A Decision Support System for Effective Academic Analysis by Using the Concept of Data Mining

Aaditya Jain[1], Megha Jain[2], Sandeep Saxena[3]

[1,2,3]College of Computing Sciences and Information Technology, Teerthanker Mahaveer University, Moradabad,
India
Email: [1]aadityajain58@gmail.com, [2]meghaj409@gmail.com, [3]sandeepsaxenambd2@gmail.com

ABSTRACT

In India, higher education institutions experience the need for effective analysis and decision support tools to assist in all academic processes. Management needs to take effective administrative decisions regarding the quality of education and efficient use of resources. Some of the higher learning institutions only manage financial, administrative and academic data used for common analysis works. The proposed conceptual system is designed to work as an integral part with its sub components for supporting academic and administrative decision making. For effective decision making, this system can be used to manage appropriate academic analytics & databases to collect, synthesize, and evaluate educational data.

Key words: Higher Education System; Data Mining; Decision Support System; HEDMS; Education Policy.

1. INTRODUCTION

In the current era of technology sophisticated tools are required to process and store educational data. The University Information System and software tools should be well equipped with functional analysis techniques to support data-driven decision making to enhance the overall academic performance. Decision Support Systems (DSS) are interactive computer based information systems that support the decision makers. It is used to identify the critical issues and their possible solutions to make efficient decisions. Decision makers that follow DSS are equally capable to work in semi-structured as well as unstructured decision making environments. They provide support for decision making so that the effectiveness of decisions can be improved. The system is able to produce the quick summary of data by detecting trends over a period of time. It compares points and patterns related to different variables and perform forecasting functions.

A DSS for better academic should collect records on all instructional procedures, provide feedback for their improvement, and provide choice-making help with excessive integration and direct interaction with all of the domain names of the hassle [1][2][3][4][5][6]. The improvement and integration of a DSS with the university

ICT systems may additionally likely decide a reduced price and time needed to remedy key issues of adopting maximum appropriate decisions, for the representative complexities of better academic systems [7]. Data mining based analysis models can be used to present the advantages of DM technology for the higher educational system to improve the effectiveness of the traditional processes [8]. Current revolution in Intelligent DSS and an integrated IT system will assist the decision makers in taking the right and effective decisions constitutes an important step towards the implementation of the new educational policies [9][10]. The Higher Educational Data Mining System (HEDMS) is concerned with the developing methods that discover useful knowledge from data that extracted from the educational system [11]. The DM based recommendation system supports learners to take suitable decisions in their academic activities. It guides to take the decision on the choice of course based on the experience of previous students with similar academic achievements [12].

A typical educational system with data mining [1][13][15] support requires following major components:

- The RDBMS, data warehouse and Server system which consists of the set of databases, data warehouses, or other types of information repositories containing the student and course information. Database Server supplies the applicable data based on the administrator and user requests.
- The knowledge base that contains metadata and the academic knowledge used to evaluate the interestingness of resulting patterns from various sources.
- The DM engine which consists of algorithms and functionality for tasks such as classification, association, cluster analysis etc.
- The user interface component that provides the user with an interactive graphical approach to the database and data mining system.

This data mining integrated decision support system requirement can be achieved by using data warehousing, OLAP, data mining and EDM (Educational Data Mining) methods for analyzing and reporting into a flexible structural design [14]. It should contain data management level where the process must be applied to clean and load data into a data warehouse, a DM, analysis and model management level with

analytical models where OLAP and DM techniques can be used for academic analysis, a user interface level where reports can be built with academic intelligence tools. To design an educational decision support system, DM techniques and methods are required to be integrated to improve performance and accuracy of analysis for forecasting.

2. OVERVIEW OF DM-HEDSS SYSTEM

This DM based integrated and interactive system is a system of users, hardware, software, procedures, an interactive query facility, data records for collecting, storing, retrieving, analyzing and transmitting information to authorized personnel. Information needs for decision support for different stakeholders in different situations with different forms of visualization and complexity have to be identified [14][16][17[18]]. For a model development and management, it is necessary to identify the boundary of the educational system and ensure the major functionalities of the system identifying, meaning to evaluate the executing tasks and environment, collect and analyze the requirements, and understand the information needed to figure out critical problems.

2.1 Components

The conceptual model of a DM integrated higher educational decision support system called DM-HEDSS is given here with its components in figure 1. Depending on the academic and technical requirements adjusted system can be designed in two tiers by integrating the database and DM system. Here system is divided into following components in 3 tiers for convenience.

- Record Management Component (RMC) includes the large repository of academic and administrative database and data warehouse.
- Academic Analysis Component (AAC) where model management and pattern evaluation functions will take place.
- User Interaction Component (UIC) where knowledge presentation functions available with different customizable graphical analysis & reporting required by users for decision-making.
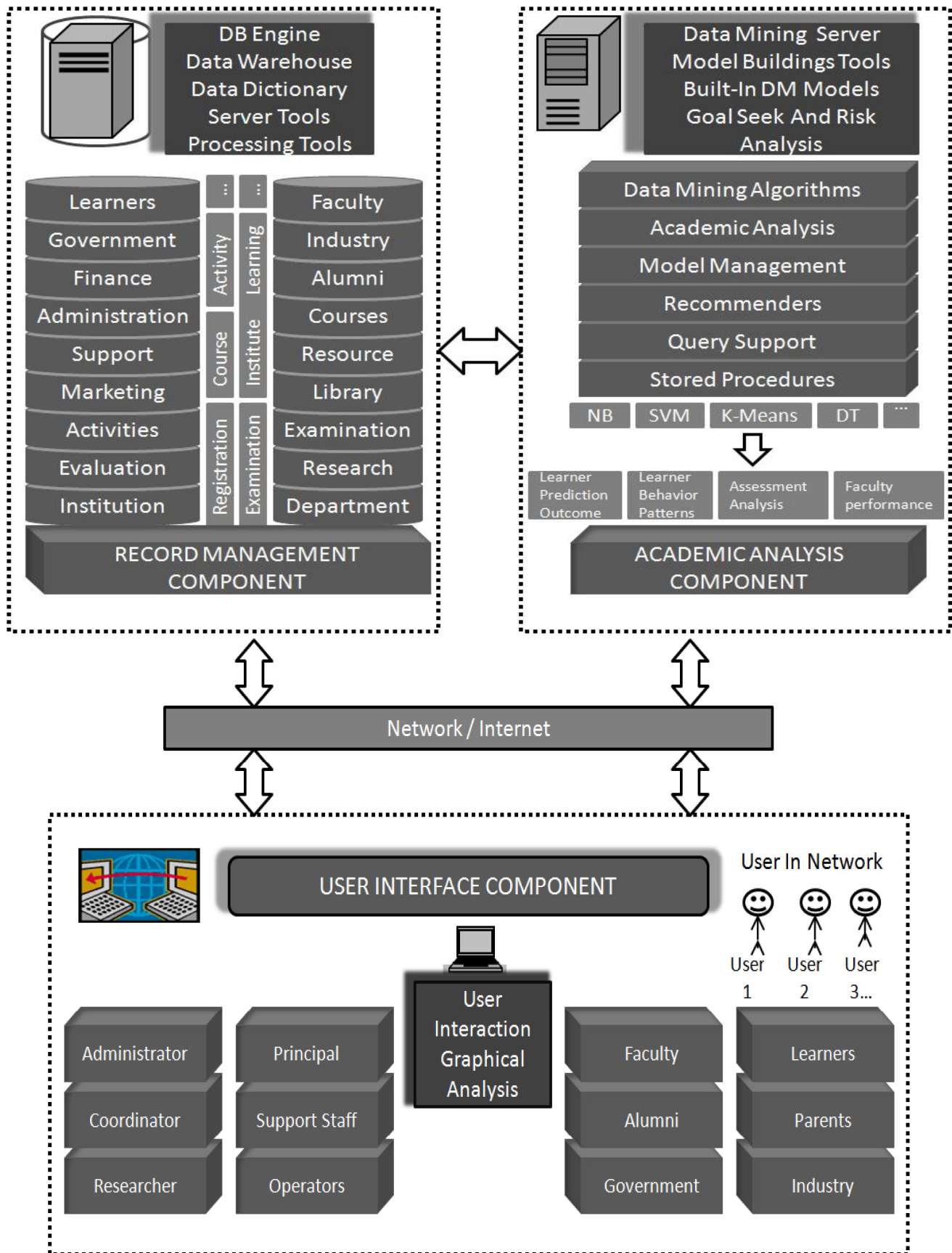
Figure 1: Higher education decision support system

2.2 Data Mining Process

The data mining processes inside DM-HEDSS is viewed as "3P" (Prepare-Produce-Present) phases diagrammatically shown in Fig. 2. Process planning inside DM-HEDSS include six phases similar to typical data mining process [15] used in knowledge discovery in databases (KDD) projects.
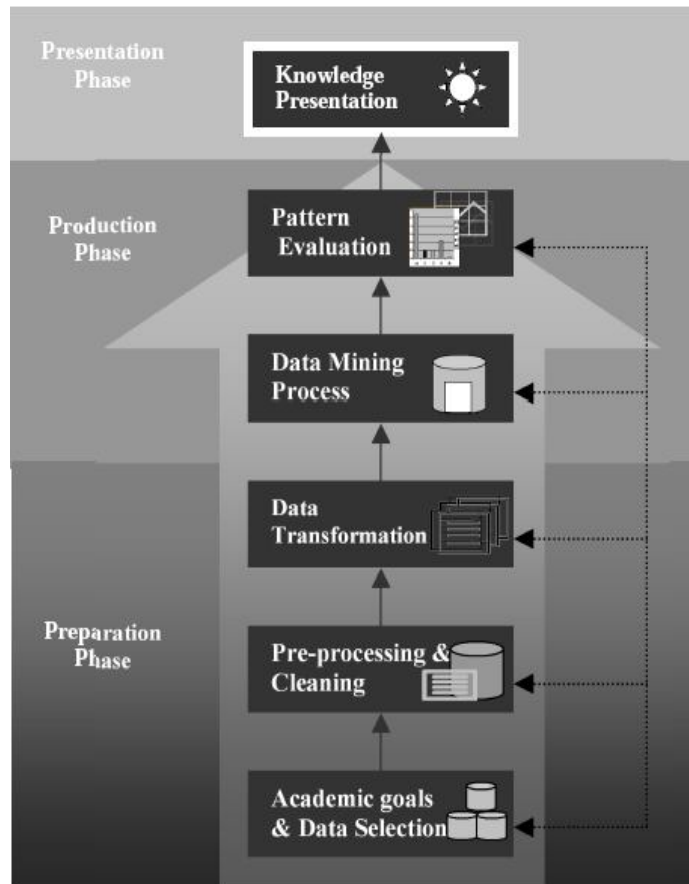

Figure 2: DM and KDD process in DM-HEDSS

- **Understanding of the higher educational academic setting sand data set selection:** Preliminary understanding transformation, representation, algorithms, etc. about the academic settings need to be considered. It is necessary to understand objectives of user, set the boundary in which the knowledge presentation process for decision making will take place. Selecting and creating the data that will be used to produce results will be performed. This step search the available data as per need and then combine all the data into one set without missing required attributes that will be considered for the process.

- **Pre-processing and cleaning:** This step includes the removal of noisy and inconsistent data to enhance the reliability of data. Data cleaning, such as handling missing values and the removal of noise or outliers involves complex statistical methods.

- **Data transformation:** This phase involves the conversion of operational data from one structure to another structure for better understanding and mining results. In this stage, depending upon the project handled, methods include dimension reduction such as feature selection and attribute transformation methods.

- **Data Mining:** It is an essential step of discovering anomalies, patterns and correlations in large amount of data sets. DM is considered as an interdisciplinary field that uses various innovative methods to predict the outcomes. In this step the use of DM algorithms, classification, clustering, regression, etc. will be decided. Basically this step includes the selection of specific methodology for finding task relevant data into pattern.

- **Pattern evaluations:** This step typically employs the searching of specific pattern. It is also used to find the interestingness score of each pattern. All evaluations in this step done to achieve the goal higher educational domain.

- **Knowledge presentation:** This step is the usage of the patterns and discovered results obtained by the DM. The gained knowledge can be including in other system for future use. Visualization and knowledge presentation tools are used to refine the knowledge for the decision making process.

2.3 Record Management

RMC is responsible for storing integrated data from various sources related to the entire leaning system. Primarily, it is composed of the data warehouse, RDBMS and data dictionary. Data entered by users from different systems are integrated here. A database system has ad-hoc query and on-line transaction processing capabilities. RDBMS covers the compilation of data with data generation, data manipulation, data updating and data maintenance. Data set for the process must be updated with the current state database and data is checked for errors. RMC has required data to develop models. It contains major database entities related to learners, learners' extracurricular activities, teaching faculty, non-teaching staff, alumni, government and accreditation, administration and finance, courses and curriculum, marketing information, examination & evaluation, library management, industry and placements, institutional resources (infrastructure details, facilities, learning materials.), projects & academic research information.

2.4 Academic Analysis

AAC supports DM techniques, analysis and visualization techniques. DM Component will facilitate decision making with the use of techniques such as neural networks, decision trees, prediction, clustering methods, etc. This helps for prediction, association tasks and anomaly detection. This

allows a variety of goal seeking and risk management applications concerned with the institution administration. This system support queries and produce recommendations related to course and learning activities. DM methods and sample academic analysis tasks are shown in Table 1. At the AAC level model builder facility is used to create, identify process, store, update and maintain various decision making models in education. It contains model dictionary for definition and use of different models. DM practitioners will handle data modeling and evaluation of analytic models and implement different these models based of academic requirements. In this level model building algorithms are used to create different analytic models and evaluation is performed for higher quality and refinements will be made at times. This level processed operational data and obtain the necessary information for assumptions and taking decisions. System supports the building of models as per analysis requirements of users. The stored data is analyzed using appropriate analysis tools to identify pattern sand behaviors. The obtained knowledge is stored in the data repository.

This model Building & management segment focuses on using DM algorithms and the specifications to build models. Output validation & interpretation is done based on the values measured are compared to the expected output, accepted difference and errors are set based on comparative analysis and why these differences have arisen. The result obtained after applying DM techniques will be used by specific stored procedures and the performance of the model is examined. These algorithms ensure data transformed into knowledge to support effective academic analysis. This component allows statistical analysis, querying, reporting, visualizing support and identifies critical success factors. Data is transformed into knowledge using artificial intelligence techniques or statistical analysis techniques in DM algorithms which allows the discovery of knowledge, correlations and rules to support decision making. Table 1 shows the Data mining methods and related analysis task in academic environment.

Table 1: DM methods and sample analysis tasks in academics

DM Method	Sample Analysis tasks
Prediction	Prediction of alumni donations
	Prediction of learner's placement in Industries
	Predicting active and inactive alumni
	Academic time table planning
	Predict academic success
	Predict student drop out chances
	Predict Change of major and optional subject
	Predict faculty performance based on student results
	Predict success rate of a course and subject considering cost and performance issues
	Factors affecting success level of learners
	Predicting learner behavior
	Predict learners who are likely to take particular optional subject and sport activity.
	Prediction of learners problem solving behavior
	Predict institution accreditation level and rank based on learner outcome
	Predict learner performance in a group
	Predict proficiency of learners in different subjects and fields
Clustering	Learners cluster based on academic success levels in different perspectives
	Cluster similar students in different perspectives
	Cluster students based on characteristics
Association	Associate learners demographic data with final academic outcomes
	Associate student assignment score, attendance with final result
	Associate learner with different academic subjects
	Associate learner attitude with their results
	Associate teaching faculty with learners examination score
	Associate lecturer performance with their personal data
	Associate course work, subject, curriculum different students
	Associate students behavior with performance inassessments
	Associate Assessment difficulty level with student grade
	Associate learners' skills with present industry requirement
	Associate student or group with a particular project or activity
Classify	Classify students based on academic success level
	Classify students based on their course topic interest
	Classify students based on their assignment and attendance performance
Outlier Analysis	Detecting abnormal behaviors in students
	Detecting malpractice in different assessments
	Detect students with specific learning difficulties

2.5 User Interface

User Interface Component has Graphical User Interface (GUI) that allows users to communicate with system and other electronic devices by using graphical icons and indicators. It is very helpful to observe information for the educational process. The GUI component acts as a user interaction layer which facilitates communication between system and user and helps to enter and extract useful information for the academic process. This component provides web based graphical menus to access the functionality of the system. This system allows administrators or persons who are involved in the decision making process to quickly understand and visualize the impact of various courses of action using graphical analysis. This helps the user to seek a quick summary of data, trend detection and forecasting activities. It supports learners, teaching & non-teaching staff, alumni groups, Government and accreditation officials, placement industry officers, academic researchers and other authorized users. Different users with basic knowledge of DM concepts will be able to use the system at UIC level.

User's interaction is done at GUI level and it helps operators and educators to communicate with the system to generate the strategic information for analyzing the results. Access via Web based system works with browser supporting interactive query and executive information reporting. Reports involve the summary and interpretation of information about learners' learning activity and academic management related information to the administrators. It allows top executives of institution display results produced by AAC in this user friendly system. Machines in the network with different authorization levels will easily interact with database server. It provides excellent user interface and graphical analysis facilitating the user interactions with the system. Data generated from the operational data stores and data mart are collected and arrange easily by using application programs in a user interface component. This level is composed of a set of queries, customized data reporting tools, customizable analysis tools, and graphical analysis interface. All components of DM-HEDSS support the use of different telecommunication equipment, network devices and mobile platforms.

3. CONCLUSION

The DM-HEDSS is useful for educational decision making for academic administrators. This proposed system design enables the creation of DM models based on present data for academic analysis so that users who have basic DM skills will be able to use at the User Interface Component level. DM models defined various patterns and rules that are used by academic administrator as the new knowledge and very helpful in effective analysis and decision making. Educators may adopt some changes in counseling approach and learning methodologies based on the knowledge acquired in DM

models. Entire system development requires a team work for implementation of complex algorithms. At present, as an illustration of part of the development process of DM-HEDSS model, a sample work involving a learner and examination database is implemented to study the effectiveness of different types of educational assessments for an undergraduate course in commerce and business administration students. For future work, we hope to develop the model in order to achieve valuable and accurate output useful for teaching faculty to improve the students' learning outcomes by valuating their E-Learning and cognitive abilities.

REFERENCES

[1] Rok Rupnik, Matjaz Kukar, Marko Bajec, Marjan Krisper, DMDSS :Data mining based decision support system to integrate data mining and decision support, 28th Int. Conf. Information Technology Interfaces ITI, 2004.

[2] Vasile Paul Bresfelean, NicolaeGhisoiu, Ramona Lacurezeanu, Dan-Andrei Sitar-Taut, Towards the Development of Decision Support in Academic Environments, Proceedings of 31st Int. Conf. on InformationTechnology Interfaces, 2009.

[3] Ayesha A., Mustafa T. and Khan M. I., 2010, "Data Mining Model for Higher Education System". European Journal of Scientific Research, Vol. 43, No.1, pp.24-29.

[4] Sonali A., Pandey G. and Tiwari M., 2010. "Data Mining In Education: Data Classification and Decision Tree Approach" Vol. 11, No. 5.

[5] Mohammad H. and Jafar H, September, 2011. "Using Educational Data Mining Methods to Study the Impact of Virtual Classroom in E-Learning", IJDKP.

[6] Edin Osmanbegović and Mirza. Suljic, "DATA MINING APPROACH FOR PREDICTING STUDENT PERFORMANCE", *Journal of Economics & Business/Economic Review*, vol. 10, pp. 3-12, 2012.

[7] Mostow, J., Beck, J., Cen, H., Cuneo, A., Gouvea, E., and Heiner, C., 2013, "An Educational Data Mining Tool to Browse Tutor-Student Interactions: Time Will tell",in: Proceedings of the Workshop on Educational Data Mining, pp.15–22.

[8] ElenaSusnea, 5th International Conference EDU-WORLD 2012 -Education Facing Contemporary World Issues, Improving DecisionMaking Process in Universities: A Conceptual Model of IntelligentDecision Support System, Procedia - Social and Behavioral Sciences (Science direct) 795 – 800, 2013.

[9] AymanKhedr, Ahmed I. El Seddawy, A Proposed Data MiningFramework for Higher Education System International Journal ofComputer Applications, 2015.

[10] M. K. Behzad Nakhkob, "Predicted Increase Enrollment in Higher Education Using Neural Networks and Data Mining Techniques", 2016.

[11] Anoopkumar M and A. M. J. M. Z. Rahman, "A Review on Data Mining techniques and factors used in Educational Data Mining to predict student amelioration," 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE), Ernakulam, 2016, pp. 122-133, doi: 10.1109/SAPIENCE.2016.7684113.

[12] R. Asif, A. Merceron, S. Ali and N. Haider, "Analyzing undergraduate students' performance using educational data mining", *Computers & Education*, vol. 113, no. Supplement C, pp. 177-194, 2017.

[13] K. Venkatachalapathy, V. Vijayalakshmi and V. Ohmprakash, "Educational Data Mining Tools: A Survey from 2001 to 2016," 2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM), Tindivanam, 2017, pp. 67-72, doi: 10.1109/ICRTCCM.2017.53.

[14] Eka Sugiyarti, Kamarul Azmi Jasmi, Bushrah Basiron, Miftachul Huda and Andino. Maseleno, "Decision Support System of Scholarship Grantee Selection using Data Mining", *International Journal of Pure and Applied Mathematics*, vol. 119, 2018.

[15] C. C. Ukwuoma, C. Bo, I. A. Chikwendu and E. Bondzie-Selby, "Performance Analysis of Students Based on Data Mining Techniques: A Literature Review," 2019 4th Technology Innovation Management and Engineering Science International Conference (TIMES-iCON), Bangkok, Thailand, 2019, pp. 1-5, doi: 10.1109/TIMES-iCON47539.2019.9024396.

[16] S. Gupta, A. Jain, P. Jeswani, "Generalized Method to Produce Balanced Structures Through k-means Objective Function", IEEE 2nd Int. Conf. on I-SMAC organized by SCAD Institute of Tech. at Palladam, Tirupur, pp. 586-590, 2018, DOI:10.1109/I-SMAC.2018.8653668.

[17] A. Jain, "Clustering of Text Streams via Facility Location and Spherical K-means", 2nd IEEE Int. Conf. on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, pp. 1209-1213, 2018, DOI: 10.1109/ICECA.2018.8474757.

[18] R. Islam, M. T. Sazid, S. R. Mahmud, C. N. Ferdous, R. Reza and S. A. Hossain, "Parametric Study of Student Learning in IT Using Data Mining to Improve Academic Performance," 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Spokane, WA, USA, 2019, pp. 286-290, doi: 10.1109/ICIEV.2019.8858551.