ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/344378601

Developing an Optimized Feature Selection Process for Designing Efficient Content Management System using Educational Data

Article · July 2020

CITATIONS
0

READS
54

2 authors, including:

Vijayalakshmi v.
Sri Manakula Vinayagar Engineering College
45 PUBLICATIONS **32** CITATIONS

SEE PROFILE

All content following this page was uploaded by Vijayalakshmi v. on 25 September 2020.

The user has requested enhancement of the downloaded file.

IJARSCT

ISSN (Online) 2581-9429

IJARST

International Journal of Advanced Research in Science & Technology (IJARST)

Volume 7, Issue 1, July 2020

Developing an Optimized Feature Selection Process for Designing Efficient Content Management System using Educational Data

V. Vijayalakshmi[1] and A. Prakash[2]

[1,2]Assistant Professor, Department of Computer Science and Engineering,
Sri Manakula Vinayagar Engineering College, Pondicherry
vivenan09@gmail.com

Abstract: As of now, the process of learning and teaching is strongly supported by educational technologies, in both offline and online learning platforms. The advancements in technology have supported educational actors by providing relevant information and thereby promote the quality and innovations in educational context. In order to obtain better quality of e-learning, the pedagogical attributes have to be evaluated continuously for achieving collaborative environment. Due to the growing interest in e-learning technologies, the acquisition of relevant knowledge from educational information is a challenging task. Data mining is the field that seamlessly helps us to explore the knowledge from different evaluation aspects. Feature selection is an important step in the data mining process. This paper explores an innovative process to select the features using pedagogical data. The proposed framework composes of two phases, namely, data pre-processing and feature selection. In pre-processing phase, single linked list is been employed to remove the duplicates values, so as to enhance the memory computational process. Canonical Correlation Analysis (CCorA) is used for deriving the relationship among the attributes and knowledge for the given education data. Experimental analysis is carried out in two datasets of UCI machine repository, namely, teaching assistant evaluation and Turkiye student evaluation dataset. Each dataset have its own attributes and the characteristics. Evidently, the results say that the proposed work is concentrated on selecting the features required for content management systems across two different datasets which paves the path for researchers in the field of e-learning.

Keywords: Educational Data, Content Management System, Correlation Analysis, E-Learning and Feature Selection.

I. INTRODUCTION

In today's environment, the growth of data is increasing without any measures due to the development of web technologies. The improvement of technological process has enabled the students to learn in different ways. The universities have launched different modes of education systems i.e. online learning and offline learning systems [1]. Regardless of growing opportunities for students and faculties, the e-learning (or) online learning brings challenges due to the absence of efficient data analyzing and interpreting systems. Online environments allow the generation of large amounts of data related to learning/teaching processes, which offers the possibility of extracting valuable information that may be employed to improve students' performance. In conventional educational systems, faculties play a vital role in learning process [2]. The students acquire basic knowledge and skills through knowledge shared by the faculty. The development of web has drastically changed the learning system by providing online courses. Besides these developments, the analysis and interpretation of educational data is still challenging tasks. The figure1 states the overview of how EDM methods are applied.

Data uncertainty and inappropriate features selection are the major challenging tasks in analyzing and interpretation of educational data. Data uncertainty is the probability of irrelevant data presented in source data. This irrelevant data degrades the knowledge pattern extraction by causing noises and redundant features. To overcome from these issues, data preprocessing should be done properly. The main reasons to preprocess the datasets are 1) reduction of the size of

IJARST
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science & Technology (IJARST)
Volume 7, Issue 1, July 2020

the dataset in order to achieve more efficient analysis, and 2) adaptation of the dataset to best suit the selected analysis method [3]. The size of the dataset can be reduced by performing feature set reduction (or) feature subset reduction. The problem is important, because a high number of features in a dataset, comparable to or higher than the number of samples, lead to model over-fitting, which in turn leads to poor results on the validation datasets. Additionally, constructing models from datasets with many features is more computationally demanding [4]. Feature selection [5] is one of the important steps in machine learning systems.

The main aim of feature selection methods is to choose the highly discriminating features. Both Feature extraction and feature selection are capable of improving learning performance, lowering

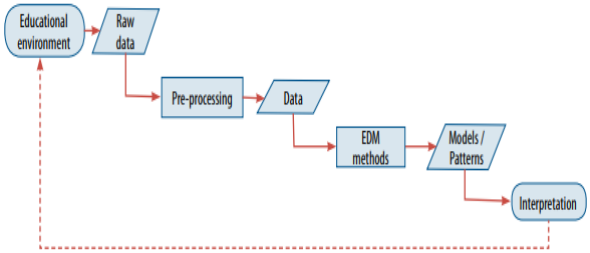

Figure 1: Applying EDM methods [2]

Computational complexity, building the generalized models, and decreasing required storage. Feature extraction [6] maps the original feature space to a new feature space with lower dimensions by combining the original feature space. It is difficult to link the features from original feature space to new features. Therefore, further analysis of new features is problematic since there is no physical meaning for transformed features obtained from feature extraction techniques. While feature selection selects a subset of features from the original feature set without any transformation, and maintains the physical meanings of the original features. In this sense, feature selection is superior in terms of better readability and interpretability [7]. This property has its significance in many practical applications such as finding relevant genes to a specific disease and building a sentiment lexicon for sentiment analysis. Typically, feature selection and feature extraction are presented separately.

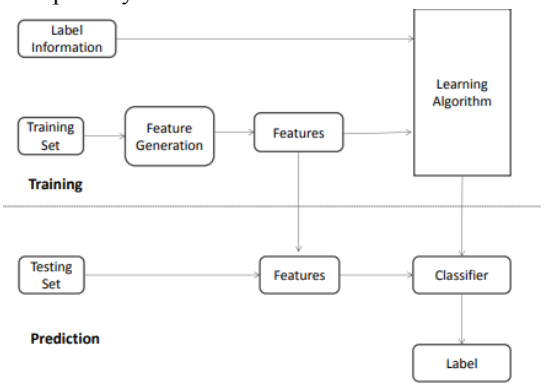

Figure 2: Principle model of Data classification methods [7]

The main contributions of this paper are:

- We observe the present need of content management systems in e-learning.
- The data collection intimates the role of our study in educational content and its context.
- To the best of our knowledge, single linked list is used as preprocessing model to enhance the computational time.
- Then, Canonical correlation analysis is employed to find the relationship among the attributes, as feature selection process, that depicts a path for designing efficient content management systems.
- Finally, the proposed system derives the knowledge from two different datasets.

IJARST
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science & Technology (IJARST)
Volume 7, Issue 1, July 2020

The rest of the paper is organized as follows: Section II presents the related work; Section III presents the proposed work; Section IV presents the experimental analysis and finally concludes at Section V.

II. Related Works

This section presents the existing techniques adopted in e-learning systems. In [8, 9] presented a model, named, RELOAD that provide lab experiments at remote for limited fields of engineering. Modality of the data is not focussed due to independency of the student. The challenges in personalized e-learning are discussed by [10]. The authors have presented architecture for student and instructor that included important attributes of the system. The exploration of knowledge via memetic agents presented by [11] studied by use of ontologies. Learning path determined the better learning services for personalized e-learning systems. But, the system lacks the optimal fitness solutions for feature subsets. In [12], the authors studied about disengagement of the students in online learning systems. At first, content delivery systems of the learning experience using educational data mining models is analyzed. The relevancy of the learners attributes is explored. It was experimented in HTML tutoring systems where the search operation fails in larger content delivery systems.

The authors in [13] studied about context aware recommender systems for e-learning systems. Technology Enhanced Learning (TEL) recommender system is been studied to identify the relevancy of the context dimensions using learning attributes. The concept of re-ranking services in recommender systems fails due to resource inefficiency and higher computational costs. Then, the study is extended to similarity pattern of the learner's attributes using adaptation strategies. It was approached in two forms namely, typology pattern of adaptation navigation and combining the similar patterns using adaptation strategies. The meta-rules generated for extracting similar patterns restricts the domain variables. In order to resolve the resource inefficiency of the e-learning systems, in [14], studied a directed hypergraph for better resource optimization that explored the significance of the process model and their relationship among the learners and the available resource environment. This model is imbalanced for large scale resource environment. European Qualification Frameworks (EQF) is been devised by [15] which enhanced the performance of teaching qualities of the teachers in European country. The deployment of cognitive agents of the visualization tasks is not defined properly.

The authors in [16] studied about the dynamic learning style using pattern recognition systems. They introduced a learning style of the student by extracting similar patterns towards their engagement in online courses. Most of the issues in cognitive overload and the learner's deviation have been resolved. Classification accuracy and the dimensionality reduction are not enhanced. Student modeling for personalized e-learning is studied by [17]. It has been studied in OSSM interface, Mastery Grids that described the student's attributes based on social dimension. Low-group variables have been restricted to discover the knowledge. In similar way, blended learning model is studied by [18] using recommender systems and web 2.0 technologies. They explored on t-test model which doesn't support the mind mapping systems. The authors in [19] studied for learning objects using profiles of the teachers. A recommendation model was designed based on the learning objects provided by the teachers that degraded the robustness in collaborative filtering models. In [20, 21], they explored the learning indicators and their engagement in online learning systems. They have explored how to combine data about self-regulated learning skills with observable measures of online activity in a blended learning course to increase predictive capabilities of student academic performance for the purposes of informing teaching and task design.

III. Research Methodology

This section presents the research methods developed to select optimal features using educational data. The proposed framework composes of several phases, namely,

A. Data acquisition & Data Preprocessing

Data acquisition is the foremost step in our framework. The selection of appropriate datasets from different sources depicts the efficiency of the systems. Thus, two datasets, namely, Teaching assistant evaluation and Turkiye student evaluation datasets from UCI machine repository have been selected [22]. The teaching assistant evaluation dataset

IJARST
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science & Technology (IJARST)
Volume 7, Issue 1, July 2020

depicts the simpler data whereas the Turkiye student evaluation datasets depicts the different attributes of teaching quality. The common aim of those datasets is to evaluate the performance of the teachers. In our study, we used to find the attributes required for designing content management system. The datasets details are been discussed in table I & table II.

Database name	No. of instances	No. of attributes	Missing values	Attribute characteristics	Dataset characteristics
Teaching assistant evaluation	151	5	No	Classification	Multivariate
Turkiye student evaluation	5820	33	No	Classification	Multivariate, sequential and time-series

Table 1: Dataset details

Attributes Names	Attribute details
Teaching Assistant knows English or not	1= English Speaker 2= Non- English speaker
Course Instructor	25 categories
Course	26 courses
Summer or regular semester	1= Summer 2= Regular
Class size	Numerical
Class attribute	1= low 2= medium 3 =high

Table 2: Attribute Descriptions 2(a): Teaching Assistant Evaluation Dataset

Attributes	Possible ranges
Instructor's identifier	1,2,3
Class (course code)	(1-13)
No. of times student handle the class	(0,1,2,…)
Attendance level	(0,1,2,3,4)
Course difficulty level	(1,2,3,4,5)
Teaching method with accurate content (Q1)	(1,2,3,4,5)
Objectives of the course is clearly stated (Q2)	(1,2,3,4,5)
Credited amount worth for the course (Q3)	(1,2,3,4,5)
Class discussion is satisfactory (Q5)	(1,2,3,4,5)
Resources were efficient and updated (Q6)	(1,2,3,4,5)
Projects and quizzes are held to enhance the knowledge (Q8)	(1,2,3,4,5)
Enjoying the sessions and participate in class discussions (Q9)	(1,2,3,4,5)
Initial expectations were followed till end of the period (Q10)	(1,2,3,4,5)
Course is beneficial for professional development (Q11)	(1,2,3,4,5)
Course changed my lifestyle and challenges (Q12)	(1,2,3,4,5)
Instructor's knowledge is updated (Q13)	(1,2,3,4,5)
Instructor came prepared for classes (Q14)	(1,2,3,4,5)
Well-planned for taking classes (Q15)	(1,2,3,4,5)
Instructor was committed to the course and was understandable (Q16)	(1,2,3,4,5)
Instructor arrived on time for classes (Q17)	(1,2,3,4,5)
Instructor has a smooth and easy to follow delivery/speech (Q18)	(1,2,3,4,5)

IJARST
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science & Technology (IJARST)
Volume 7, Issue 1, July 2020

Instructor made effective use of class hours. (Q19)	(1,2,3,4,5)
Instructor explained the course and was eager to be helpful to students (Q20)	(1,2,3,4,5)
Instructor demonstrated a positive approach to students (Q21)	(1,2,3,4,5)
Instructor was open and respectful of the views of students (Q22)	(1,2,3,4,5)
Instructor encouraged participation in the course (Q23)	(1,2,3,4,5)
Instructor gave relevant homework assignments/projects, and helped/guided student (Q24)	(1,2, 3,4,5)
Instructor responded to questions about the course inside and outside of the course (Q25)	(1,2,3,4,5)
Instructor's evaluation system effectively measured the course objectives (Q26)	(1,2,3,4,5)
Instructor provided solutions to exams and discussed them with students (Q27)	(1,2,3,4,5)
Instructor treated all students in a right and objective manner (Q28)	(1,2,3,4,5)

2(c): Turkiye Student Evaluation Dataset

The two sorts of datasets ensures absence of missing values, but, the duplicates removal process is done to enhance the memory computation that is increases accuracy by decreasing the running time of the optimized feature selection model. Single linked list is used as data preprocessing model to remove the duplicates presented in the database. The pseudo- code is given as follows:

Algorithm 1: Preprocessing model
Input: Integer data
Operator: Deletion
Output: Pre-processed data
Steps:

a) Let each data be a 'node'.
b) Create a linked structure for each node. i.e data(information) and next (Address of the next node)
c) The first node is set to head i.e. firstnode. Head → NULL
d) Deleting the specific node via duplicate removal function

```
Void deleteNode ( Struct Node* & start)
Node* temp = New node;                  //creating a location for temporary node
Temp→next= head                         // Temporary node points to head
Node * prev =temp                       // Ensure no duplicates
Node * present = Head ;
While (present !=NULL)
{
    While (present → Next != NULL && prev→ Next →   data == present →Next→ data )
  Present= Present → next;
If ( prev→next== present)
Present = present→ next;
Else
Prev→ next= current→ next;
Present= present→ next
}
 Head = Temp→ next                      // update original head to next of temp node
```

Dataset	No. of instances (Before duplication removal)	No. Of instances (After duplication removal)	No. of duplicates
Teaching assistant evaluation	151	111	40
Turkiye student evaluation	5820	5820	No duplicates

Table 3: Pre-processed data

IJARST
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science & Technology (IJARST)
Volume 7, Issue 1, July 2020

B. Feature Extraction & Selection

Feature selection is an important step that determines the efficiency of the systems. The main task of feature selection is to select the relevant features that enhance the reliability of the systems. Canonical correlation analysis is been designed to select optimal features across the two datasets. Though, it is one of the oldest techniques, the discovery of correlation across the sets explores the efficiency of this method. Let X is the set of teaching assistant evaluation data and Y is the set of education process mining data. The correlation between these two datasets are been explored to find the optimal features of the educational systems for upcoming content management systems. The canonical correlation is defined as the correlation between the corresponding pairs of canonical variables. The proposed steps are as follows:

i) Defining the linear transformations of the given data.

Let $X = \{x_1, x_2 \ldots x_m\}$ be the set of dataset 1 & $Y = \{y_1, y_2 \ldots y_n\}$ be the set of dataset 2. The linear transformations of two variables are given:

A. 1st transformation:

$$K_1 = p_{11}x_1 + p_{12}x_2 + \cdots . p_{1m}x_m$$
$$K_2 = p_{21}x_1 + p_{22}x_2 + \cdots . p_{2m}x_m$$
$$K_m = p_{m1}x_1 + p_{m2}x_2 + \cdots . p_{mm}x_m$$

B. 2nd transformation:

$$L_1 = q_{11}y_1 + q_{12}y_2 + \cdots .. q_{1n}y_n$$
$$L_2 = q_{21}y_1 + q_{22}y_2 + \cdots .. q_{2n}y_n$$
$$L_n = q_{n1}y_1 + q_{n2}y_2 + \cdots .. q_{nn}y_n$$

Therefore, (K_m, L_n) is the canonical variate pair.

ii) Defining the canonical variables

The covariance between members of each canonical variate pairs is formulated as:

$$Var(K_i) = \sum_{u=1}^{a}\sum_{v=1}^{b} a_{iu}b_{iv}\, Cov\,(x_u, x_v)$$
$$Var(L_j) = \sum_{u=1}^{a}\sum_{v=1}^{b} a_{ju}b_{jv}\, Cov\,(y_u, y_v)$$
$$Cov(K_i, L_j) = \sum_{u=1}^{a}\sum_{v=1}^{b} a_{iu}b_{jv}Cov(x_k, y_l)$$

The correlation between (K_i, L_j) is given as:

$$Cor(K_i, L_j) = \frac{Cov\,(K_i, L_j)}{\sqrt{Var\,(K_i)}\,\sqrt{Var\,(L_j)}}$$

iii) Defining the canonical variation:

The canonical variation across the two sets is given as:

$$\rho_i^* = \frac{Cor\,(K_i, L_j)}{\sqrt{Var\,(K_i)}\,\sqrt{Var\,(L_j)}}$$

Where ρ_i^* defines the relationship between the variables of two datasets. Figure 3 is the proposed diagram.

ISSN (Online) 2581-9429

IJARST

International Journal of Advanced Research in Science & Technology (IJARST)

Volume 7, Issue 1, July 2020

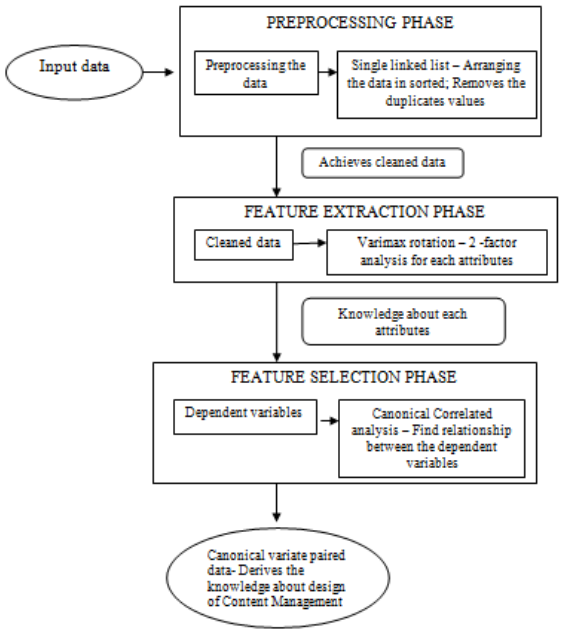

Figure 3: Proposed workflow

IV. RESULTS AND DISCUSSION

This section presents the experimental analysis of our proposed framework. The two different datasets are been obtained from UCI Machine Repository, namely, teaching assistant evaluation and Turkiye student evaluation to evaluate the efficiency of our proposed framework. By using single linked list, the data is pre-processed.

Varimax rotation is a simpler rotation process of small number of large loadings and large number of zero loadings. By doing so, the interpretations of the attributes are easy. In addition, the factors can often be interpreted from the opposition of few variables with positive loadings to few variables with negative loadings. The variance of the each attributes is estimated in factors in order to achieve the optimal attribute selection process. According to Eigen values and 2-factors solution are processed, and their values are given in Table 4.

Attributes	Rotated values	
	Factor 1	Factor 2
Q1	0.339	0.831
Q2	0.474	0.798
Q3	0.559	0.707
Q4	0.455	0.800
Q5	0.505	0.794
Q6	0.498	0.778
Q7	0.469	0.816
Q8	0.459	0.815
Q9	0.541	0.711
Q10	0.527	0.789
Q11	0.560	0.695
Q12	0.478	0.768
Q13	0.763	0.549
Q14	0.804	0.508

IJARSCT
IJARST
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science & Technology (IJARST)
Volume 7, Issue 1, July 2020

Q15	0.801	0.506
Q16	0.718	0.600
Q17	0.842	0.381
Q18	0.775	0.529
Q19	0.799	0.510
Q20	0.828	0.473
Q21	0.844	0.449
Q22	0.846	0.446
Q23	0.764	0.557
Q24	0.724	0.584
Q25	0.835	0.456
Q26	0.764	0.531
Q27	0.711	0.558
Q28	0.824	0.443
Course instructors	0.653	0.536
Class size	0.563	0.435

Table 4: Attribute analysis – Factor analysis

The table 4 depicts the analysis of each attributes under 2 factors. Both factor 1 and factor 2 should satisfy the greater than threshold value, 0.5 are extracted for further process.

Datasets	**Time (ms)**
Teaching assistant evaluation dataset (1)	1.256
Turkiye student evaluation dataset (2)	5.369

Table 5: Preprocessing time analysis

The table 5 presents the computational time analysis of the educational data. It is observed that 1.256ms taken by teaching assistant evaluation dataset and 5.369ms taken by Turkiye student evaluation dataset. The attributes in dataset 2 is of 33 which yields higher computational time.

Datasets	Categorization	Selected attributes	Canonical Correlation (CCorA) **(Proposed)**	Neighbourhood similarity **(Existing)**
Teaching assistant evaluation dataset (1)	Speaker (Native English (1) & Non-English speaker (2))	Course Instructors & Class size	3.615	7.6
Turkiye student evaluation dataset (2)	3 instructors	Q3, Q5, Q9, Q10, Q11, Q13, Q14, Q15, Q16, Q18, Q19, Q23, Q24, Q26, Q27	5.234	6.9

Table 6: Information analysis

The table 6 & figure 4 depicts the information analysis process of two datasets. It is observed that the proposed CCorA yields better information than the existing work. The dataset 1 explores 3.615 and the dataset 2 gives 5.234 which show that the relationship between two datasets is quite correlated. In order to design efficient content management systems, the e-learning content has to satisfy both the teachers and the students.

IJARST
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science & Technology (IJARST)
Volume 7, Issue 1, July 2020

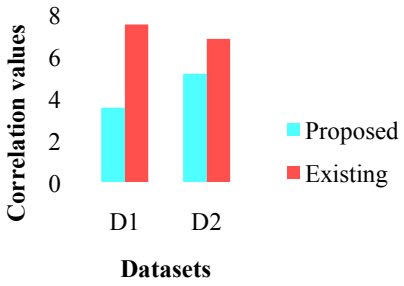

Figure 4: Information analysis between proposed and existing system

V. CONCLUSION

This paper researches the demand for innovative learning technologies in Content Management Systems of online learning systems. Thus, an optimized parameterization is required to design or update the content and context of e-learning systems. The advancement made in e-learning is associated with technical improvement with better affordability. In this paper, we propose an optimized feature selection model that depicts the significance of each attributes in pedagogical data. Initially, the two dimensional teaching dataset is been collected from UCI machine repository. The collected datasets are subjected to irrelevant data which is pre-processed using Single linked list. Each pedagogical data is indexed and analyzed by comparing the records with previous records. If more than a single value of the data matches with existing records, then it's declared as duplicate values and it is discarded. By doing so, we achieve better processing time. The cleaned data is then analyzed by Varimax rotation which shows the significance of each attributes. From this, the canonical correlation analysis (CCorA) is deployed as feature selection model that aids to derive the knowledge by finding the relationship across two datasets.

The knowledge derived from those results is:

- **a)** Students are not satisfied by the amount credited to the course that dictates classes are effectively handled by the instructors in terms of presentation skills.
- **b)** The students are not actively participated in the class discussions, quizzes and field work and their expectations are met.
- **c)** The instructor effectively delivers the course content that helped student look at life and the world with a new perspective.
- **d)** The designing process should be identified with valuable information, interaction, interface and presentation.
- **e)** The deployment of content and information needs to be conveyed efficiently.
- **f)** System evaluation should be done to test the feasibility of the technical aspects of the systems.

REFERENCES

[1]. Popp R L, Pattipati K R, Bar-Shalom Y. "m-Best S-D assignment algorithm with application to multitarget tracking". IEEE Trans. on AC, 2001, 37 (1):22 - 38.

[2]. Changzhong wang et al, "Feature selection based on neighbourhood discrimination index", IEEE transactions on neural networks and learning systems, 2017.

[3]. Peña-Ayala A, "Educational data mining: A survey and a data mining-based analysis of recent works", Expert systems with applications, 41(4), 2014.

[4]. Andonie R. "Extreme data mining: Inference from small dataset", International Journal of Computers Communications and Control, 5(3), 2010.

[5]. Romero C, Ventura S. "Educational data mining: A survey from 1995 to 2005", Expert systems with applications, 33(1), 2007.

[6]. Minaei-Bidgoli B, Punch WF," Using genetic algorithms for data mining optimization in an educational web based system". Genetic and Evolutionary Computation, Springer Berlin Heidelberg, 2003.

IJARST
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science & Technology (IJARST)
Volume 7, Issue 1, July 2020

[7]. Ashish Dutt et al, "A Systematic Review on Educational Data Mining", IEEE transactions on content mining", Iss.99, 2017.
[8]. S. M. Merchán and J. A. Duarte, "Analysis of Data Mining Techniques for Constructing a Predictive Model for Academic Performance", IEEE LATIN AMERICA TRANSACTIONS, 14(6), 2016.
[9]. Camilo E. López G et al, "A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining", IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, 10(3), 2015.
[10]. Carlos Márquez-Vera et al, "Predicting School Failure and Dropout by Using Data Mining Techniques", IEEE Journal of Latin-American Learning Technologies, 8(1), 2013.
[11]. Mihaela Cocea and Stephan Weibelzahl, "Disengagement Detection in Online Learning: Validation Studies and Perspectives", IEEE Transactions on Learning Technologies, 4(2), 2011.
[12]. Mustafa Agaoglu, "Predicting Instructor Performance Using Data Mining Techniques in Higher Education", IEEE transactions on content mining, 4, 2016.
[13]. Xin Chen et al, "Mining Social Media Data for Understanding Students' Learning Experiences", IEEE Transactions on Learning Technologies, 7(3), 2014.
[14]. Crist´ obal Romero et al, "Educational Data Mining: A Review of the State of the Art", IEEE Transactions on Systems, Man, And Cybernetics, 40(6), 2010.
[15]. Jui-Long Hung et al, "Identifying At-Risk Students for Early Interventions–A Time-Series Clustering Approach", IEEE Transactions on Emerging Topics in Computing, 5(1), 2017.
[16]. Petra Vrablecová, and Marián Šimko, "Supporting Semantic Annotation of Educational Content by Automatic Extraction of Hierarchical Domain Relationships", IEEE Transactions on Learning Technologies, 9(3), 2016.
[17]. Baradwaj, B.K. and Pal, S., "Mining Educational Data to Analyze Students' Performance", International Journal of Advanced Computer Science and Applications, 2(6), 2011.
[18]. Ahmed, A.B.E.D. and Elaraby, I.S., "Data Mining: A prediction for Student's Performance Using Classification Method", World Journal of Computer Application and Technology, 2(2), 2014.
[19]. Pandey, U.K. and Pal, S., "Data Mining: A prediction of performer or underperformer using classification", International Journal of Computer Science and Information Technologies, 2 (2), 2011.
[20]. Bhardwaj, B.K. and Pal, S, "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security, 9(4), 2011.
[21]. Yadav, S.K., Bharadwaj, B. and Pal, S, "Data Mining Applications: A Comparative Study for Predicting Student's Performance", International Journal of Innovative Technology & Creative Engineering, 1(12), 2011.
[22]. Dataset : https://archive.ics.uci.edu/ml/datasets.html

View publication stats