iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** IV **Month of publication:** April 2018

DOI: http://doi.org/10.22214/ijraset.2018.4142

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

An Efficient Education System to Enhance Student Employability through Mining of Academics and Event Performance

Rashmi M R[1], Reshitha R[2], Shubha R[3], Tejaswini Y S[4], Vidya V[5]
[1] *Rashmi M R, Assistant Professor, Department of Computer Science and Engineering, The National Institute of Engineering, Mysuru, Karnataka*
[2]*Reshitha R, Department of Computer Science and Engineering, The National Institute of Engineering, Mysuru, Karnataka*
[3]*Shubha R, Department of Computer Science and Engineering, The National Institute of Engineering, Mysuru, Karnataka*
[4]*Tejaswini Y S, Department of Computer Science and Engineering, The National Institute of Engineering, Mysuru, Karnataka*
[5]*Vidya V, Department of Computer Science and Engineering, The National Institute of Engineering, Mysuru, Karnataka (*
[1., 2, 3, 4, 5] *Department of Computer Science and Engineering The National Institute of Engineering Mysuru*

Abstract: ***Education system in the current world focuses on multiple dimensions required for overall student development. Ensuring an ability to acquire a job opportunity is the basis of training given at an undergraduate and graduate level. Academic abilities are not alone important when it comes for placement, their co-curricular and extra-curricular trainings also plays a vital role. The pattern of behavior that exists in the student community is that they have tendency to attend program where a lot of people participate, and it is popular. Nobody cares whether the program suits their caliber or is there any benefit for them in attending such programs. The research here concentrates in developing an analytical model which helps in predicting the student community the most suitable program they can attend based on their knowledge level and also provides the rating of the various events in an educational institute. The basic step taken is to categorize the students and events based on various criteria and parameters using a Decision tree algorithm. Naïve Bayesian classification is applied on top the dataset in order to make a classification of event based on student grade and outcome they gained by attending an event. Overall system does the analysis of students and events and creates an analytical model for student community to choose the most beneficial programs that they should attend which enhances their chances of getting placed.***
Keywords: educational data mining, decision support system, event analysis, Naïve Bayesian classifier.

I. INTRODUCTION

Mining is the process of discovering patterns in large datasets involving methods at the intersection of machine learning, statistics and database systems. It is an essential process where intelligent methods are applied to extract data patterns. The main goal of data mining process is to extract information from a dataset and transform it into an understandable structure for future use. It involves database and data management aspects, data preprocessing, model and inference considerations, complexity considerations, visualization. [1] A unique pattern is used to analyze student academic performance and behavior. Data is obtained from the school records. Classification, training and testing is done on the basis of two categories, pass or fail. The research helps in knowing the students' interest in different subjects which helps in increasing success rate and decision making. [2] The research focuses on identifying attributes which are responsible for the

loss in student academic performance. The data from academic records and socio-economic data helps in finding the attributes responsible at the first four enrollments. [3] The research focuses on K-12 students' low academic performance. A generic framework is developed on various aspects like subjects, topics within subjects, concepts within topics to predict performance. It not only helps students but also teachers in getting prepared to change their teaching ways guiding and focusing on particular student. [4] Traditional and online offerings help students personally to improve and make timely decisions. Online videos, studying course material help in determining the attrition rate. The data collected helps in predicting student performance and also make student confident of selecting the correct course. [5] The research aims at predicting the at-risk online learning students at the early stage and even when they do not seek assistance. The data can be used to amend instructors' understanding of the relationship between specific learning behaviors at specific times and their students' performance.

II. LITERATURE REVIEW

[6] An experiment was conducted on 700 students with 19 attributes in Amritha Vishwa Vidyapeetham, Mysuru where student academic history was taken as input and students upcoming performances was predicted on the basis of semester. Some attributes like student attendance in class, hours spent on daily basis after college, family income and mother's education showed high co-relation with the performance of the student. [7] A student's yearly performance is predicted in the form of Cumulative Grade Point Average (CGPA) using supervised neural network which digs out information from student's data. The experimental results show that the academic performance largely depends on family background and student's engagements with social media interactions. [8] Characterization and predictive models signifying the student dropouts helping in determining the probability of dropouts.

[9] Research analysis was carried out for 700 UCLA undergraduate students who have opted for digital signal processing course where results obtained from pilot course suggested that early in-class assessment results in timely performance prediction of each student, thus enabling timely interventions by the instructor. The obtained grades showed how well a student can understand and apply the knowledge conveyed in a course. [10] The results obtained from repeated iterations is evaluated on what is expected to what is obtained. Factors like population studied and specific needs manifested by institution are considered for development for predictive model. The research benefits students in learning, making timely decisions to prevent desertion.

[11] The research was conducted on data collected from 307 students of 3rd year CS (Computer Science) under graduates, India. Online learning skills, day scholars, time management, sports participation are some of the attributes considered which determines behavior, educational environment and interest of students. The prediction helps in enhancing student academic status.

The research includes a project-based learning activity to ensure that engineering students are motivated to work on multinational projects. The need for the collaborative work among engineers is calling for changes in educational environment to educate engineers to become globally competitive and to operate in interdisciplinary design project teams. [13] Students are classified into four types using cluster analysis. The research addresses how we can take advantage of what differentiates students in a gamified learning environment to predict their behavior early in the term. Some gamification elements like experience points, challenges, badges and leader boards encourage students to adopt behaviors that help them learn better. Such students can perform better in practical assignments and can have higher overall score. [14] The research includes a customized online tutoring system for each student for analyzing student performance where a tutor increases support when a student makes a mistake or reduces the amount of help when the student makes progress, thus helps to prevent student retention.

III. PROPOSED SYSTEM FOR EVENT ANALYSIS AND STUDENT PREDICTION

Study of existing literature provides information that there is no major focus given on the role of co-circular and extra- circular activities in equipping a student job ready.

The research concentrates on developing an inference engine capable of predicting co-circular and extra-circular activities which are essential for a student.

The initial path begins with a data collection through a survey conducted at The National Institute of Engineering. The inference engine, the vital functionality of the system performs a set of analytical operations to categorize students, events and correlate matching factors which are important in prediction. The system also efficiently helps the trainers in conducting programs which are beneficial to student community. Identification of target audience gives a real time use of the research in any academic institution.

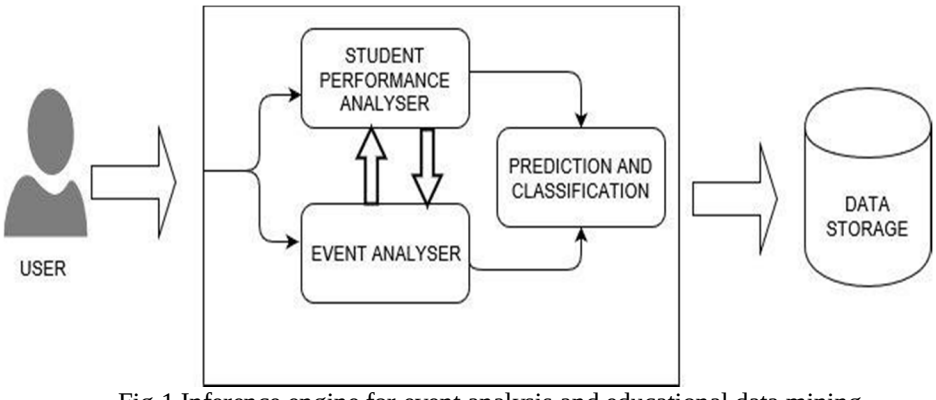


Fig 1 Inference engine for event analysis and educational data mining

Fig 1, the architecture diagram explains the various functional modules that integrate together into an efficient system. The user inputs are taken in the form of a survey and fed into the inference engine. Inference engine is constituted of event analyzer, student performance analyzer and prediction system. Prediction system takes inputs from both the analyzer and classifies the events based on student's category.

IV. ALGORITHMS

A. *Decision Tree Algorithm*

1) *Input:* Data partition, D, which is a set of training tuples and their associated class labels.
2) *Step 1 :* create a node N;
3) *Step 2 :* if tuples in D are all of the same class, C then return N as leaf node labeled with class C;
4) *Step 3 :* if event|Liststudent List is empty then return N as leaf node with labeled with majority class in D;
5) *Step4 :* apply attribute_selection_method (D, eventList | studentList) to find the best splitting_criterion; label node N with splitting_criterion;
6) *Step 5 :* if splitting_attribute is discrete-valued and multiway splits allowed then attribute_list = splitting attribute; remove splitting attributefor each outcome j of splitting criterio
7) Step 6 : let Dj be the set of data tuples in D satisfying outcome j; // a partition
8) *Step 7 :* if Dj is empty then attach a leaf labeled with the majority class in D to node N; else attach the node returned by Generate decision tree (Dj, attribute list) to node N; End for return N;

B. *Naïve Bayesian Algorithm*

To classify the event according to the student performance and benefits of being part of an event a Naïve Bayesian algorithm is used Naïve Bayesian takes the student academic score and event rating done by student to classify the event in to various levels according to student benefits. Naïve Bayesian is a conditional probability model: given a problem instance to be classified, represented by a vector $\mathbf{x} = (x_1, \ldots, x_n)$ representing some n features (independent variables), it assigns to this instance probabilities for each of
K possible outcomes or classes. p(Ck|x1,.....,xn)p(Ck|x1,.....,xn)

The problem with the above formulation is that if the number of features n is large or if a feature can take on a large number of values, then basing such a model on probability tables is infeasible. We therefore reformulate the model to make it simpler. Using Bayes theorem, the conditional
probability can be decomposed as − p(Ck|x)=p(Ck)p(x|Ck)p(x)p(Ck|x)=p(Ck)p(x|Ck)p(x)

This means that under the above independence assumptions, the conditional distribution over the class variable C is − p(Ck|x1,.....,xn)=1Zp(Ck)∏i=1np(xi|Ck)p(Ck|x1,.....,xn)=1Zp(Ck)∏i=1np(xi|Ck).

V. RESULTS AND DISCUSSION

	A	B	C	D	E	F	G
1	pucgrade	CGPA	Grade	Name	eventtype	eventname	rating
2	1	2	2	Vidya.V	Science/Technical fest	TEQIP - Introduction to network security and cloud computing	3
3	1	1	1	Tejaswini.Y.S	Hands-on-session/Workshops	Current trends in digital image processing techniques and applications	5
4	2	2	2	Suhas.k	Science/Technical fest	Operation and programming of kuka robot	3
5	1	2	2	Shubha.R	Conference/Seminar	Tech talk on 3D hologram technology	4
6	1	1	1	Reshitha.R	Hands-on-session/Workshops	TEQIP - Introduction to network security and cloud computing	1
7	1	1	1	Rakesh.T	Hands-on-session/Workshops	Operation and programming of kuka robot	4
8	2	3	3	Lavanya R	Science/Technical fest	Idea hackathon	3
9	3	4	3	Aishwarya P	Cultural events	Technieks - logo quiz	3
10	4	4	4	Sahana R	Hands-on-session/Workshops	Weekend statup	3

Fig 2 Dataset

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

Fig 2 shows the data collected from 400 students at The National Institute of Engineering. It includes the parameters like pre university grade (PUC), cumulative grade point average (CGPA), name, event type, event name and ratings.

The pre-processed data is fed into a decision tree based analyzer which considers CGPA, PUC marks obtained to formulate a score on which the data can be trained and students can be divided into various classes.

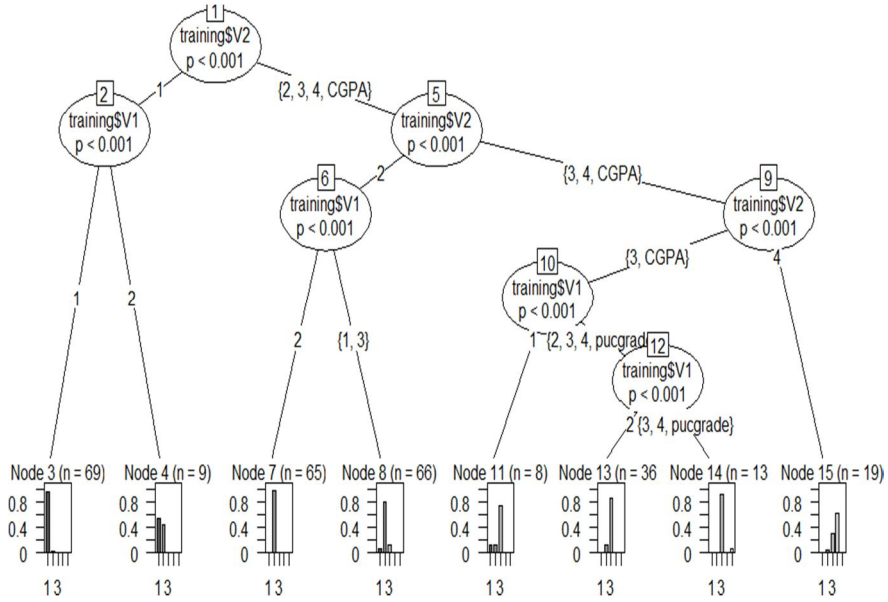


Fig 3 Student Performance

The decision tree algorithm also considers event parameters that are event type, event rating and outcome of the event. The algorithm classifies the event into various classes and generates the decision tree.

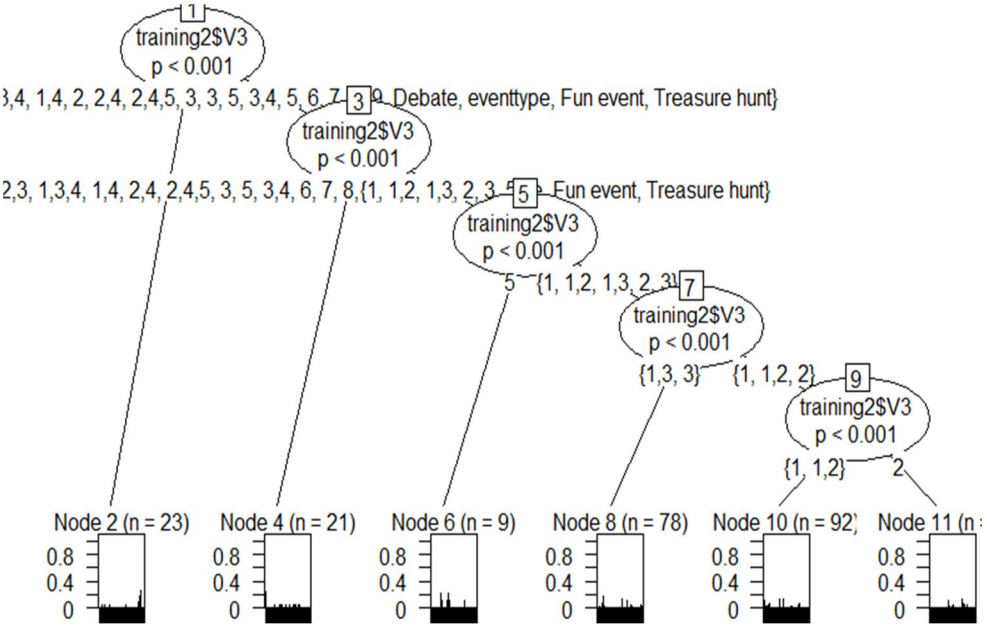


Fig 4 Event classifier

In order to correlate the event outcomes with various student scores, a Naive Bayesian classifier is deployed. It uses the conditional probability model and segregates the correlated data into classes. Even it has a prediction, helping student choose an event based on CGPA. The predictor analyses distribution between event and student class and is shown below.

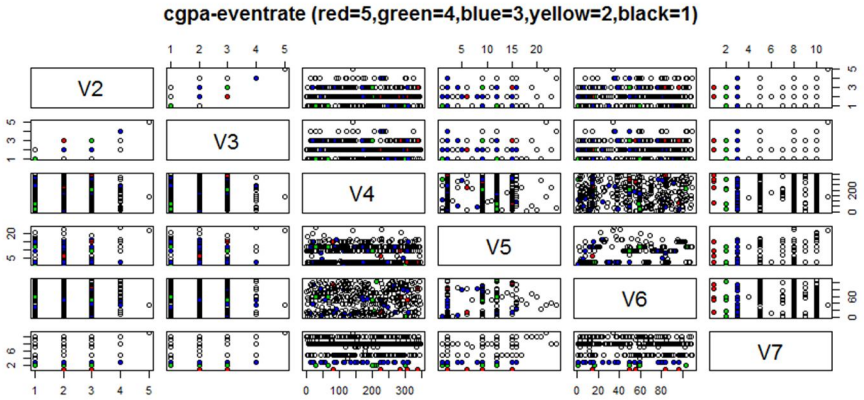


Fig 5 Correlation

VI. CONCLUSION AND FUTURE WORK

In order to help the student to choose an efficient program based on observations collected from CGPA and event attending pattern. The system is both useful for the student as well as for the trainers in conducting and participating for various extra-circular and co-circular activities.

The system can be enhanced with a mobile based application in future which lists the events, their ratings and gives suggestion for the registered students.

REFERENCES

[1] S. S. Athani, S. A. Kodli, M. N. Banavasi, and P. G. S. Hiremath, -Student academic performance and social behavior predictor using data mining techniques,‖ in2017 International Conference on Computing, Communication and Automation (ICCCA), 2017, pp. 170–174.

[2] C. E. Lopez Guarin, E. L. Guzman, and F. A. Gonzalez, —A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining,‖ Rev. Iberoam. Tecnol. del Aprendiz., vol. 10, no. 3, pp. 119–125, 2015.

[3] S. Ikbal, A. Tamhane, B. Sengupta, M. Chetlur, S. Ghosh, and J. Appleton, -On early prediction of risks in academic performance for students,‖ IBM J. Res. Dev., vol. 59, no. 6, 2015.

[4] A. Elbadrawy, A. Polyzou, Z. Ren, M. Sweeney, G. Karypis, and H. Rangwala, -Predicting Student Performance Using Personalized Analytics,‖ Computer (Long. Beach. Calif)., vol. 49, no. 4, pp. 61– 69, 2016.

[5] J.-L. Hung, M. C. Wang, S. Wang, M. Abdelrasoul, Y. Li, and W. He, -Identifying At-Risk Students for Early Interventions—A Time-Series Clustering Approach,‖ IEEE Trans. Emerg. Top. Comput., vol. 5, no. 1, pp. 45–55, 2017.

[6] Devasia, Vinushree T P, and V. Hegde, -Prediction of students performance using Educational Data Mining,‖ in 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE), 2016, pp. 91–95.

[7] M. F. Sikder, M. J. Uddin, S. Halder, and Ieee,
-Predicting Students Yearly Performance using Neural Network: A Case Study of BSMRSTU,‖ 2016 5th Int. Conf. Informatics, Electron. Vis., pp. 524–529, 2016.

[8] Y. Amaya, E. Barrientos, and D. Heredia, -Student Dropout Predictive Model Using Data Mining Techniques,‖ IEEE Lat. Am. Trans., vol. 13, no. 9,pp. 3127–3134, 2015.

[9] Y. Meier, J. Xu, O. Atan, and M. Van Der Schaar,
-Predicting grades,‖ IEEE Trans. Signal Process., vol. 64, no. 4, pp. 959–972, 2016.

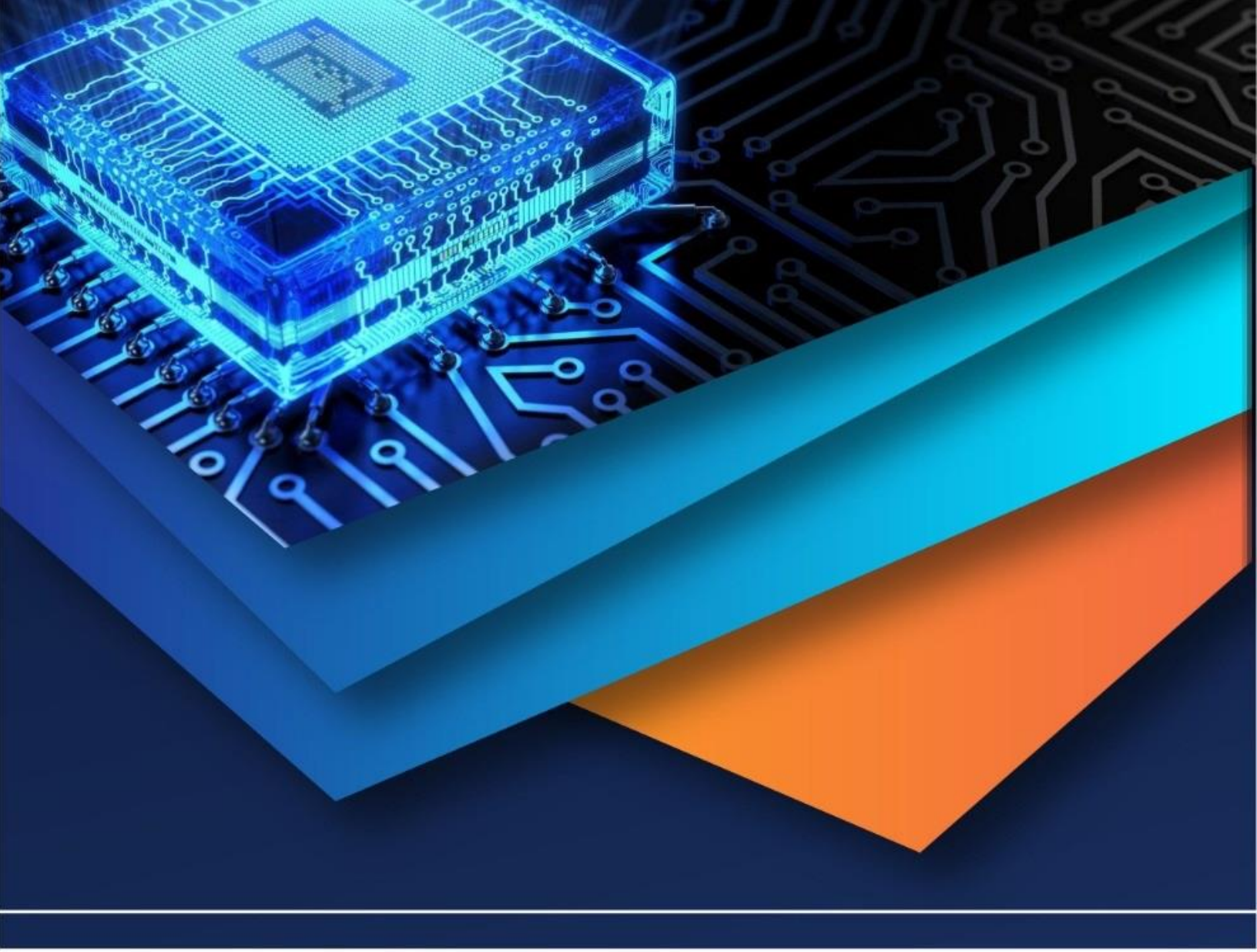
[10] S. M. Merchán and J. A. Duarte, —Analysis of Data Mining Techniques for Constructing a Predictive Model for Academic Performance,‖ IEEE Lat. Am. Trans., vol. 14, no. 6, pp. 2783–2788, 2016.

[11] S. Venkatramaphanikumar, S. C. Kumar, E. D. Chowdary, and K. V. K. Kishore, -M5P model tree in predicting student performance: A case study,‖ in 2016 IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, RTEICT 2016 - Proceedings, 2017.

[12] I. E. Esparragoza et al., -Study on interest and perception of value in multinational collaborative design projects among engineering students,‖ in Advances in Intelligent Systems and Computing, 2018, vol. 627, pp. 16–29.

[13] G. Barata, S. Gama, J. Jorge, and D. Goncalves,
-Early Prediction of Student Profiles Based on Performance and Gaming Preferences,‖ IEEE Trans. Learn. Technol., vol. 9, no. 3, pp. 272–284, 2016.

[14] D. Riofrio-Luzcando, J. Ramirez, and M. Berrocal- Lobo, -Predicting Student Actions in a Procedural Training Environment,‖ IEEE Trans. Learn. Technol., pp. 1–1, 2017.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY