sample - error = \frac{1}{(K-1)N} \sum_{t=1}^{K} \sum_{n=1}^{N} \left| C_n(t) - \hat{C}_n(t) \right|
\]

Eq(5)

where \(C_n(t)\) is the original value of the concept \(n\) at time \(t\), \(\hat{C}_n(t)\) is the calculated value of that concept using the candidate FCM and Eq (1) and Eq (2), \(K\) is the number of input observations and \(N\) is
the number of concepts. The following is the pseudocode of FCM-MARO algorithm (Salmeron et al., 2019).

\[
\begin{align*}
t &= 1 \quad // \text{Initial time setting} \\
loc &= 1 \quad // \text{Initial local number setting} \\
P &= \text{initialize}(L,U) \quad // \text{Creating a randomly generated weight matrix as parent, bounded between lower and upper thresholds} \\
\text{errorP} &= \text{Cost}(P) \quad // \text{Calculated regarding Eq}(5) \\
\textbf{While} \text{ stopping_conditions} \neq \text{true} \textbf{do} \\
& \quad \text{bud}_t = \text{reproduce}(P) \quad // \text{Generating a randomly modified offspring from the parent (P)} \\
& \quad \text{error}_t = \text{Cost} (\text{bud}_t) \\
& \quad \textbf{if} \text{ error}_t < \text{errorP} \textbf{then} \text{ P}= \text{bud}_t, \text{loc} = 1; \\
& \quad \textbf{else if} \text{ error}_t + \Delta_t > \text{errorP} \textbf{then} \text{ P}= \text{bud}_t; \\
& \quad \textbf{else} \text{ clear(bud}_t), \text{loc} ++
\end{align*}
\]

Where \( t \) is time, \( loc \) shows the number of times in which the algorithm has no improvement, \( P \) is a randomly generated weight matrix as a parent, and the cost function is simply calling Eq (5). This algorithm starts with generating a random offspring from the pattern using a strategy namely budding (fully explained in Mansouri et al. [2011] and Farasat et al. [2010]) and calling the cost function for this generated offspring. If the error of the new solution is less than its parent, we accept it, remove the parent and make it as a new parent. However, we will accept the generated solution if its error is not greater than \( \Delta_t \) amount from its parent error. \( \Delta_t \) is calculated using Eq (6). Finally, we discard the offspring if it meets none of these conditions.

\[
\Delta_t = \frac{\ln(loc)}{\sqrt{t}} \quad \text{Eq}(6)
\]

The algorithm ends when it gets to the stopping condition. The final solution is the aimed FCM’s weight matrix.

**DEVELOPING THE DATA GATHERING INSTRUMENT**

A review of the recent relevant literature has been conducted to prepare the research data gathering instrument. Six distinct categories of antecedents are identified as course characteristics (six items), LMS characteristics (four items), student characteristics (five items), student engagement (12 items), student support (seven items), and institutional factors (three items), along with measurement items within each category. Furthermore, we assessed the student’s overall performance using three items of OP1: student satisfaction score, OP2: knowledge construction level, and OP3: student GPA (Grade Point Average). Excluding measurement items under the student engagement category (SE1 to SE12, or part B of the questionnaire) extracted directly from the Moodle LMS, others (Part A) were surveyed through online questionnaires (see Appendix A for the list of measured items).

**CREATING THE LFCM MODEL**

To create the LFCM model, we used the 37 antecedents as the model’s first layer and three performance indicators as the second layer (see Figure 2 as a simplified representation of the developed LFCM model). Longitudinal data from three consecutive semesters has been exploited to model the interdependencies.
**DATA GATHERING PROCEDURE**

A longitudinal set of data has been gathered and used to run the research model. The same set of 45 students in an Information System masters program (only course-based enrollments) at an online university (during a four-semester period) has been targeted as the research sample. This university uses Moodle LMS, which is an open-source e-learning platform. A questionnaire (for Part A) has been developed and distributed using the “questionnaire” plugin of Moodle, capable of linking and tracing responses back to students.

At the end of each semester, the sample students were asked to fill the questionnaire’s Part A. Data for part B (i.e., SE1 to SE12) has been extracted from the Moodle LMS. To date, to the best of the authors’ knowledge, there are no publicly available codes/reports at Moodle to extract this data set. Therefore, SQL queries have been developed and ran to secure access to the required fields. Besides, as most of the measurement items were sums or averages across different courses’ data, we ran such calculations on separate spreadsheets. For the fourth semester, the students were asked only to answer OP1, OP2, and OP3. Data for the fourth semester will be used to evaluate the accuracy of the fourth semester’s predicted performance scores. At the end of the period, we had 30 (out of 45) valid sample data, as some responses to questionnaire items were not valid, and some of the students quit their program. The data-gathering period was: (1) autumn 2018, (2) spring 2018, (3) autumn 2019, and (4) spring 2019. The individual student performance is predicted for the fourth period, i.e., spring 2019, and compared with the same period’s actual student performance data. For the sake of privacy and ethical issues, data has been anonymized before the analysis.

**FINDINGS**

**PREDICTING STUDENT PERFORMANCE**

In this section, we considered each student separately to construct his/her own FCM regarding performance antecedents (Ai) and overall performance (Pi) for three semesters, using the algorithm described in the Learning FCM section above. We applied the resulted FCM as a special type of Recurrent Neural Network for predicting the overall performance in the last (fourth) semester for each student case. MSE and Standard deviation for three performance indicators (OP1, OP2, and OP3) are presented in Table 1. These values display a good performance of the predicting model, having in mind the small sample size.
### Table 1. MSE and Standard deviation for performance variables

<table>
<thead>
<tr>
<th>Measure</th>
<th>OP1</th>
<th>OP2</th>
<th>OP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.07783908</td>
<td>0.07463091</td>
<td>0.05041522</td>
</tr>
<tr>
<td>StDev</td>
<td>0.05817604</td>
<td>0.08766665</td>
<td>0.05861402</td>
</tr>
</tbody>
</table>

**THE EFFECTS OF ANTECEDENTS ON STUDENT PERFORMANCE**

Based on the augmented adjacency matrix among student performance antecedents \((A_i)\), and student’s overall performance \((P_i)\), and equations 1-4, the path effect from student performance antecedents to student overall performance \((A_i \sim P)\) can be calculated. The results of the calculations are presented in Figures 3 and 4.

*Note: To plot part D, means of OP1, OP2, and OP3 are used.*

**Figure 3. The direct and indirect effects of student performance antecedents on performance**
It is observable from the values in Figures 3 and 4 that the indirect (maximum of minimums drawn by Eq (3) and Eq (4), shown in red line) effects for almost all variables are larger than the associated values for direct effects (blue lines). It should be noted again that the maximum of minimum operator traverses all connecting paths between two concepts, assign the minimum graph value to each path and select that path containing the highest value. Through this approach, the operator can capture indirect effects between two concepts.

For instance, according to Figure 3, part A, S11 has a very strong direct effect on OP1. However, when we count the network of interrelations among all factors, its indirect effect is minor. On the other hand, according to Figure 3, part B, SC4 depicts a minor direct effect while a strong indirect effect on OP2. Such results depict that the causal relations and indirect effects are important and can be intensified or diminished through interrelationships among the proposed model’s factors.

Figure 4. The direct and indirect effects of categories of student performance antecedents on performance
Figure 4 illustrates the magnitude of the categories’ effects on each performance indicator and overall performance. It can be observed from the figures that student engagement carries the largest effect (see part D), while student characteristics experienced the highest augmentation, as the effect size is scaled up from 0.124 in the direct effect to 0.300 in the max of mins effect.

**THE CAUSAL INTERRELATIONS AMONG PERFORMANCE ANTECEDENTS**

Figure 5 also illustrates the causal interrelations among student performance antecedents (see Appendix B for further details). The X-axis indicates each variable’s causal magnitude (rows in Appendix B). The Y-axis shows the magnitude of effects (columns in Appendix B). Based on this, we classify factors into four clusters. The cluster in quadrant 1 includes “autonomous” factors that have weak cause and weak effect. These factors are relatively disconnected from the system. It means that the level or magnitude of the factors belonging to this cluster has nothing to do with other variables’ magnitude. The cluster in quadrant 2 includes “independent” factors with a strong cause but a weak effect. As the title suggests, this cluster’s factors act independently, and based on the positive or negative relationship, intensify or weaken the magnitude of other variables. The cluster in quadrant 3 consists of dependent factors that have a weak cause but a strong effect. Finally, quadrant 4 includes “linkage” factors that have strong cause and effect.

Regarding this, CC#: Course Characteristics, and LMS#: LMS Characteristics mainly fall under quadrant 1, indicating that those variables neither causing nor getting affected by other factors. Variables associated with SE#: Student Engagement and SS#: Student Support also mainly fall under quadrant 2, indicating a strong cause on others, with a moderate effect from other variables. Besides, according to quadrant 2, it can be observed that two out of three institutional factors (IF2 and IF3) fall under this category, proposing that the quality of the infrastructure, hardware, and software of the institute are affecting other variables while receiving the least influence from others. It is also apparent from quadrant 1 of the figure that CC#: course characteristics’ items fall here. Meanwhile, they have the least impact on others, indicating that an improved level of the underlying variables will not lead to enhanced levels at other variables.

![Figure 5. Causal interrelations among student performance antecedents](image-url)
DISCUSSION

This research aims to propose a brand-new approach for student performance prediction using the LFCM approach. To build our LFCM model, we use a longitudinal quantitative and qualitative dataset of 30 masters students in an online program for three consecutive semesters, trying to predict every individual student's performance for the fourth semester. Employing the LFCM approach, we also map the causal map of effects among the observed factors and their impact on student performance. Below, the results obtained is discussed in detail.

Among all factors we identified in the literature, the results show that SE7: total time online, and SE10: the regularity of learning interval in LMS, have the largest effect on overall performance (see Figure 3, part D). It is not a surprising result, as the more time and effort learners regularly dedicate to learning, the more likely they will construct and develop their body of knowledge and outperform other inactive students. This argument is well in line with the findings of prior research (Siegle et al., 2010).

According to Figure 4, Part D, and Figure 5, our findings showed that the course characteristics (e.g., course plan and learning material) category has the least direct influence on overall performance. This finding is completely different from previous findings (Cheng & Chau, 2016). We also found that this category falls under the autonomous quadrant, meaning that it has the least influence on other factors. In contrast, it has been affected moderately by other factors. Similarly, student characteristics (e.g., attitude, self-efficacy, and experience) has the second least direct influence on overall performance. Items of student characteristics do not depict a significant impact on other variables. We also found that the LMS characteristics category is placed third in terms of the least direct effects on a student’s overall performance, with items that fall under the first quadrant. We can conclude that the items belonging to these three categories act autonomously and do not strongly intensify other variables. It does not mean that the related variables are not important in the interrelations among the observed variables. Nevertheless, it argues that the intensity of the relationship is weaker compared to other influencing variables. Therefore, we do not contradict previous research that proposed significant relationships between different elements presented in our model. For instance, previous studies highlighted the role of LMS on the quality of learning. They mentioned that LMS could support ubiquitous learning by giving students access to learning materials anytime, anywhere, and everywhere. It can also help students acquire team-based skills and support collaborative learning, active engagement, and participation in course activities (Nwokeji et al., 2019). However, we argue that in the network of interrelations in our model, the role of other factors outweighs the role of LMS on student performance and on the magnitude of other variables.

Interestingly, the student engagement category has the highest direct effect on a student’s overall performance (see Figure 4, Part D). Student engagement primarily focuses on the time and effort put by the students on the educational activities to achieve the desired learning outcomes and is considered as a proxy for learning outcomes (Pye et al., 2015). Therefore, students’ learning effectiveness and satisfaction should be improved by designing systems and instilling strategies that facilitate, encourage, and reward their engagement (Hu & Hui, 2012). Moreover, learning engagement, which is an important antecedent for learning outcome, is lower for technology-mediated learning than face-to-face learning (Hu & Hui, 2012). It is a good point because, from a pedagogical perspective, e-learning systems have several difficulties transferring emotion or even engage students during a course session (Muntean, 2011). To do so, LMS has to focus on raising students’ engagement in classes to reach a higher level of performance. In this regard, the learning process should encompass more interactive and gamified material to gain greater attention from learners. It also strives to encourage users to engage in desired behaviors in connection with the applications. For instance, we can convert points or badges into virtual goods or even get discounts for tuition fees. Students will engage more with the application and be motivated to earn more points to benefit from these advantages (Muntean, 2011).
We observed a strong impact of student support on overall performance according to Figure 4, Part D, and a significant effect on other variables, according to Figure 5. Most items in this category fall under the fourth quadrant, meaning that they mostly act independently. In other words, improving the level of student support will indirectly result in improvements on other items. One of the interesting findings of this study is that SS7: subjective norm has the strongest influence on other factors (see Figure 5). Subjective norm refers to the person's perception that most people who are important to the person think the person should or should not perform the behavior in question (Ashrafi et al., 2020). This clearly shows that students' perception has a great influence on the overall performance and other factors. To be clear, the subjective norm makes users compare their initial expectations with their modified perceptions. The strong impact of student support elements on other variables is not surprising, as a proper support level can positively affect student engagement, attitudes towards using the LMS, and attending actively in the course. This finding is well in line with prior research results (Grant-Vallone et al., 2003; Hughes & Kwok, 2006).

Establishing action plans to boost student performance should focus on the performance antecedents and the interrelations among them. Figure 5 presents valuable insights for decision-makers when limited resources come into the decision-making context, and educational institutes need to choose which educational aspects they should focus. The presented four-dimensional profile in Figure 4 could be constructive here by highlighting the most important categories and individual factors. More specifically, variables in quadrant 2 and 4 of Figure 5 (with the highest cause) should be of more focus while establishing improvement action plans, as any improvements here will also result in an improved level of affected factors, without direct improvement plans over there.

CONCLUSION

This research proposes a new approach for student performance prediction using the LFCM approach. As mentioned before, although the available set of common prediction approaches, such as ANNs or logistic regression, work well with large datasets, they face challenges dealing with small sample sizes, limiting their practical applications in real practices. Given this limitation, it is important to present a new approach for student performance prediction using new approaches such as LFCM. From this perspective, the proposed LFCM model constitutes practical and managerial implications for the e-learning community. We found that the student engagement category has the highest direct effect on a student's overall performance and student support strongly intensifies other items in our proposed model. Even though our study offers some contributions, it suffered from several limitations that must be tackled in future studies. First, this study is by no means comprehensive to address all student performance antecedents. Even though we reviewed the most relevant and recent literature to identify the performance determinants, we were confined with the length of the questionnaire (especially in longitudinal research) and a set of indices we could extract from Moodle. Second, we do not aim to claim our findings' generalizability, as our results are based on a single master's program and a single university. So simply, the interrelationships are extracted and interpreted for a limited context. As a potential for future work, the proposed approach could be used in other contexts to test its applicability. Third, it is important to stress that the predictive power of LFCM models is far from perfect and far from other techniques such as ANN and regression. However, LFCM performs well on limited sample sizes. Future studies could improve the performance level of the proposed LFCM model by tuning the model's elements.

REFERENCES


Student Performance Prediction using LFCM


Kireev, V. S. (2016). Development of fuzzy cognitive map for optimizing e-learning course. International Conference on Data Analytics and Management in Data Intensive Domains. https://doi.org/10.1007/978-3-319-57135-5_4


Student Performance Prediction using LFCM


Student Performance Prediction using LFCM


# Appendix A. The Measurement Items Used for Data Gathering

<table>
<thead>
<tr>
<th>Measurement items</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Part A</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Course Characteristics (CC)</strong></td>
<td></td>
</tr>
<tr>
<td>CC1. Course plan (syllables and course assessment method)</td>
<td>(Hadullo et al., 2017; Tarus et al., 2015)</td>
</tr>
<tr>
<td>CC2. List of textbooks/ references</td>
<td></td>
</tr>
<tr>
<td>CC3. List (and CV) of lecturer(s)</td>
<td></td>
</tr>
<tr>
<td>CC4. List (and CV) of teaching assistant(s)</td>
<td></td>
</tr>
<tr>
<td>CC5. Quality learning material</td>
<td></td>
</tr>
<tr>
<td>CC6. Interactive/Gamified learning material</td>
<td></td>
</tr>
<tr>
<td><strong>LMS Characteristics (LMS)</strong></td>
<td></td>
</tr>
<tr>
<td>LMS1. Perceived usefulness</td>
<td>(Kaba &amp; Osei-Bryson, 2013; Panigrahi et al., 2018; Ros et al., 2015; Sharma et al., 2017)</td>
</tr>
<tr>
<td>LMS2. Perceived ease of use</td>
<td></td>
</tr>
<tr>
<td>LMS3. Perceived enjoyment</td>
<td></td>
</tr>
<tr>
<td>LMS4. Perceived LMS system quality (Reliability, availability, response time, …)</td>
<td></td>
</tr>
<tr>
<td><strong>Student Characteristics (SC)</strong></td>
<td></td>
</tr>
<tr>
<td>SC1. Attitude</td>
<td>(Al-Azawei et al., 2016; Hadullo et al., 2017; Hwang, 2015; Kisanga, 2016; Kuo &amp; Kituyi, 2012; Makokha &amp; Mutisya, 2016; Panigrahi et al., 2018)</td>
</tr>
<tr>
<td>SC2. Self-efficacy</td>
<td></td>
</tr>
<tr>
<td>SC3. motivation, incentives</td>
<td></td>
</tr>
<tr>
<td>SC4. experience</td>
<td></td>
</tr>
<tr>
<td>SC5. Self-regulation</td>
<td></td>
</tr>
<tr>
<td><strong>Student Support (SS)</strong></td>
<td></td>
</tr>
<tr>
<td>SS1. Interactions with peers</td>
<td>(Baloyi, 2014a, 2014b; Chawinga &amp; Zozie, 2016; Makokha &amp; Mutisya, 2016; Muuro et al., 2014; Panigrahi et al., 2018; Queiros &amp; de Villiers, 2016)</td>
</tr>
<tr>
<td>SS2. Interactions with teaching assistants</td>
<td></td>
</tr>
<tr>
<td>SS3. Interactions with lecturer(s)</td>
<td></td>
</tr>
<tr>
<td>SS4. Assignment feedback</td>
<td></td>
</tr>
<tr>
<td>SS5. Exam feedback</td>
<td></td>
</tr>
<tr>
<td>SS6. Technical support for LMS and Virtual Classes with convenient operating hours</td>
<td></td>
</tr>
<tr>
<td>SS7. Subjective Norm</td>
<td></td>
</tr>
<tr>
<td><strong>Institutional Factors (IF)</strong></td>
<td></td>
</tr>
<tr>
<td>IF1. Perceived quality of maintenance of infrastructure, seminars, and workshops</td>
<td>(Hadullo et al., 2017; Kashorda &amp; Waema, 2014; Ngandu &amp; Brown, 2015; Ssekakubo et al., 2011; Tarus et al., 2015)</td>
</tr>
<tr>
<td>IF2. Perceived quality of institutional infrastructure</td>
<td></td>
</tr>
<tr>
<td>IF3. Perceived quality of institutional hardware and software</td>
<td></td>
</tr>
<tr>
<td><strong>Overall Performance (OP)</strong></td>
<td></td>
</tr>
<tr>
<td>OP1. User satisfaction score</td>
<td>(Ferguson &amp; DeFelice, 2010; Lai et al., 2015; Misopoulous et al., 2018; Raspopovic et al., 2014; Zhang &amp; Cheng, 2012; Zhu, 2012)</td>
</tr>
<tr>
<td>OP2. Knowledge construction level</td>
<td></td>
</tr>
<tr>
<td>OP3. Student GPA (Grade Point Average)</td>
<td></td>
</tr>
</tbody>
</table>
## Measurement items

<table>
<thead>
<tr>
<th>Part B*</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student Engagement (SE)</strong></td>
<td></td>
</tr>
<tr>
<td>SE1. Total studying time in LMS (min)</td>
<td>Dias et al., 2015; Hadullo et al., 2017</td>
</tr>
<tr>
<td>SE2. Total login frequency in LMS</td>
<td></td>
</tr>
<tr>
<td>SE3. Number of discussion posts views</td>
<td></td>
</tr>
<tr>
<td>SE4. Number of resources viewed</td>
<td></td>
</tr>
<tr>
<td>SE5. Number of course page views</td>
<td></td>
</tr>
<tr>
<td>SE6. Average time per session (min)</td>
<td></td>
</tr>
<tr>
<td>SE7. Total time online (min)</td>
<td></td>
</tr>
<tr>
<td>SE8. Numbers of links viewed</td>
<td></td>
</tr>
<tr>
<td>SE9. Number of content page views</td>
<td></td>
</tr>
<tr>
<td>SE10. The regularity of learning interval in LMS</td>
<td></td>
</tr>
<tr>
<td>SE11. Total number of discussion posts</td>
<td></td>
</tr>
<tr>
<td>SE12. Average assessment quiz grade</td>
<td></td>
</tr>
</tbody>
</table>

* Seven point Likert scale (strongly low to strongly high) for all measurement items of Part A (excluding OP3: GPA) is used; Data for Part B is extracted from Moodle LMS at the end of each semester

## APPENDIX B. INTERRELATIONS AMONG VARIABLES

Appendix B can be downloaded from [https://doi.org/10.28945/4760](https://doi.org/10.28945/4760) or from [https://www.informingscience.org/Publications/GetFile/6977?ArticleFileID=7639](https://www.informingscience.org/Publications/GetFile/6977?ArticleFileID=7639)

## AUTHORS

**Taha Mansouri** is a Postdoc Researcher at the School of Science, Engineering, and Environment, the University of Salford, Manchester, UK. He has published in acclaimed journals, such as the Knowledge-Based Systems, Expert Systems with Applications, Applied Soft Computing, Optic Letters, and Telecommunication Systems. His research interests include Machine Learning, Deep Learning, Reinforcement Learning, Computer Vision, and Complex Systems Modelling.
Ahad ZareRavasan is an Assistant Professor at the Department of Corporate Economy, Faculty of Economics and Administration Masaryk University, Brno, Czech Republic. He has published in acclaimed journals, such as the International Journal of Information Management, Expert Systems with Applications, Information Systems, International Journal of Production Research, Journal of Business and Industrial Marketing, and the Journal of Enterprise Information Management. His research interests include Information Systems, e-learning, Business Analytics, and Blockchain.

Amir Ashrafi is currently a Ph.D. student of business & management in the Alliance Manchester Business School at the University of Manchester. He has published a number of papers in acclaimed journals, such as the International Journal of Information Management, Journal of Enterprise Information Management, Journal of Business and Industrial Marketing, Interactive Learning Environment, and Journal of Organizational and End User Computing. His research focuses on Digital Innovation & Transformation, Big data analytics, Artificial Intelligence, and IT Business Value.