iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** V **Month of publication:** May 2021

DOI: https://doi.org/10.22214/ijraset.2021.34173

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: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

Application of Machine Learning for Higher Education

Sampada Taralgatti[1], Aparna Nair[2], Tejaswini Sawant[3], Priya Parate[4]
[1, 2, 3, 4]*Student, Department of Computer Engineering, Rajiv Gandhi Institute of Technology, Mumbai, India*

Abstract: ***In recent years, the demand for higher education has increased drastically due to evolution in technology, student demographics, increased competition and employer demands. However, the future of higher education relies on the ability to become more adaptable, flexible and graceful in the way they approach the development and learning design. So the most common issue faced by the University students is to take the right decision in relation to their academic path based on available information for e.g.: courses, schedules etc. In this paper, we present a model for selection of the correct study program for higher education students. The main idea is an early analysis so that student's successes on each study program and recommendation of a study program where a student will likely to be succeeded. So software could be utilized to predict and recommend relevant universities.***

Keywords: Machine **Learning, Prediction, Recommendation, Linear regression, Tensor Flow.**

I. INTRODUCTION

The future of upper education is intrinsically connected with upcoming new technologies and calculating capacities of the new intelligent machines. University educations are an important part of most people's preparation for working life. An admission to university is therefore an important challenging topic. Hence, effective university admission prediction and recommendation services are needed for helping students to enter the right university. However, due to the huge number of students requiring to attend the university every year, this decision making process became a very complex problem. Since, this process is not merely relying on student's test score but also their background profile and other qualification weighing criteria. There are several online portals that provide admission related statistics, but do not provide information regarding individual profile data, thus, leaving the student with the only option of guessing and hoping for the admission or rejection. Universities on the other hand, also receive thousands of applications each year. Handling this process manually could be monotonous task and could also lead to errors. Since the numbers of applicants have drastically increased in the past years, there is a need to deploy a more effective and efficient automated method to handle the admission process. The aim of this project is to construct trusted software based on machine learning algorithms and knowledge discovery rules to achieve the enrolment of student's admission in the universities impartially, according to the standard criteria of universities. This criteria includes Undergraduate GPA (CGPA), which is out of 10; Graduate Record Exams (GRE) score-, the score will be out of 340 points; Test of English as foreign language (TOFEL) score, which will be out of 120 points; Statement of purpose (SOP), which is a document written to show a candidate's life, ambitions and the motivations for the chosen university, the score will be out of 5 points; Letter of recommendation strength (LOR) which verifies candidate's professional experience, builds creditability, boots confidence and ensure your competency, the score is out of 5 points; research paper publication and previous internships. The system also helps university decision makers to provide needed facilities and resources that are required for achieving highly qualified education. The software is adaptable and reliable, since it allocates students according to their profile entered, hence, achieving students satisfaction. In this project we focused on modelling admissions in United States' Universities for higher education. The project provides answers to some of the questions for students: a) what are my chances of admission in the universities? b) What are the universities I will be admitted in? c) What are the top 10 universities? Our goal of the models is to solve the issue faced by the students and the universities by developing the machine learning models.

II. OBJECTIVE

The main objective of the software is to perform prediction of chances of admission and recommend universities according to the student profile which will help students solve their issue and even universities can develop machine learning models based on previous years admits/rejects of students.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

III. COMPARATIVE ANALYSIS

A. Linear Regression

Linear regression is one of the simplest and easiest machine learning algorithms. It is a statistical procedure which does predictive analysis for continuous/real or numeric variables. Since this regression shows the linear relationship, which suggests, it finds how the worth of the variable is changing with the worth of the experimental variable. Mathematically, linear regression is represented as:

1) $y= a_0+a_1x+ \varepsilon$
2) Y= Dependent Variable (Target Variable)
3) X= Independent Variable (predictor Variable)
4) a0= intercept of the line (Gives an additional degree of freedom)
5) a1 = Linear regression coefficient (scale factor to each input value).
6) ε = random error
7) The values for x and y variables are training datasets for Linear Regression model representation.

B. Bayesian Ridge

Bayesian regression allows a natural mechanism to survive insufficient data or poorly distributed data by formulating linear regression using probability distributors rather than point estimates. The output or y is assumed to be drawn from a probability distribution instead of approximate value as one.
Mathematically, to obtain a complete probabilistic model the response y is assumed to be Gaussian distributed around Xw*X* as follows:
$p(y| X,w,\alpha)=N(y| Xw,\alpha)$

C. ARD Regression

ARD is Automatic Relevance Determination Regression, which is an efficient tool for pruning large numbers of irrelevant features resulting in a sparse explanatory subset. However, popular upgraded rules used for ARD are either difficult to increase to more general problems of interest or are characterized by non-ideal convergence properties.
To derive it as lightly as possible the mathematical basics of an easy linear model is repeated because of the idea of regularization to stop over-fitting.
Based on this idea, the Bayesian Ridge Regression is introduced. Generalizing the concept of Bayesian Ridge Regression even more gets us eventually to the idea behind ARD.

D. Theil-sen Regressor

Theil–Sen estimator may be a method for powerfully fitting a line to sample points within the plane (simple linear regression) by selecting the median of the slopes of all lines through pairs of points.
Theil–Sen estimator of a set of two-dimensional points (xi,yi) is the median m of the slopes (yj − yi)/(xj − xi) determined by all pairs of sample points. Once the slope m has been discovered, one may determine a line from the sample points by positioning the y-intercept b to be the median of the values yi − mxi. The fit line is then the road y = mx + b with coefficients m and b in slope–intercept form.
After implementation of the above algorithms it was observed that linear regression algorithm gave 94.742% accuracy and bayesian ridge gave 94.4% accuracy, whereas ARD regression gave 93.6% and theilsen regressor gave 93.3%. Thus, the analysis of algorithms made it clears to approach with linear regression algorithm for the prediction of chances of admission.

IV. METHODOLOGY

The software flow begins with registration. All the students need to register themselves by entering the required details. Then they are required to enter the parameters according to their profile and then accordingly the chances of admission are predicted and universities are recommended according to their profile. Also, they get to view the list of top 10 colleges.

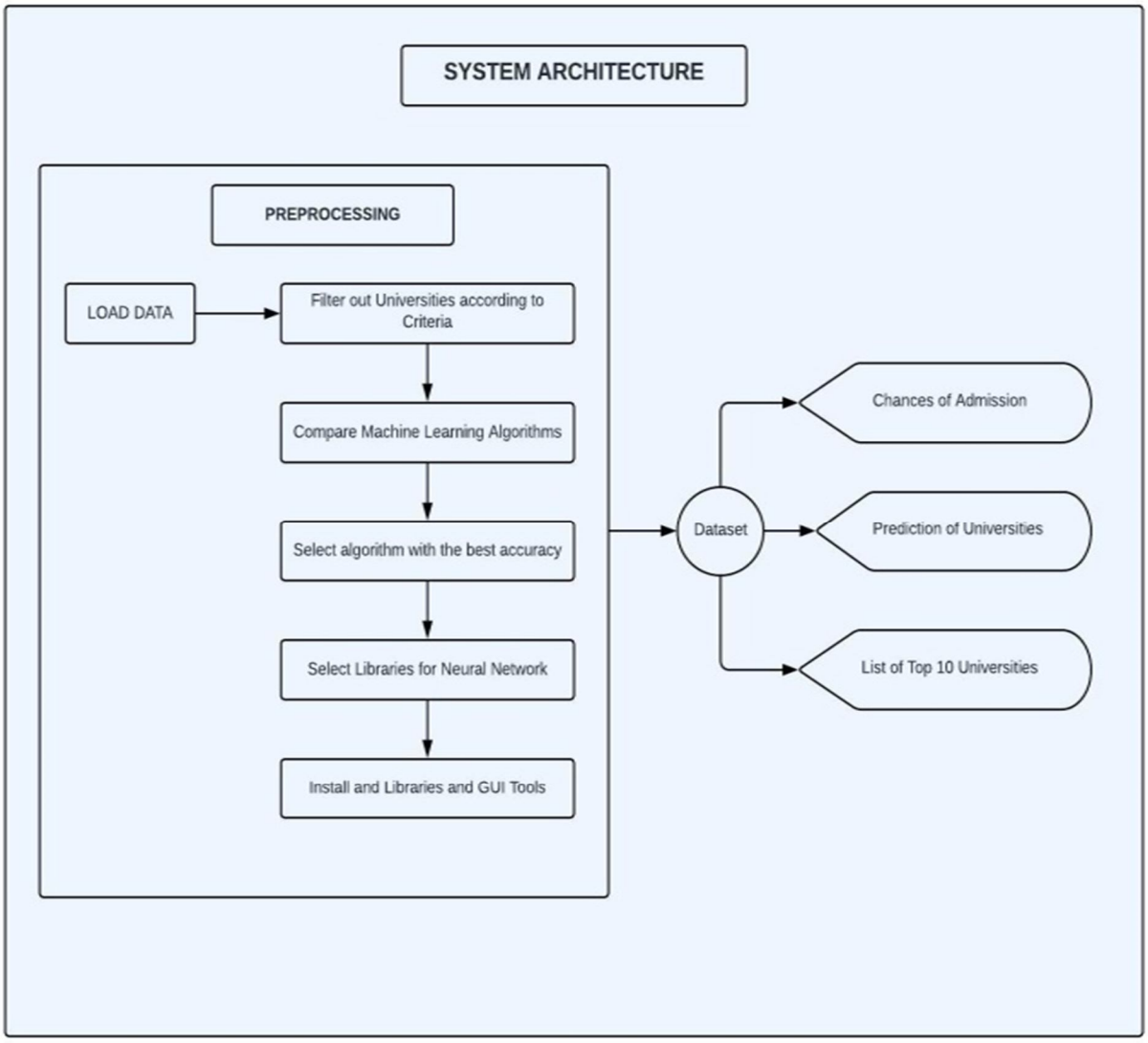


Figure 1: System Architecture

The whole system process can be divided into 6 parts:

A. Student Registration

A student new to the software has to register by entering the required details. These details include username and a password to set. This student registration page is build with the help of tkinter, a graphical user interface tool for python. After the successful registration, the student is taken to the main page of the software.

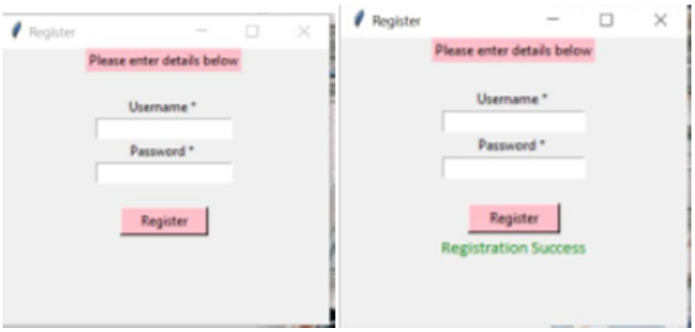


Figure 4: Register Figure 5: Registration success

B. Student Login

An existing student to the software has to login by entering the required details. These details include username and password. This student registration page is build with the help of tkinter, a graphical user interface tool for python. After the successful registration, the student is taken to the main page of the software.

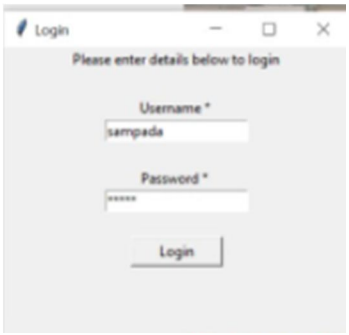


Figure 6: Login

C. Input Details

After successful registration and login of the software, the student is taken to the main page where in a student has to enter his/her details which includes GRE score, TOFEL score, SOP, LOR, CGPA, research paper publication and internship opportunity.

Figure 7: Input Details

D. Chances of Admission

On entering the score, our model performs to find linear relationship between the target and the predictors; it does this using linear regression algorithm which is a way to model a relationship between two sets of variables. Undergoing this process the chances of admission is predicted for the particular student profile.

Figure 8: Chances Of Admission

E. Prediction Of Universities

On entering the score, our model undergoes relu activation function for input layer which is easier train and achieves better performance and sigmoid function for output layer which is used to get multiple true values. Performing this process the universities are predicted according to the student profile.

Figure 9: Prediction Of Universities

F. List of top 10 Universities

The top 10 universities button is provided just so that students could know that what the top 10 universities are in general. We display them by using tkinter, a graphical user interface tool for python.

Figure 10: Top 10 universities

V. RESULTS AND DISCUSSION

We used linear regression algorithm which is a supervised machine learning algorithm where the anticipated output is continuous and characterises a constant slope. It is used to predict values within an endless range, instead of trying to classify them into categories. It performs a regression task**.** Regression represents a target forecast value based on independent variables. It is mostly used for locating out the connection between variables and forecasting. Linear regression performs the work to anticipate a dependent variable value (y) based on a given independent variable (x). So, this regression technique finds out a linear relationship between x and y. Based on this the chances of admission of a student is predicted.

We have also used neural networks which maybe a computational learning system that uses a network of functions to know and translate a knowledge input of 1 form into a desired output, usually in another form. The idea of the synthetic neural network was influenced by human biology and therefore the way neurons of the human brain function together to know inputs from human senses. The layers of functions between the input and therefore the output are what which makes up the neural network.

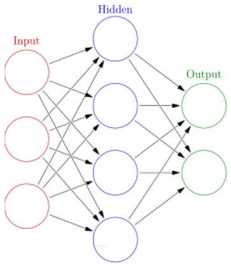


Figure 11: Layers of neural network

We use relu activation function for our input layer which is liable for transforming the summed weighted input from the node into the activation of the node or output for that input.

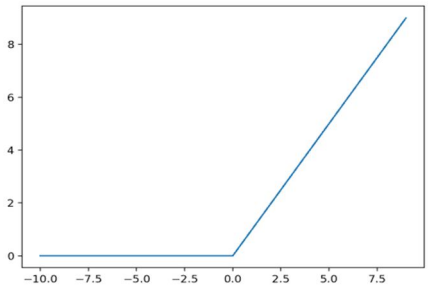


Figure 12: Graph of Relu function

We use sigmoid function for output layer which has the property that they map the entire number line into a small range such as between 0 and 1, or -1 and 1, so one use of a sigmoid function is to convert a true value into one which will be interpreted as a probability.

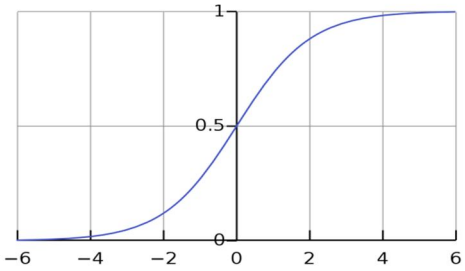


Figure 13: Graph of sigmoid function

The software contains login system for students where they can login themselves to access all the functionalities of the software. Students new to the software have to register first.

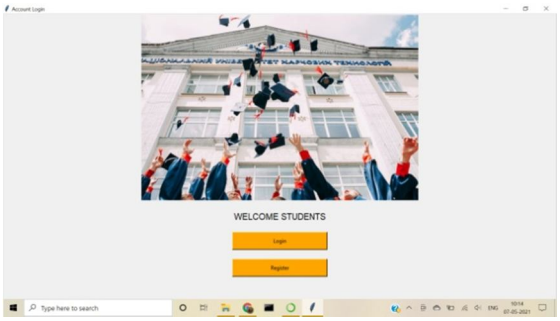


Figure 14: Main page

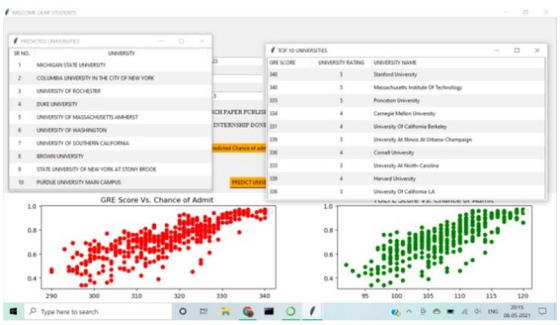


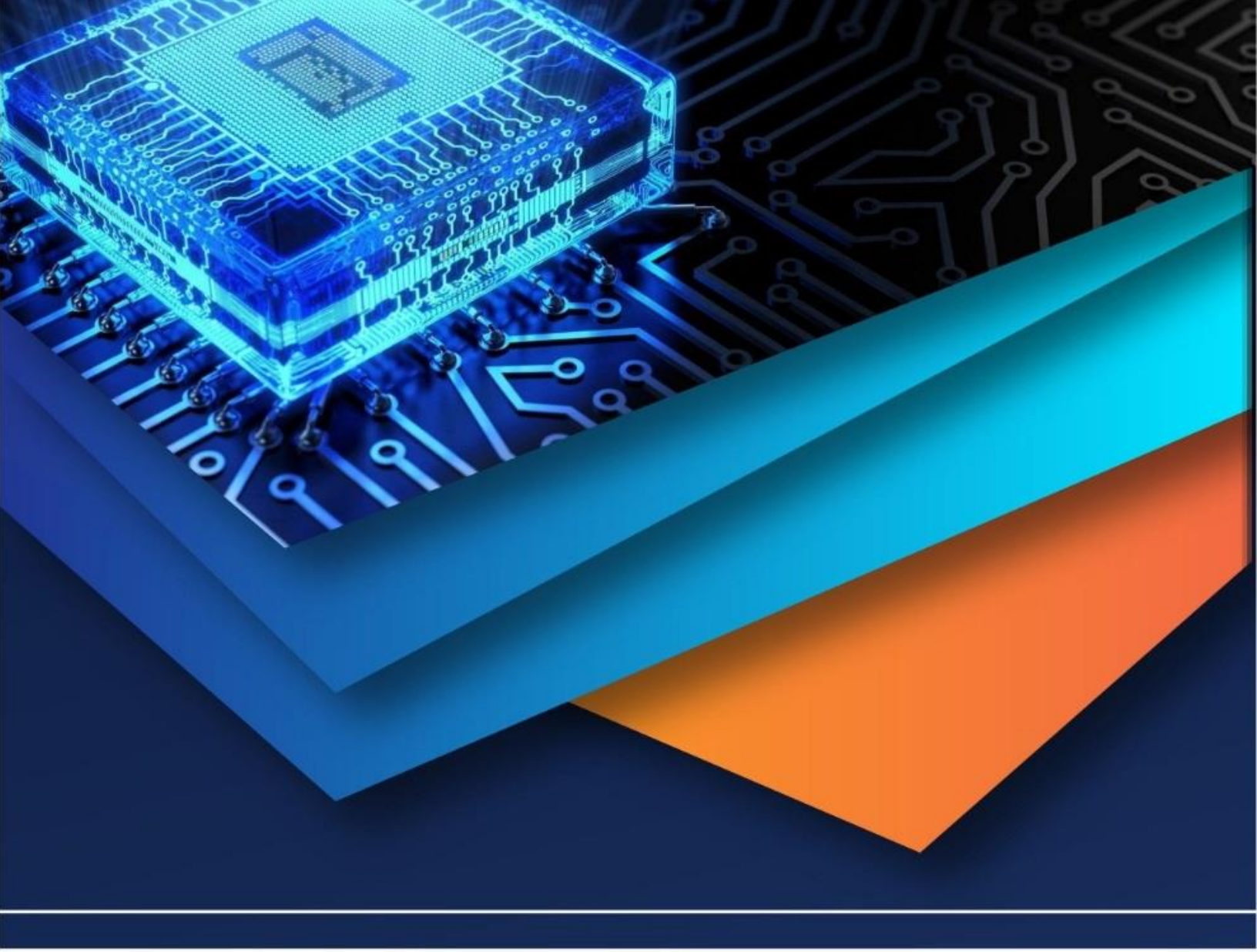
Figure 15: Final Result

VI. CONCLUSIONS

In this project, all the algorithms described are compared with respect to their precision rates. This comparative analysis depicted the strength and the weakness of each one of them in different versions of the dataset. In this project, we presented models which are capable of learning, predicting and recommending the universities for students for their higher education. The model design was explained, with an eye to key functional principles of machine learning and neural networks. Further, the outputs of the model were analysed during a quantitative and qualitative fashion.

REFERENCES

[1] Bilal Zorić, A. (2019). Predicting Students' Success Using Neural Network. RENET-Society for Advancing Innovation and Research in Economy, Zagreb, pp. 58-66, 2019.
[2] Kou-yuan Huang, Senior Member, and Wen-lung Chang, "A Neural Network Method for Prediction of 2006 World Cup Football Game," Proc. of International Joint Council on Neural Networks, pp. 259–266, 2010.
[3] Singh, R. & Pal, S. Application of Machine Learning Algorithms to Predict Students Performance. International Journal of Advanced Science and Technology, Vol 29, No. 5, pp 7249-7261, 2020.
[4] Hasan, H. R., Rabby, A. S. A., Islam, M. T., & Hossain, S. A. (2019, July). Machine Learning Algorithm for Student's Performance Prediction. In 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) pp. 1-7. IEEE, 2019.
[5] Ahmed A, Elaraby I. Data mining: A prediction for student's performance using classification method. World Journal of Computer Application and Technology. Vol. 2, Issue 2, pp:43-47, 2014.
[6] Livieris I. & Pintela P. An improved spectral conjugate gradient neural network training algorithm. International Journal on Artificial Intelligence Tools. Vol. 21, Issue. 1, pp. 1250009-1-21,2012. https://doi.org/10.1142/S0218213011004757
[7] J.S.Moore, "An expert system approach to graduate school admission decisions and academic performance prediction". ScienceDirect, Volume 26, Issue 5, October1998, Pages 659-670, https://www.sciencedirect.com/ science/ article/pii/ S030504-8398000085



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY