Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7593066

STUDENT'S PERFORMANCE PREDICTION USING HYBRID OPTIMIZATION ALGORITHM-BASED MAP REDUCE FRAMEWORK

[1]Mahendra S Sawane , [2]Pharindra Kumar Sharma
[1]Research Scholar, Nirwan University Jaipur,AP, G H Raisoni College of Engg. & Mgt. Pune
*Corresponding Author Email id: mahendra.sawane@gmail.com
[2]Assistant Professor, Nirwan University, Jaipur, Rajasthan, India
E-Mai id : cs@nirwanuniversity.ac.in

ABSTRACT
Learning analytics (LA) is an examination field that is extending, and its objectives incorporate choosing, dissecting, and revealing student information (in their cooperation with the web-based learning climate), finding designs in student conduct, and displaying important data in intriguing configurations. A definitive objective of this exploration is to foresee student execution, enhance the educational platform, and carry out customized mediations. Learning analytics (LA) can be characterized as "the estimation, assortment, examination, and announcing of information about students and their specific circumstances, for the reasons for understanding and improving learning and the conditions where it happens," as expressed by the General public of Learning Analytics Research1. LA can likewise be alluded to as "learning insight." The point is exceptionally interdisciplinary, including not just information from learning sciences, teaching method, and social science, yet in addition procedures for AI, educational information mining, measurable examination, informal organization investigation, and normal language handling. Modern outlines of the area can be seen as in. Learning analytics can offer help for different educational assignments, including the accompanying: investigation and perception of information; giving input to supporting teachers; giving suggestions to students; anticipating student execution; student demonstrating; distinguishing unfortunate student ways of behaving; gathering students; interpersonal organization examination; creating idea maps; building courseware; planning and booking. In a similar style, the seven essential objectives of learning analytics can be portrayed as follows: observing and examination; prediction and mediation; coaching and tutoring; evaluation and criticism; transformation; personalization and suggestion; and reflection.

Keywords: Learning, Student, Prediction, Map, Educational.

I. INTRODUCTION

Education is a vital component in the structural support of every society. Students have their minds and personalities developed to a higher level as a result of this. It's possible that the current education system isn't up to snuff when it comes to meeting the ever-evolving demands of the society. Predicting student achievement in advance is a significant component of the new paradigm being implemented in the educational system. Because students are the most important stakeholders in educational systems, academic institutions may be able to satisfy the ever-changing requirements of society if they perform data analysis on student information and come up with a variety of hypotheses based on that information. In addition, the outcomes of these forecasts can be of use when formulating plans to enhance the standard of educational

provision. The kids who receive an education of higher quality are more likely to develop their skills and characteristics. This gives the attention to analyze the academic facts. Models that predict student performance are helpful in conducting an analysis of student data using a variety of data mining approaches. In addition, several different models for predicting student success have been presented in order to make it easier to make predictions about student performance. Both the scholarly community and the educational sector have given a great amount of thought to the development of models that predict the academic achievement of students. Models for predicting student performance take on the challenge of predicting students' grades [1, GPA (Grade Point Average) [2, CGPA [3, and Pass/Fail Course [4]. In EDM (Educational Data Mining), the purpose of students' performance prediction models is not only to achieve a high level of accuracy in prediction models, but also to assist educational stakeholders in the process of forecasting the performance of students. Students are the most valuable resources of any community, and the primary objective of any academic institution is to ensure that its students receive an education of the highest possible standard. In addition, the pupils who receive a good education are more likely to have developed skills and characteristics. Models that predict student performance are helpful in conducting an analysis of student data using a variety of data mining approaches. A significant amount of effort has been put into developing models that can accurately predict the performance of pupils. There are two primary approaches that can be taken when constructing models to predict student achievement. The first approach is the supervised method, and the second is the unsupervised approach. The process of classification is an example of a supervised learning method. According to the findings of [5,] around 71.4% of research articles on models for predicting students' performance use a classification approach. It has been determined to be the most effective approach for the performance prediction models [6]. The target variable in the classification approach is explicitly described as the factor on the basis of which we wish to make predictions regarding grades, GPA, CGPA, or whether students pass or fail. Because of this, we were inspired to concentrate on developing a model that could anticipate the students' performance using the categorization approach.

The performance of students has been predicted over the course of the last few decades by evaluating the influence of various factors, such as emotional characteristics, family characteristics, study schedules, institutional characteristics, and students' scores on assignments, quizzes, and final examinations [7-11]. These kinds of systems have useful applications in a wide variety of fields within academia, such as estimating the success or failure of students as a result of influential factors [12-16]. In this study, the earlier contributions are separated into two distinct categories. The first group is comprised of insightful psychological studies, discoveries made through data mining, and insights obtained through data analysis that in some way contribute to the optimization of systems used to anticipate students' success. The findings of the first group were used as a basis for the second group's report, which details how current prediction systems were improved using the information obtained from the first group. However, the considerable synchronization that took place between the two groups is still a mystery, and as a result, the performance prediction systems of the students became more reliant on a dataset that was taken from the actual world. During the process of optimizing and collecting data, such synchronization can bring helpful

ideas that can be implemented. Additionally, it clears the path for the creation of an assumption-based dataset, which will verify the viability of pilot project implementations and expedite the process of modelling the emotional characteristics of pupils. Participating in an academic competition is not only an efficient method for evaluating the performance of teaching and learning, but it also has a beneficial influence on the academic motivation and study habits of students [17]. The academic competition is a series of activities to locate and solve difficulties through practical activities outside the classrooms. This is an efficient measure to recognize and train young talents [18-21]. Therefore, pupils will benefit much from participating in academic tournaments. The academic competitions have the potential to promote the students' individual efficacy during study activities, as well as the students' understanding of the importance of collaboration and communication. Before the tournament, students will receive training from their allotted tutors. On the other hand, the organization of the competition is fraught with two challenges. Students who choose not to take part in the competition will miss out on the opportunity to win the reward. Because there are so many students, it is challenging for tutors to choose individuals who have the potential to succeed. It is of the utmost importance to choose potential pupils to participate in the competition in order to get good results. Students who have performed very well can be identified through the tutor's observation or through practice examinations. The knowledge and experiences of the tutors introduce an element of bias into these procedures. In addition, they are not appropriate for use if the selection process involves an excessive number of pupils. One of the most important factors contributing to the overall success of all other domains is education (e.g. Medical science, Business). As a result, the achievement of success in all other domains is directly correlated to the quality of the educational system that is in place.

In today's world, educational institutions like universities have amassed a large amount of data that is readily apparent, such as the fundamental information of students, their family circumstances, their grades and scores in a variety of subjects, as well as the particular information of students' speech rate and correct rate in answering questions in the classroom, as well as their attendance in the classroom. This data is used to improve the educational experience of students. Since it is obvious that these data in the field of education are always changing and will grow explosively as a result of the development of education informatization, the question of how to extract useful information from these complex and tedious data for the purpose of data analysis will have a good research value. Using contemporary technology to effectively uncover previously concealed information while processing data resources with enormous data sizes can be a challenging undertaking. (e the development of data mining technology over the past few years has made available powerful technical support for overcoming difficulties of this kind. (e the application of data mining technology to the discovery of valuable data or relevant knowledge law information from a large amount of data, and the application of the analysis and mining results to the teaching management, will have far-reaching significance to improve the quality of school teaching and education management level. Data parallelization processing has become one of the main issues of the current day. One of the most significant difficulties associated with data mining work is figuring out how to integrate data parallel processing mechanisms into the processing workflow. (e combination of data mining algorithm and data parallelism mechanism can make

full use of the parallelism mechanism to accelerate the fast processing of data and also to dig the hidden rules or principles behind the massive data according to a deeper level so as to understand the nature of things more deeply.

Data mining is one of the most essential fields in recent technological developments, and its primary purpose is to retrieve useful information from vast quantities of unstructured and distributed data through the use of parallel processing of data [22]. The techniques of data mining are being utilized in a variety of different fields in order to extract unique information from large data sets. These days, educational institutions are the most common users of data mining tools. The vast majority of researchers have turned to data mining strategies in order to unearth relevant information from the realm of education. The application of data mining in the sphere of education is now the most important focus of study being done today. In order to pursue new studies, the researchers are looking for enthusiasm from the educational community [23]. The practice of data mining has found use in a variety of contexts. Recently, the technique of data mining has been applied in the field of education in order to extract the information that is concealed within educational data sets.

This review paper presents the findings of an analytical investigation as well as a literature-based survey in order to paint a comprehensive picture of the previous research on the subject of student performance analysis and prediction.

II. BIG DATA IN EDUCATION

The expanding capabilities of huge data sets has the potential to improve both logical inquiry and instructional procedures. By assisting teachers in analyzing what their students actually know and identifying which enlightening strategies will typically benefit every understudy, big data might supplement the conventional educational framework. As a result, educators have the chance to learn about and try out novel approaches to teaching that are directly applicable to their profession [24]. When it comes to web-based education, the use of big data is not as indirect as one might think. Recent years have seen tremendous growth in the realm of distance learning. The training sector is becoming increasingly vulnerable [25]. The scholarly community has discovered a plethora of uses for big data. The goals of educational learning analytics and educational information mining are to sift through large amounts of accumulated educational data in search of useful information, to gain insight into the nature of learning, and to uncover solutions that will enable the development of learning implementation and proof of its effectiveness. Insights, software engineering, and AI all contribute to the development of these disciplines.

Fig.1. Big Data in Education

Figure 1 explains how a major information application has been used in several areas of educational information mining, such as to improve students' academic achievements and decrease their likelihood of dropping out of school. Over the past few years, information overload has emerged as a critical issue in the realm of science and innovation. When discussing massive amounts of data, the "3Vs" perspective emphasizes the significance of three factors: velocity, variety, and volume. There is a lot of it: The organization regularly places orders for content from various online business and leisure websites. It's the primary identifier for a huge percentage of the existing data sets. Over the years, information storage capacities have increased dramatically, from kilobytes (KB), megabytes (MB), and gigabytes (GB). Everything on the chart is exponentially larger. The exponential growth of available data is astounding [26]. Another factor that will be discussed is speed. It alludes to the speed with which new datasets are being made available. The rate of progress suggests splitting your attention in two. The first demonstrates an unusually rapid rate of new knowledge production. The second demonstrates the speed with which data is processed. The third model expands on the informational differences that can be made in respect to the various routes, arrangements, and designs for conveying information, but which are outside the scope of present information management organization. It's possible to retrieve data from a wide variety of storage formats, such as structured, quantitative information in traditional data sets and unstructured content records, email, video, and audio.

III. RELATED WORK

The following sections provide an overview of some relevant research based on the findings of an analysis of student performance using a classification and map-lessen approach in massive data. According to Alcala et al. [27], the growing number of approaches to determining proof from information shook up educational information mining. This was caused by the model's tendency to emphasize both the strengths and weaknesses of a given information source. Information about children gathered from the internet, as proposed by Clint McElroy et al. [28], can help teachers quickly identify students who are at risk and others who are thriving. It's likely that these figures will continue to rise. Authors Anna Lea Dickon, et al. [29]

We have introduced a set of learning analytics tools that can assist educators in making informed comparisons and decisions about student behavior, credit, and graphical representations of learning. George Leora et al. [30] offered an open source learning analytics platform that will generate data from many sources and provide essential value to all users in order to guide their decisions about the learning system. et al. visited [31] This study minimally disseminated four distinct characteristics of Los Angeles and associated fields and mapped them onto a reference model. We also looked at the geography of Los Angeles and made a list of the city's problems and potential solutions. Scientists Ourania Petropoulos et al. [32] The Hadoop MapReduce planning technique is explained, and given a brief overview. This is especially useful for scientists. Specifically, MapReduce has significant planning challenges in the following three areas: area, board, and value. The primary aim of booking calculations is to minimize the time needed to complete multiple applications at the same time. It was found by Ferreira et al. [33] Finding the executives' conduct will be less of a hassle with the assistance of educational investigation execution due to the need to combine a wide variety of information sources and data that is simple to generate and is related to the difficulty of treatment.

Mohan et al. [34] presented an undergraduate survey and research of risk-taking based on data collected from students at Indian supplementary schools for optional training. The review focused only on learners from India. A evaluation conducted by Alshammari et al. [35] determined the likelihood of employing a cloud platform to steer an investigation of the genome Bigdata. The vast quantities of data are described with great care. Zhang et al. [36] suggested a model for anticipating student execution that is tied to the entire learning process. The model consists of four distinct phases: data collection and preprocessing; analysis of learning behaviours creation of a calculation model; and final determination. In addition, we employ a refined version of the calculated relapse process to assess the students' behavior and make inferences about their overall performance and character. Oyelade et al. [37] will soon provide a quantifiable approach to obtaining students' ratings based on their performance. Widyahastufi et al. [38] uses WEKA's direct relapse and multi-facet discernments to evaluate the test's plausibility with respect to accuracy, efficiency, and error rates in order to provide a measure of students' growth over the course of the year. A natural prediction model for student accomplishment based on student performance was developed by Wang et al. [39]. The purpose of the model is to foretell students' performance. They used a strategy based on relapse teaching to disseminate knowledge, giving new credence to the educational concept. To estimate the endorsement value open doors available to each student, Ramos et al. [40] use information mining algorithms to determine the student profile and cooperation designs at a higher distance course. Using two different models, Al-Shehri et al. [41] predicted the students' performance on the final exam. Predictions are made about the student's presentation based on the data collected, and then the accuracy of those predictions is evaluated using both the Vector Machine calculation and the KNearest Neighbor technique.

Prachuabsupakij et al. [42] designed rule-based learning as an effective and efficient method for comparing and contrasting the two pretreatment strategies for Destroyed and Releif, with the goal of determining which is more likely to result in a successful graduation rate and thereby improving the quality of education. It is an online tool that uses straightforward Bayesian methods to extract useful information from the records of 700 College of Amrita students.

Devasia et al. [43] developed the software. The findings suggest that the credible Bayesian computation is more accurate than alternative prediction calculations. The investigation by Archanaa et al. [44] provides a close evaluation of many controls. Determination procedures for selection can be found in learning calculations such result checking and trustworthy highlights, packaging methods - classifier subset evaluation. Calculations for learning, including things like clustering, trees, Bayesian methods, result checking, and solid building blocks The idea for a proposal framework that can determine which of five resettlement expresses a student is in was conceived by Thangavel et al. [45]. Not acceptable: dream, centre, and mass. You should also be aware that they have the best odds of achieving the individual placement of students. Students who participate in online discussions from afar could be monitored and graded using a computerized examination system, as proposed by Dinesh et al. [46]. As a result, it is helpful for the teacher to record and evaluate classroom footage in order to get a full picture of the students' progress. As we have seen in the preceding sections, the feasibility of the computations is insufficient, and there are few works that investigate the prediction of findings across a confined dataset. Part of the reason why there is a lack of investigation in this area is because of this. It takes into account a limited number of variables, but will base its forecast on a wide range of information. The computations they've used have their own set of constraints, such as limitations and difficulty. Lots of musts and must not exist.

IV. MAPREDUCE

One must first have a firm grasp on what MapReduce is and why employing it, as opposed to building local multi-strung code for each specific application, is more time and resource efficient. As a versatile calculation, MapReduce can be put to use in a number of contexts, with the creation of a single library and the implementation of a few essential features providing all that is required to run MapReduce on a custom data index. If the client can define both a Map capability and a Decrease capability for the application, MapReduce has the potential to carry out a wide variety of tasks. Early iterations of MapReduce were designed to function on big, appropriation-based frameworks with several hundred physical machines. The organisation connecting these computers may operate on a global (Internet) or local (intranet) scale (intranet). In the next sections, you'll learn more specifics about the components that make up MapReduce. In the beginning, regular personal computers were put to use. There were two processing hubs and two to four gigabytes of random access memory [47]. Because normal IDE hard drives are inexpensive and quick enough for the tasks at hand, they were installed in the machines. These computers are obviously linked together using ethernet ports, with transfer rates ranging from 100 Mbps to 1 GB/s. In the modern era, these groups remain one of the most cost-effective and efficient ways to amass a sizable subscriber base. Individual computers that work together as a group are often referred to as hubs. Generally speaking, a single hub is equivalent to multiple processors or an actual string. All of a company's questions can be answered by a single Expert hub. This Expert hub maintains a database of the several hubs that make up the larger organization. A labourer hub is any node in the network that mostly consists of workers. Risks will be assigned to them, and they will carry them out. All of the necessary tasks have been broken down into specific roles. In MapReduce, depending on the nature of the job, it will be classified as a map task or a decrease task.

One of the numerous benefits of adopting MapReduce is that it simply requires a single library for a wide range of applications. Word count, which tallies the number of words found in records, and other applications that manage massive amounts of data will be simple to develop once the MapReduce library has been completed. According to MapReduce's central notion, just the Map and Reduce capabilities need to be defined for every given programme. MapReduce automatically handles monitoring for machine failures and parallelization issues. Therefore, the Mapper's returned results should be presented to the Minimizer in a matching list format. Due to the theoretical concept of how MapReduce is implemented, it is possible to train the calculation to calculate a large range of things by making super small changes to the code. Sandford University [48] shown that because MapReduce was designed with the ability to recover from performance disappointments, the amount of code required to construct an application can be reduced to a more reasonable level. The Expert hub periodically polls specialists to see if they are active. Laborers can be prompted to get to work by means of a ping. The Expert hub will redirect the affected specialist's work to another employee if it detects a problem with that employee. When it appears there are no more tasks to assign, the Expert hub will start assigning work that is currently being done by other representatives to idle workers. The time lost when a worker stops doing their part is reduced when this method is used. This benefit is in addition to the fact that an expert who has been assigned a task as a backup is more likely to complete it than the worker who was assigned the task initially. Given that all of the organization's machines have the same level of processing capacity, this speedup is not particularly noteworthy. One of the numerous benefits of adopting MapReduce is that it simply requires a single library for a wide range of applications. Word count, which tallies the number of words found in records, and other applications that manage massive amounts of data will be simple to develop once the MapReduce library has been completed. According to MapReduce's central principle, just the Map and Diminish capabilities need to be described for any application. MapReduce automatically handles parallelization and machine errors. Therefore, the Mapper's returned results should be presented to the Minimizer in a matching list format. The conceptual basis for how MapReduce is implemented makes it possible to train the calculation to determine a broad variety of things with only minimal alterations to the code. According to studies conducted at Stanford University[9], one major advantage of this method is a reduction in the amount of code needed to create an application. By design, MapReduce can quickly bounce back from performance setbacks. If a specialist is currently active, the Expert hub will determine this automatically. One way to accomplish this goal is through "pinging" the experts. As soon as the Expert hub detects an issue with a specialist, it will switch the task the affected worker was working on to a different specialist inside the business. When the time comes that there are no longer any fresh tasks to assign, the Expert hub will begin assigning work that is currently being done by other representatives to idle workers. Utilizing this tactic has the advantage of reducing the amount of time lost if a worker stops producing results. In addition to this perk, an expert who has been assigned a job as a backup is more likely to complete the project quickly than the worker who was assigned the job initially. Given that all of the organization's machines have the same level of processing capacity, this speedup is not particularly noteworthy.

V. EXISTING MODELS AND PERFORMANCE PREDICTION SYSTEM OPTIMIZATION

Researchers in the field of using simulated brain networks to predict the outcomes of events in different contexts have garnered significant attention. Articles are also brimming with examples of deep learning algorithms that make predictions and highlight at-risk kids. There aren't many other innovations that provide the opportunity to definitively evaluate performance while also slowing the rate of failure [49-51]. similarly helpful in guiding kids who, because to traumatic events, are unable to focus on their schoolwork (for instance, Coronavirus). This meant that, as we directed the writing audit, we came across a number of fascinating prediction frameworks created by the students. In spite of this, the prediction frameworks were not able to numerically exhibit deep attributes and coordinate them with institutional qualities, focus schedules, and family credits. Family and institutional factors were also excluded from the prediction models.

Research presented in the papers reveals new information about the ways in which various factors have an effect on pupils as a whole. This investigation goes beyond the traditional boundaries of AI and software engineering to investigate the practice of predicting how code will actually be run.

5.1 Related Methods for the Prediction of Performance

As was referenced before, many examinations throughout years and years had the option to settle significant difficulties in the exploration field of anticipating students' exhibition. The previous investigations made various significant commitments, including brain works, proposal frameworks, course suggestions, and execution assessment frameworks for students. Past investigations have shown that broad work has been finished on the advancement of execution prediction frameworks for students. These frameworks utilize data acquired during the students' communications with the qualities of the organizations. This sort of data gives legitimate rules, which are essential for numerically thinking about the normal activities of a student's variables. One of the main elements is that the students' all's data handling frameworks have a similar construction, and that implies that they can be utilized reciprocally to fabricate a learning calculation (mental design). Be that as it may, displaying the relationship between students' feelings (disappointment, stress, and so on) and students' exhibition gets next to no consideration [52-53]. Writing studies are overflowed with numerous discoveries that basically add to prediction calculations and numerical models.

Moreover, the distributed examinations on demonstrating feeling can't be extrapolated to a completely evolved prediction framework. Thusly, it is vital to assess the crucial engineering of the different existing prediction calculations. To assess the presentation of understudies, it is important to explore the subjective discoveries of mental examination and the disclosures made through information investigation. Furthermore, the iterative calculation of the close to home effect on execution during fundamental mental exercises will profit from this data.

Magnificent mental abilities are the main thing that can make a remarkable scholastic achievement conceivable. These capacities are vital for the consummation of any work that requires the utilization of critical thinking systems, thinking, or memory the board. Nonetheless, assuming that an individual's mental capacities are less than ideal, it will be unimaginable for them to get a high score in any of the mental undertakings that might be

expected of them, like composed tests, tests, or tasks. They request students to coordinate their learning, process new information, and review recently put away data (from memory) for resulting use. Hence, anticipating execution while calculating the extreme effect of different gatherings of variables is fundamental not just for mentors to guarantee viable showing strategy, yet additionally for students' achievements and compelling scholarly approaches. This is on the grounds that it considers more precise evaluation of the relationship between various variables. Prior research have given different strategies that can estimate the exhibition of students; in spite of this, they have made the way for new snags that should be overwhelmed by educational frameworks to find true success. The students' capacity levels are continuously moving because of their learning and failing to remember new data. The educational framework requires a framework that is equipped for dealing with the unique way of behaving of students while they are participated in mental exercises.

Different examinations have described the students' characters as well as the basic qualities that make up an individual's character. As per the discoveries, the exhibition prediction configuration can be divided into subsections that, in contrast with different techniques, are more exact. What's more, it gives the way to the improvement of execution prediction plans that are direct and easy to understand. This audit shows that the exhibition prediction framework requires coordination between prediction engineering, mental tests, semantical research, factual investigation, and numerical formulation. This is vital to conjecture future execution precisely. Scientists working in the fields of execution prediction, bioinformatics, information mining, and information coordination were offered an exceptional chance because of this review.

VI. CHALLENGES

Evaluation of student performance can present a number of issues, including the following:

• The data is separated into groups using similarity estimates between the relevant parts of the data set, as determined by the bunching technique. Therefore, we may rest assured that a thorough analysis of the data will be conducted. One of the trickiest challenges of grouping is dividing the data set into distinct categories [54].

• At the time, it was difficult to predict how well a student would perform in a given course without considering how well they had performed in other classes; one of the most challenging aspects of developing a robust indicator is finding a way to accommodate students' diverse geographic and academic backgrounds [55].

• While improved accuracy and shorter computation times are achieved in [56], two significant challenges were not investigated: investigating dynamic neuro fluffy models and examining the assembly rate. We are facing a major roadblock because of this.

• The most challenging barrier to overcoming when employing the Educational Information Mining technique [57] is increasing the generalizability of the results, which will take into account improved prediction performance.

• The AI Approach could provide useful information to academic advisors, who could then recommend the students take the resulting courses and carry out educational mediation measures; nevertheless, the presentation prediction to elective courses was not evaluated.

VII. PROPOSED METHODOLOGY

Developing a framework for anticipating students' intellectual presentations is central to this research. The key promise that can be made by the evaluation is to develop another prediction approach that uses the enhancement calculation. There are two main steps to the suggested cycle: preparation and testing. There will be provision in the proposed method for the impressions that students made on their assignments. The mapper will get the data during the preparation phase. After this information has been gathered, homotropy will be used to select the elements. The following bridging data was created using the elements. After that, the minimizer capability, which employs a hybrid form of the NFS-BCS (Neuro Fluffy framework Bat based chicken multitude) calculation, will be fed the data derived from the middle ground. The NFS-BCS computation is a fusion of the Neuro fuzzy framework, the chicken multiplicity calculation, and the bat calculation. From then on, we can use that as a template for future planning. During the assessment phase, the map capacity, built with the help of the produced combination NFS-BCS calculation, will be fed data about the students. The production of the moderate data will occur simultaneously. The moderate data is then provided to the minimizer capability, which provides a rough estimate of the students' final grades. The proposed method will be implemented in Python, and evaluation will be enabled. The suggested strategy will be evaluated based on a number of metrics, such as the standard deviation of errors and the root-square mean error. Similar to the exhibition measurements, a correlation analysis will be performed, and the results from each methodology will be compared to one another. In particular, we'll compare and contrast the methods now utilized to anticipate student executions with the one we advocate. Figure 2 is a block diagram depicting the proposed model for predicting students' scholastic achievements.

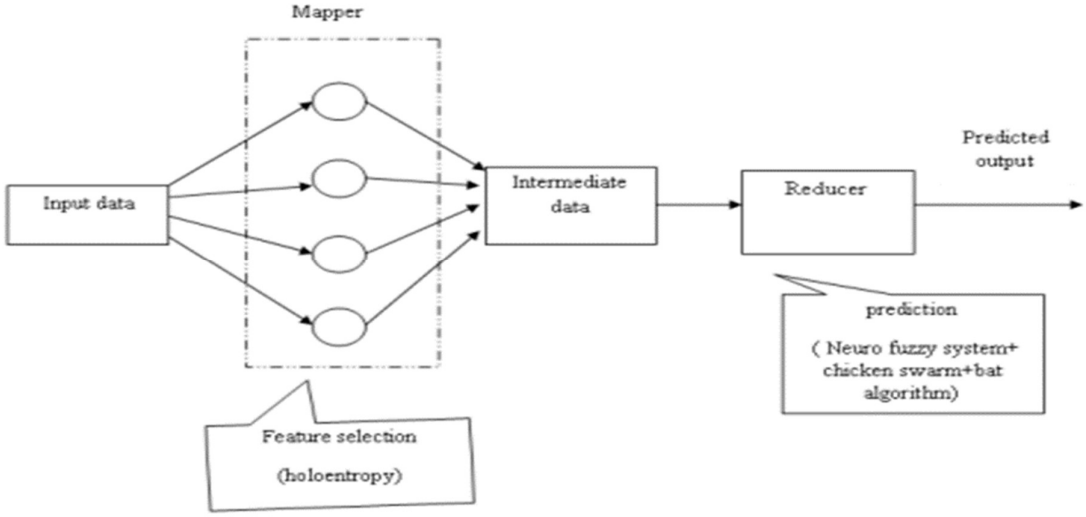

Figure 2: Block diagram of proposed student performance prediction model

VIII. CONCLUSIONS

The survey that has been given causes to notice the conceivable examination valuable open doors that exist to further develop the exhibition prediction frameworks utilized by the students. Foreseeing the scholastic outcome of students has arisen as a noticeable report center. Most of the as of now available prediction models were developed utilizing a strategy called AI. Broad examination is directed on student scholastic execution in advanced education (HE) to battle

scholarly underachievement, developing college dropout rates, and graduation delays, among other troublesome hardships. To put it another way, student execution can be perceived as how much both present moment and long haul educational objectives are accomplished. They are more worried about the exactness of the predictions; however, they try to ignore the interpretability. Subsequently, we can reach the determination toward the finish of this paper that the procedure of map diminish calculation that was proposed is useful to anticipate the students' presentation examination and plan the preparation action to them for them to get propelled in their intrigued space as per their decision.

REFERENCES

[1] M. A. Al-Barrak and M. S. Al-Razgan, "Predicting students 'performance through classification: A case study," Journal of Theoretical & Applied Information Technology, vol. 75, no. 2, pp. 167–175, 2015.

[2] A A. Aziz, N. H. Ismail, F. Ahmad and H. Hassan, "A framework for students 'academic performance analysis using naïve Bayes classifier," Jurnal Teknologi (Sciences & Engineering), vol. 75, no. 3, pp. 13–19, 2015.

[3] N. Buniyamin, U. bin Mat and P. M. Arshad, "Educational data mining for prediction and classification of engineering students achievement," in Int. Conf. on Engineering Education, Japan, IEEE, pp. 49–53, 2015.

[4] L. Ramanathan, S. Dhanda and S. Kumar, "Predicting students' performance using modified ID3 algorithm," International Journal of Engineering and Technology, vol. 5, no. 3, pp. 2491–2497, 2013.

[5] C. Del Río and J. P. Insuasti, "Predicting academic performance in traditional environments at higher-education institutions using data mining: A review," Ecos de la Academia, vol. 2, no. 4, pp. 185–201, 2016.

[6] P. Thakar, "Performance analysis and prediction in educational data mining: A research travelogue," International Journal of Computer Application, vol. 110, no. 15, pp. 975–8887, 2015.

[7] Sumanasekera I., Abd Hamid J., Khatibi A., Azam S. F. Involvement and style of parents on student motivation towards student performance with the moderating effect of academic causal factors: development of a conceptual model. Global Journal of Management and Business Research . 2021;21(1):10–24.

[8] Osburn M. Z., Stegman C., Suitt L. D., Ritter G. Parents' perceptions of standardized testing: its relationship and effect on student achievement. Journal of Educational Research & Policy Studies. 2004;4(1):75–95.

[9] Hoover-Dempsey K. V., Battiato A. C., Walker J. M. T., Reed R. P., DeJong J. M., Jones K. P. Parental involvement in homework. Educational Psychologist. 2001;36(3):195–209. doi: 10.1207/s15326985ep3603_5.

[10] Huang H., Liang G. Parental cultural capital and student school performance in mathematics and science across nations. The Journal of Educational Research. 2016;109(3):286–295. doi: 10.1080/00220671.2014.946122.

[11] Taylor L. C., Hinton I. D., Wilson M. N. Parental influences on academic performance in african-american students. Journal of Child and Family Studies. 1995;4(3):293–302. doi: 10.1007/bf02233964.

[12] Wößmann L. Schooling resources, educational institutions and student performance: the international evidence. Oxford Bulletin of Economics and Statistics. 2003;65(2):117–170.

[13] Mutodi P., Ngirande H. The impact of parental involvement on student performance: a case study of a south african secondary school. Mediterranean Journal of Social Sciences. 2014;5(8): p. 279. doi: 10.5901/mjss. 2014.v5n8p279.

[14] Cabus S. J., Ariës R. J. What do parents teach their children? - the effects of parental involvement on student performance in Dutch compulsory education. Educational Review. 2017;69(3):285–302. doi: 10.1080/00131911.2016.1208148.

[15] Harris A., Goodall J. Do parents know they matter? engaging all parents in learning. Educational Research. 2008;50(3):277–289. doi: 10.1080/00131880802309424.

[16] Khan R. M. A., Iqbal N., Tasneem S. The influence of parents educational level on secondary school students academic achievements in district rajanpur. Journal of Education and Practice. 2015;6(16):76–79.

[17] Van Nul, W.T.D.; Roach, V.A. Head-to-head: The role of competition in undergraduate education. Anat. Sci. Educ. 2015, 8, 404–412.

[18] Campbell, H.W.J.R.; Walberg, H.J. The theory of a general quantum system interacting with a linear dissipative system. Annals of Physics. 2000, pp. 547–607.

[19] Campbell, J.R.; Walberg, H.J. Olympiad studies: Competitions provide alternatives to developing talents that serve national interests. Roeper Rev. 2010, 33, 8–17.

[20] Goldstein, D.; Wagner, H. After school programs, competitions school Olympics, and summer programs. Int. Handb. Res. Dev. Gift. Talent 1993, 33, 593–604.

[21] Urhahne, D.; Ho, L.H.; Parchmann, I.; Nick, S. Attempting to predict success in the qualifying round of the international chemistry olympiad. High Abil. Stud. 2012, 23, 167–182.

[22] Pratiyush Guleria, Manu Sood ", Big Data Analytics: Predicting Academic Course Preference Using Hadoop Inspired MapReduce", IEEE, 2017.

[23] A. Dinesh Kumar, Dr. V. Radhika," A Survey on Predicting Student Performance", International Journal of Computer Science and Information Technologies (IJCSIT), Vol 5.

[24] D.M West, "Big Data for Education: Data Mining, Data Analytics, and Web Dashboards," Government Stud. Brook US Reuters, 2012. 13.

[25] Athanasios S. Drigas, P. Lelipoulous, "The Use of Big Data in Education", Internal Journal of Computer Science Issues, Vol-11, September 2014

[26] R.Swathi, N.Pavan Kumar, L.KiranKranth, "Systematic Approach on Big Data Analytics in Education Systems", International Conference on Intelligent Computing and Control Systems (ICICCS), 2017.

[27] Alcalá-Fdez, J., Sanchez, L., Garcia, S., del Jesus, M. J., Ventura, S., Garrell, J. M., ... & Fernández, J. C. (2009). KEEL: a software tool to assess evolutionary algorithms for data mining problems. Soft Computing-A Fusion of Foundations, Methodologies, and Applications, 13(3), 307-318.

[28] McElroy, C. (2011). The online student profile learning system: A learner-centered approach to learning analytics.

[29] Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. Journal of Educational Technology & Society, 15(3), 58.

[30] Lepouras, G., Katifori, A., Vassilakis, C., Antoniou, A., & Platis, N. (2014, July). Towards a learning analytics platform for supporting the educational process. In Information, Intelligence, Systems and Applications, IISA 2014, The 5th International Conference on (pp. 246-251). IEEE.

[31] Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. International Journal of Technology Enhanced Learning, 4(5-6), 318-331.

[32] Ourania Petropoulou, Katerina Kasimatis, Ioannis Dimopoulos, and Symeon Retalis, LAe-R: A new learning analytics tool in Moodle for assessing students' performance, Bulletin of the IEEE Technical Committee on Learning Technology, Volume 16, Number 1, January 2014

[33] Ferreira, S. A., & Andrade, A. (2014). Academic analytics: mapping the genome of the University. IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, 9(3), 98-105.

[34] Mohan, M. M., Augustin, S. K., & Roshni, V. K. (2015, December). A BigData approach for classification and prediction of student result using MapReduce. In Intelligent Computational Systems (RAICS), 2015 IEEE Recent Advances in (pp. 145-150). IEEE

[35] Alshammari, H., Bajwa, H., & Lee, J. (2014, May). Hadoop based enhanced cloud architecture for bioinformatic algorithms. In Systems, Applications, and Technology Conference (LISAT), 2014 IEEE Long Island (pp. 1-5). IEEE.

[36] Zhang, W., Huang, X., Wang, S., Shu, J., Liu, H., & Chen, H. (2017, June). Student Performance Prediction via Online Learning Behavior Analytics. In Educational Technology (ISET), 2017 International Symposium on (pp. 153-157). IEEE.

[37] Oyelade, O. J., Oladipupo, O. O., & Obagbuwa, I. C. (2010). Application of k Means Clustering algorithm for prediction of Students Academic Performance. arXiv preprint arXiv:1002.2425.

[38] Widyahastuti, F., & Tjhin, V. U. (2017, July). Predicting students' performance the final examination using linear regression and multilayer perceptron. In Human System Interactions (HSI), 2017 10th International Conference on(pp. 188- 192). IEEE.

[39] Li, X., Xie, L., & Wang, H. (2016, August). Grade Prediction in MOOCs. In Computational Science and Engineering (CSE) and IEEE Intl Conference on Embedded and Ubiquitous Computing (EUC) and 15th Intl Symposium on Distributed Computing and Applications for Business Engineering (DCABES), 2016 IEEE Intl Conference on (pp. 386-392). IEEE.

[40] Ramos, T., Gomes, A., Lucena, M., Nunes, I., Valentim, R., & Nóbrega, G. (2017, June). Use of educational data mining to identify distance learning students' profiles and patterns of participation. In Information Systems and Technologies (CISTI), 2017 12th Iberian Conference on (pp. 1-6). IEEE.

[41] Al-Shehri, H., Al-Qarni, A., Al-Saati, L., Batoaq, A., Badukhen, H., Alrashed, S., ... & Olatunji, S. O. (2017, April). Student performance prediction using Support Vector Machine and K-Nearest Neighbor. In Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on (pp. 1-4). IEEE.

[42] Prachuabsupakij, W., & Doungpaisan, P. (2016, October). Matching preprocessing methods for improving the prediction of student's graduation. In Computer and Communications (ICCC), 2016 2nd IEEE International Conference on (pp. 33- 37). IEEE.

[43] Devasia, T., Vinushree, T. P., & Hegde, V. (2016, March). Prediction of students performance using Educational Data Mining. In Data Mining and Advanced Computing (SAPIENCE), International Conference on (pp. 91-95). IEEE.

[44] R. Archanaa, V. Athulya, T. Rajasundari and M. V. K. Kiran, "A comparative performance analysis on network traffic classification using supervised learning algorithms," 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, 2017.

[45] Thangavel, S. K., Bkaratki, P. D., & Sankar, A. (2017, January). Student placement analyzer: A recommendation system using machine learning. In Advanced Computing and Communication Systems (ICACCS), 2017 4th International Conference on (pp. 1-5). IEEE.

[46] Dinesh, D., & Bijlani, K. (2016, August). Student analytics for productive teaching/learning. In Information Science (ICIS), International Conference on(pp. 97-102). IEEE. International Journal of Pure and Applied Mathematics Special Issue 14

[47] Jeffrey Dean and Sanjay Ghemawat. Map-Reduce: Simplified Data Processing on Large Clusters 0018-9162/95. D OSDI IEEE, 2004.

[48] Unknown. The Phoenix System for MapReduce programming. http: //mapreduce.stanford.edu.

[49] Li M., Chen Y., Lal C., Conti M., Alazab M., Hu D. Eunomia: anonymous and secure vehicular digital forensics based on blockchain. IEEE Transactions on Dependable and Secure Computing . 2021

[50] Li M., Chen Y., Zheng S., Hu D., Lal C., Conti M. Privacy-preserving navigation supporting similar queries in vehicular networks. IEEE Transactions on Dependable and Secure Computing . 2020;19 doi: 10.1109/TDSC.2020.3017534.

[51] Khan S., Zhang Z., Zhu L., Li M., Khan Safi Q. G., Chen X. Accountable and transparent tls certificate management: an alternate public-key infrastructure with verifiable trusted parties. Security and Communication Networks . 2018;2018:16. doi: 10.1155/2018/8527010.8527010

[52] Ali S R. M., Perumal S. Multi-class lda classifier and cnn feature extraction for student performance analysis during covid-19 pandemic. International Journal of Nonlinear Analysis and Applications. 2022;13(1):1329–1339.

[53] Fard A. P., Mahoor M. H. Facial landmark points detection using knowledge distillation-based neural networks. Computer Vision and Image Understanding. 2022;215 doi: 10.1016/j.cviu.2021.103316.103316

[54] Liu, L., Sun, L., Chen, S., Liu, M., Zhong, J.: K-PRSCAN: a clustering method based on PageRank. Neurocomputing 175, 65–80, 2016.

[55] Jie Xu, Kyeong Ho Moon, and Mihaela van der Schaar, " A machine learning approach for tracking and predicting student performance in degree programs," IEEE Journal of Selected Topics in Signal Processing, 11(5), pp.742-753, 2017.

[56] Le Hoang Son and Hamido Fujita," Neural-fuzzy with representative sets for prediction of student performance," Applied Intelligence, 2018.

[57] Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. Computers & Education, 113, 177–194