IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 7, July 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907224 |

Cyber Schooling with Data Synthesis Using Machine Learning

Mr Sundar K , Keerthivasan R, Mohammed Durvesh K M, Sri charan K V

Assistant Professor, Department of Computer Science and Engineering, Velammal Engineering College , Surapet, Chennai, Tamil Nadu, India

UG Student, Department of Computer Science and Engineering, Velammal Engineering College , Surapet, Chennai, Tamil Nadu, India

ABSTRACT: E-Learning is the process of sharing knowledge through various channels such as e-books, CDs, webinars and more. It has revolutionized the conventional method of chalk and board style of learning imparted to the students. Unlike this, e-Learning education makes giving and receiving simpler, prolific, and productive. E-learning has worked towards bringing learners, tutors, experts, practitioners, and other interest groups to one place. Thus, there is a good practice of knowledge sharing followed through different online platforms. This is important in current times as competition is rising and the world is also growing. However, most of the existing systems are costly and do not offer a performance analyzer. The proposed system is implemented with linear regression to predict the important factor that help both administration and end user

KEYWORDS: Data Synthesis, Predictive analysis.

I.INTRODUCTION

E-learning is evolving from a simple web-based tool into a complex managed environment. Built into new e-learning systems is the ability to monitor a learner's course selections and progress. In an era where organizations seek to tempt innovation from knowledge, the impact of tightly supervised e-learning environments becomes a key consideration. Drawing from ideas stemming from agent and social exchange theories, this paper demonstrates that under conditions where organizational relationships exist involving trust and loyalty, monitoring can have a detrimental effect on the utilization of e-learning environments. Conversely, monitoring may proclaim a positive impact on e-learning usage in circumstances where jobs are simple or mundane, and an abstract relationship defines that between the supervisor and supervisee. The concept of E-learning has been here for more than two decades. What once was just a radical idea now evolved into mainstream phenomenon. The manner in which the e-learning methodology evolved can be approached as a chain process. When the internet started to change, people using the internet started to change and just importantly e-learning pedagogy techniques originated to evolve. Modern e-learning trends are stated as “learner oriented design”. Not only can the student control the appearance of the virtual elements but also have full control over the entire learning process (Downes, 2005). E-learning has evolved a lot from its older style; the tools are combined, making the content creation easy and delivering directly to the Web with increased integrated collaborations, describing future e-learning makes learning “More organic”. In corporations, e-learning is used in elevating sales, technical expertise, professional capability, training and legal compliance preparation.

II.LITERATURE SURVEY

2.1 DESCRIPTIVE STATISTICAL ANALYSIS IN THE PROCESS OF EDUCATION DATA MINING

Proposed System Decision Tree based soft-learning calculations are that appear to be truly outstanding and most helpful prescient of data calculations. Decision Tree based strategies enable understudy prescient data models with high precision, dependability and simplicity of elucidation. In contrast to straight models, will probably delineate direct connections among understudies’ properties great. They are versatile at taking care of any sort of issue close by (characterization or relapse). It is a famous rule utilized in a wide range of information science issues. It works for both clear cut and consistent info and yield factors. Choice tree is one of the quickest methods to recognize most huge factors and connection between at least two factors. Clearly, with the developing cooperation of enormous information on cloud condition, essential choice tree calculations show numerous limitations. Executing a decision tree can take a

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907224 |

great deal of time when the size of understudy dataset is very large. Subsequently new prescient system ought to be utilized for such datasets gathering. Our purpose behind proposing calculation is on the grounds that it does: i. treatment of both discrete qualities, and persistent characteristics; ii. It can process somewhat finish preparing informational collections with qualities not present; iii. Pruning should be possible while building the trees to forestall overfitting issues.

Reference: *"Predictive Analytics For E- learning System Using Machine Learning Approach"*

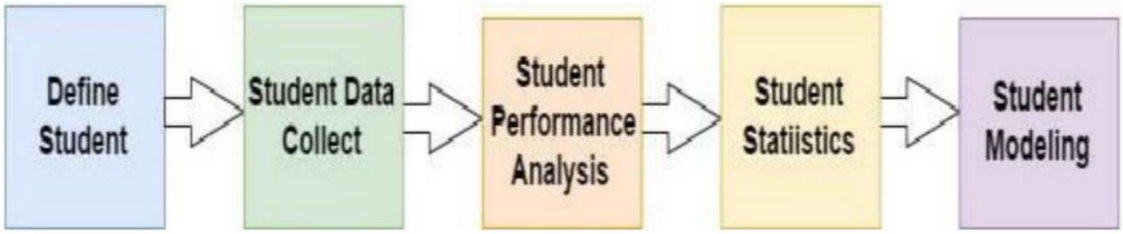


Fig2: Predictive Analytics Process on Data Collections

III. WORKFLOW

The data set used was created by combining multiple sources of Computer Graphics course held at the High School of Electrical Engineering and Computer Science of Applied studies in Belgrade during the summer term in the academic year of 2015/2016. Training set (276 records) of a blended learning environment is created by integrating data extracted from distributed sources into the appropriate organizational form. Extracting student activity records from the Computer Graphics course was done by generating SQL queries over the corresponding tables in the Moodle database. The points that students won on laboratory exercises and lectures are taken from Google Doc file. Integration of data selected from Computer Graphics Moodle database, Google Doc file and IT system of educational institution was carried out. The input data set containing 276 rows of student records was created for the analysis. Integrated data are organized in a tabular form so that the columns indicate the realized activities of students and the rows represent the complete record for each student in a blended learning environment. In the phase of preparation and creation of data set, a descriptive statistical analysis of numerical features was performed. The following measures are calculated: minimum value (Min), maximum value (Max), range (Range), arithmetic mean (X), median (Me), standard deviation (S),mode (Mo) and the skewness (α3). The obtained results of the analysis are noted in a Table. The aim of this research was to achieve greater accuracy and efficiency of the prediction model by determining an appropriate method of discretization of numerical features in the preprocessing phase of training of an educational data set.

Reference: *"Descriptive Statistical Analysis in the Process of Educational Data Mining"*

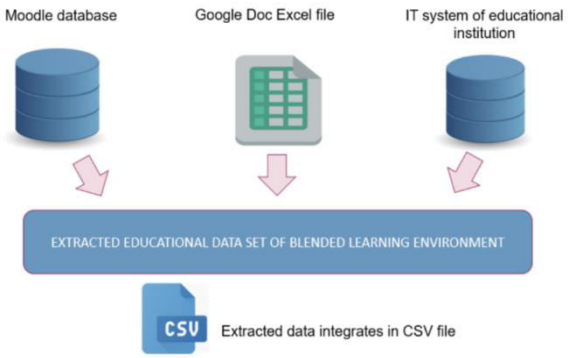


Fig2.1: Creating an educational data set from distributed sources.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907224 |

IV. PROPOSED WORK

4.1 System Analysis:

Machine learning, more specifically the field of predictive modeling is primarily concerned with minimizing the error of a model or making the most accurate predictions possible, at the expense of explainability. In applied machine learning we will borrow, reuse and steal algorithms from many different fields, including statistics and use them towards these ends.

As such, linear regression was developed in the field of statistics and is studied as a model for understanding the relationship between input and output numerical variables, but has been borrowed by machine learning. It is both a statistical algorithm and a machine learning algorithm. The reason is because linear regression has been around for so long (more than 200 years). It has been studied from every possible angle and often each angle has a new and different name. Linear regression is a **linear model**, e.g. a model that assumes a linear relationship between the input variables (x) and the single output variable (y). More specifically, that y can be calculated from a linear combination of the input variables (x).When there is a single input variable (x), the method is referred to as **simple linear regression**. When there are **multiple input variables**, literature from statistics often refers to the method as multiple linear regression. Different techniques can be used to prepare or train the linear regression equation from data, the most common of which is called Ordinary Least Squares.

Making Predictions with Linear Regression

Given the representation is a linear equation, making predictions is as simple as solving the equation for a specific set of inputs.

Let's make this concrete with an example. Imagine we are predicting weight (y) from height (x). Our linear regression model representation for this problem would be:

y = B0 + B1 * x1 or weight =B0 +B1 * height

Where B0 is the bias coefficient and B1 is the coefficient for the height column. We use a learning technique to find a good set of coefficient values. Once found, we can plug in different height values to predict the weight.

For example, let's use B0 = 0.1 and B1 = 0.5. Let's plug them in and calculate the weight (in kilograms) for a person with a height of 182 centimeters.

weight = 0.1 + 0.5 * 182

weight = 91.1

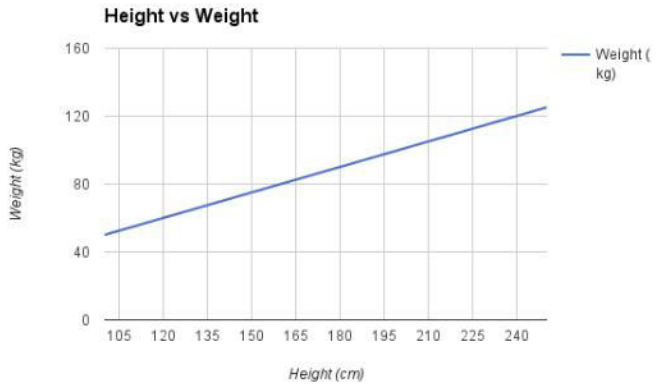


Fig3.1.1: Representing the simple Linear Regression

Preparing Data For Linear Regression

Linear regression has been studied at great length, and there is a lot of literature on how your data must be structured to make best use of the model. As such, there is a lot of sophistication when talking about these requirements and expectations which can be intimidating. In practice, you can use these rules more as rules of thumb when using Ordinary Least Squares Regression, the most common implementation of linear regression.

Linear Assumption. Linear regression assumes that the relationship between your input and output is linear. It does not support anything else. This may be obvious, but it is good to remember when you have a lot of attributes. You may need to transform data to make the relationship linear (e.g. log transform for an exponential relationship).

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907224 |

- **Remove Noise**. Linear regression assumes that your input and output variables are not noisy. Consider using data cleaning operations that let you better expose and clarify the signal in your data. This is most important for the output variable and you want to remove outliers in the output variable (y) if possible.
- **Remove Collinearity**. Linear regression will over-fit your data when you have highly correlated input variables. Consider calculating pairwise correlations for your input data and removing the most correlated.
- **Gaussian Distributions**. Linear regression will make more reliable predictions if your input and output variables have a Gaussian distribution. You may get some benefit using transforms (e.g. log or BoxCox) on your variables to make their distribution more Gaussian looking.
- **Rescale Inputs**: Linear regression will often make more reliable predictions if you rescale input variables using standardization or normalization.

V. ALGORITHM AND TECHNIQUES

5.1 Predictive Analysis Technique:

Predictive analytics is the branch of advanced analytics which is used to make predictions about unknown future events. Predictive analytics uses many techniques from data mining, statistics, modeling, machine learning, and artificial intelligence to analyze current data to make predictions about the future.

Predictive analytics is often defined as predicting at a more detailed level of granularity, i.e., generating predictive scores (probabilities) for each individual organizational element. This distinguishes it from forecasting. For example, "Predictive analytics—Technology that learns from experience (data) to predict the future behavior of individuals in order to drive better decisions." In future industrial systems, the value of predictive analytics will be to predict and prevent potential issues to achieve near-zero break-down and further be integrated into prescriptive analytics for decision optimization. Descriptive models quantify relationships in data in a way that is often used to classify customers or prospects into groups. Unlike predictive models that focus on predicting a single customer behavior (such as credit risk), descriptive models identify many different relationships between customers or products. Descriptive models do not rank- order customers by their likelihood of taking a particular action the way predictive models do. Instead, descriptive models can be used, for example, to categorize customers by their product preferences and life stage. Descriptive modeling tools can be utilized to develop further models that can simulate large number of individualized agents and make predictions.

5.2 Necessity for Predictive Analysis:

By analyzing data about learners' tasks and visualizing outcomes, Learning Analytics can provide real-time feedback about learners and give them tips on improving learning.

1. Identify Individual Training Needs 2. Retain Top Talent

3. Track Career Performance

5.3 Recent Development:

1. Adaptive Learning

Adaptive learning is a style of education where resources, activities, projects, and assignments are tailored to each student's individual needs. In the context of eLearning, the implementation of adaptive learning is usually performed by way of established algorithms and assessments, as opposed to the potentially arbitrary determinations of teachers themselves.Thus far, adaptive learning has been largely experimental, with companies and competitors having spent the past couple of years working out the kinks and engaging in small-scale execution. As eLearning continues to develop, the experiments will end and the widespread adoption will begin. Already, the major eLearning platforms are offering adaptive learning services, and there's no reason that the trend won't continue for the foreseeable future.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907224 |

2. Social Learning

Social learning takes the base components of human interaction and group dynamics and applies them to the modern technological age. Online forums, class-wide chat rooms, file-sharing platforms – with social learning in the electronic space, collaboration has never been more productive, efficient, and seamless. Now, teammates can offer insight and support from anywhere, whether it be their classroom, their homes, or their nearby coffee shops.As social learning applications continue to develop, more and more collaborative tools will likely enter the fray for market dominance.

3. Learning Management Systems

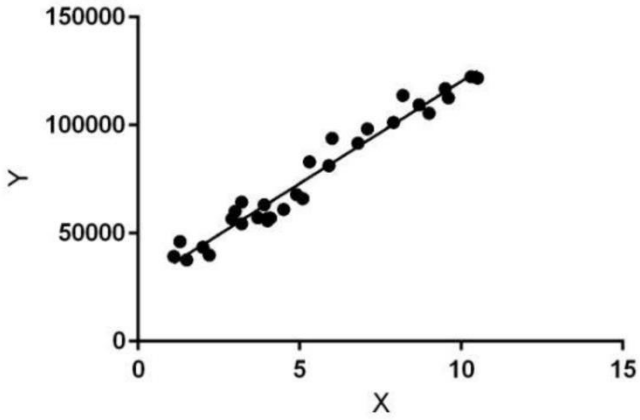
In many collaborative environments and workplaces, employers and managers commonly implement what is known as a content management system, or CMS, to create and store digital content. Recently, this concept has expanded into the world of eLearning. With the advent of the Learning Management System, instructors and other eLearning practitioners are able to develop, document, and administer the courses and curriculums that are produced.Considering the behind-the-scenes nature of LMSs, it has become easier than ever to simultaneously plan ahead and course-correct.ng away with the old analog methods for good. As such, LMSs are here to stay.

VI. DESIGN AND IMPLEMENTATION

6.1 Methodologies

Linear Regression

machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on –



Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x). So, this regression technique finds out a linear relationship between x (input) and y(output). Hence, the name is Linear Regression.In the figure above, X (input) is the number of hours and Y (output) marks in percentage of that subject. The regression line is the best fit line for our model.

TTL:

TTL indexes are special single-field indexes that MongoDB can use to automatically remove documents from a collection after a certain amount of time or at a specific clock time. Data expiration is useful for certain types of information like machine generated event data, logs, and session information that only need to persist in a database for a finite amount of time.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907224 |

Ajax:

AJAX stands for Asynchronous JavaScript and XML. AJAX is a new technique for creating better, faster, and more interactive web applications with the help of XML, HTML, CSS, and JavaScript. AJAX is the most viable Rich Internet Application (RIA) technology so far. It is getting tremendous industry momentum and several tool kit and frameworks are emerging. But at the same time, AJAX has browser incompatibility and it is supported by JavaScript, which is hard to maintain and debug.

6.2 Flow diagram of the entire process:

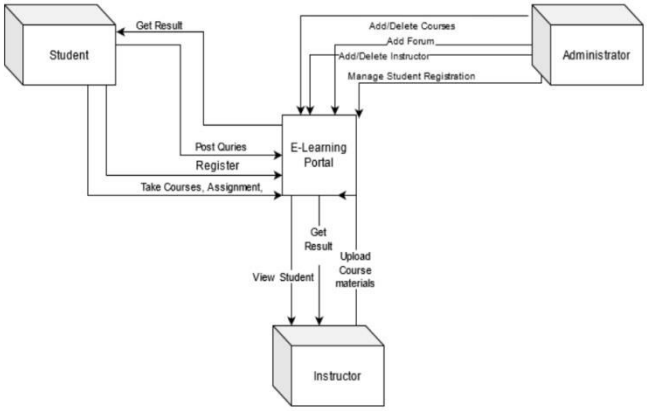


Fig 5.3: Flow Diagram

6.3 Modules Description:

6.3.1 Student Registration Module:

This Login Module is a portal that allows users to type a username and password to log in. You can add this module onto any module tab to allow users to log in to the system. This login/signup module contains a simple form that accepts username and password from the user and stories in Database for the first time and later on the users can use their respective credentials to access the portal.

6.3.2 Administration Module:

Admin module allows system administrator to set up back-end of the system and perform basic system configuration, mainly definition of predefined drop-down fields, definition of classes time schedule, etc. Here mostly Instructor is administrator where they can schedule classes or update timetable or add/delete course also site traffic management can be monitored with the support of Admin Module.

6.3.3 User Interface Module:

Student are the end users of the system so as they login into the portal they get to select the courses they need and check the notes video lecture links also chat with the concerned staff. As student's login into the portal they will have enquiry form that helps admin to collect all the required information about the requirements from the users.

VII. CONCLUSION AND FUTURE WORKS

E-learning is a fast-developing educational model that is likely to replace traditional learning. This model will change the methods of instructions, as it is more learner-centered relative to the traditional model. Learners will access information via computers, interact via computers, and be evaluated via computers. At present time the university is running various programs as full-time courses. dfficult to study for students who are doing some jobs or are unable to attend regular classes due to any other reason.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907224 |

Stake Holder	Major Benefits	Attitudes	Win Conditions	Constraints
College Management	Earn Revenue by reaching more students	Strongly Supportive	Usability, extensibility and reliability	System should be made extensible for future purpose
Faculty Member	----------	Neutral	Performance and Security	----------
Students	Able to gain quality education at home	Strongly Supportive	Usability, Efficiency, Performance	System should be easy to use operate.

Sum Up Table

In conclusion, online learning is beneficial to the students, tutors and the institution offering these courses. I would therefore recommend that online learning be implemented on all learning institutions and research on how to improve this learning process should be carried out.

Future Works:

E-learning is here to stay. As computer ownership grows across the globe e- learning becomes increasingly viable and accessible. Internet connection speeds are increasing, and with that, opportunities for more multimedia training methods arise. With the immense improvement of mobile networks in the past few years and the increase in telecommuting, taking all the awesome features of e-learning on the road is a reality with smartphones.

REFERENCES

[1] A. A. Aziz, S. Osman, S. Widyarto and S. Marjudi, "The Descriptive Data Analysis for E-learning Cloud-based Factor Adoption," 2020 6th International Conference on Interactive Digital Media (ICIDM), Bandung, Indonesia, 2020, pp. 1-3, doi: 10.1109/ICIDM51048.2020.9339645.
[2] G. Dimić, D. Rančić, O. P. Rančić and P. Spalević, "Descriptive Statistical Analysis in the Process of Educational Data Mining," 2019 14th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS), Nis, Serbia, 2019, pp. 388- 391, doi: 10.1109/TELSIKS46999.2019.9002177
[3] Imran, S. Ahmad and D. H. Kim, "Quantum GIS Based Descriptive and Predictive Data Analysis for Effective Planning of Waste Management,"
in IEEE Access, vol. 8, pp. 46193-46205, 2020, doi: 10.1109/ACCESS.2020.2979015.
[4] ROLE OF ICTS IN EDUCATION AND DEVELOPMENT: POTENTIAL, PITFALLS AND CHALLENGES, BY USHA VYASULU REDDI
[5] Mobile Chatting Server for GPRS networks. UsmanAli, TasnimAhmed. 12 May 2008.
[6] Multi-User Chat Application. R. Gayathri, C. Kalieswari. June 202
[7] E-learning development
[8]https://ieeexplore.ieee.org/document/7518886

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details