



**ARTICLE**

# A New Framework for Scholarship Predictor Using a Machine Learning Approach

**Bushra Kanwal<sup>1</sup>, Rana Saud Shoukat<sup>2</sup>, Saif Ur Rehman<sup>2,\*</sup>, Mahwish Kundi<sup>3</sup>, Tahani ALSaedi<sup>4</sup> and Abdulrahman Alahmadi<sup>4</sup>**

<sup>1</sup>School of Science and Technology, Foundation University Islamabad, Rawalpindi, 46000, Pakistan

<sup>2</sup>University Institute of Information Technology, PMAS-University of Arid Agriculture, Rawalpindi, 46000, Pakistan

<sup>3</sup>Maynooth International Engineering College, Maynooth University, Maynooth, W23 F2H6, Ireland

<sup>4</sup>Applied College, Taibah University, Madina, 42353, Saudi Arabia

\*Corresponding Author: Saif Ur Rehman. Email: saif@uaar.edu.pk

Received: 03 June 2024 Accepted: 16 July 2024 Published: 31 October 2024

## ABSTRACT

Education is the base of the survival and growth of any state, but due to resource scarcity, students, particularly at the university level, are forced into a difficult situation. Scholarships are the most significant financial aid mechanisms developed to overcome such obstacles and assist the students in continuing with their higher studies. In this study, the convoluted situation of scholarship eligibility criteria, including parental income, responsibilities, and academic achievements, is addressed. In an attempt to maximize the scholarship selection process, numerous machine learning algorithms, including Support Vector Machines, Neural Networks, K-Nearest Neighbors, and the C4.5 algorithm, were applied. The C4.5 algorithm, owing to its efficiency in the prediction of scholarship beneficiaries based on extraneous factors, was capable of predicting a phenomenal 95.62% of predictions using extensive data of a well-esteemed government sector university from Pakistan. This percentage is 4% and 15% better than the remainder of the methods tested, and it depicts the extent of the potential for the technique to enhance the scholarship selection process. The Decision Support Systems (DSS) would not only save the administrative cost but would also create a fair and transparent process in place. In a world where accessibility to education is the key, this research provides data-oriented consolidation to ensure that deserving students are helped and allowed to get the financial assistance that they need to reach higher studies and bridge the gap between the demands of the day and the institutions of intellect.

## KEYWORDS

Education; data mining; C4.5 algorithm; decision support system; scholarship guarantee; machine learning

## 1 Introduction

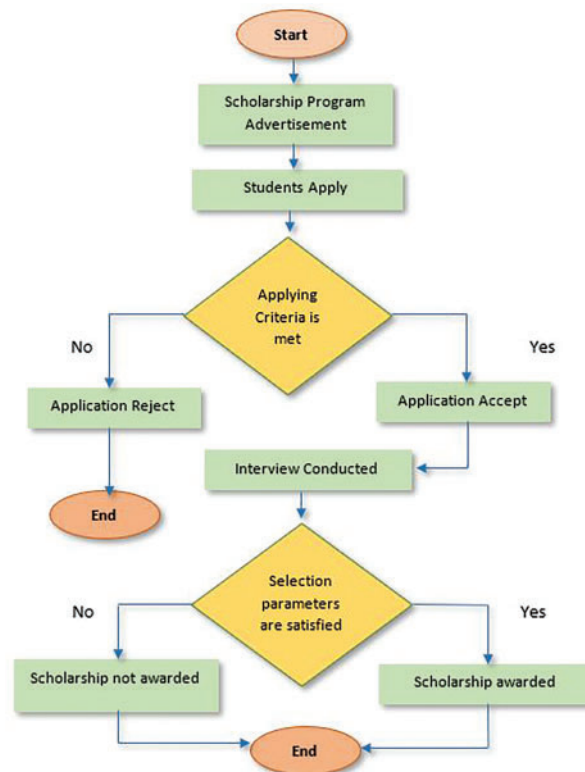
Education is widely regarded as one of the most essential and challenging issues in human life, influencing both individual and societal well-being [1,2]. Education enlightens one's intellect and way of thinking, aiding students in preparing for a profession or further education after university graduation. Educated individuals tend to feel, think, and act in ways that contribute to their success,



boosting both their personal happiness and the happiness of their communities. Financial literacy, defined as the ability to use money intelligently, is crucial in today's financial environment, yet current infrastructure tools are insufficient for young people to comprehend money and make sound financial decisions [3]. It is critical for young university students to have adequate financial literacy to manage their finances wisely and prepare for entering the workforce [4,5].

Many students leave their studies due to financial issues. Educational expenses make it difficult for parents to cover direct educational costs such as tuition, fees, and books, impacting educational outcomes because the child is either missing or under-prepared for school [6]. Financial issues often force students to work alongside their studies, which affects their academic performance. These situations impact students' educational outcomes, such as enrollment, achievement, attendance, and performance [7]. Therefore, solid support structures are needed to help students continue their studies [8]. Most universities provide scholarships to needy students to enable them to complete their education.

Most government sector universities follow a procedure, as shown in Fig. 1, for allocating scholarships to needy students. The Higher Education Commission (HEC) announces need-based scholarships with conditions such as parental income below 45,000, number of family dependents, and family expenditure. Many students apply for the scholarships; some meet the criteria, while others do not. The university rejects applications from students who do not meet the defined criteria and selects those who do for interviews. The final decision is then made to award scholarships to deserving students.

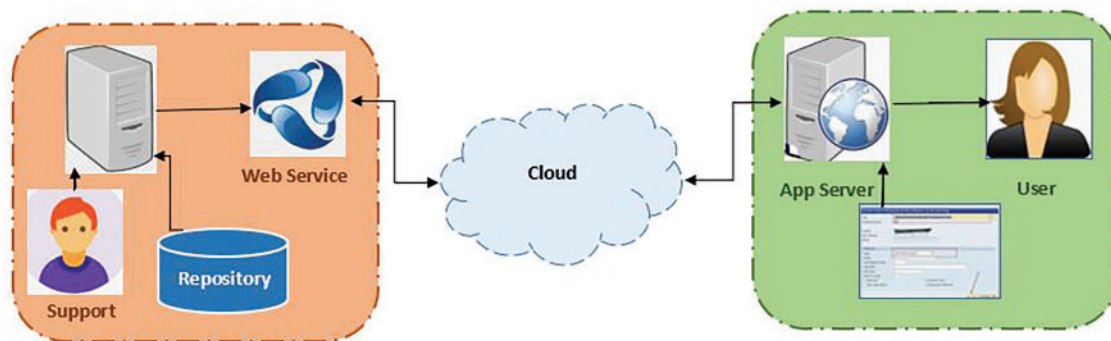


**Figure 1:** Flow chart of scholarship selection

Scholarships are a way of providing financial support to students for their educational purposes. They are awarded by governments or institutions to help students cope with the rising costs of higher education [9]. Scholarships enable students with financial difficulties to continue their studies amid rapid technological development [10,11]. Tuition fee aid, for example, helps underprivileged students complete their studies. Universities aim to develop effective Decision Support Systems (DSS) to select the most eligible students for financial aid programs [12].

Determining which students are eligible for scholarships can be challenging for universities. A DSS is needed to ensure that students are qualified for scholarships. A DSS is a computer-based interactive system that assists but does not replace decision-makers [13]. The DSS can tackle semi-structured and unstructured issues for decision-making. Advances in web technology have made it possible to provide and exchange DSS applications more efficiently. The design and operation of web-based DSS are now centered on online technologies, offering a robust tool that assists users in identifying data and services available as intranet web services [14]. According to authors in [12], the selection of a web-based scholarship guarantee is based on data mining, as these applications can be employed with large-scale databases. Variables are dynamically weighted, allowing the application to identify the most important variables, such as academic performance and financial circumstances.

Fig. 2 shows the architecture of a web-based DSS. DSS scholarship programs aim to help students continue their studies by providing financial aid through various seminars and training activities [15]. This assistance not only covers educational expenses but also serves as an award for students based on their achievements.



**Figure 2:** DSS web based architecture [16]

DSS scholarship is a development program to help students in order to continue their studies, in the form of financial aid. Similarly, through numerous seminars and training activities [15]. Moreover, this helps with the educational expense as well as acts as an award for the students, because there exists a category of scholarships based upon an individual's success.

The difficulties in obtaining scholarships arise from criteria established for specific study programs, such as non-academic achievements, academic achievements, parental income, majors, and number of dependents [17]. Therefore, not all applicants will receive scholarships. Only those who meet the criteria set by the authorities will be awarded scholarships. Hence, a DSS is required to determine the eligibility of applicants [9,18].

Students often need scholarships due to financial constraints and a lack of jobs. The number of applicants for scholarships from the Higher Education Commission is increasing, necessitating a system that can efficiently process all applications. An automatic DSS is required to facilitate

universities in selecting deserving students [19]. This study addresses the challenges of selecting deserving scholarship recipients in higher education. Many university students face financial constraints that hinder their education. The current scholarship selection process involves complex eligibility criteria that can be daunting for both students and institutions [20]. To address this issue, machine learning algorithms, including the C4.5 algorithm, have been employed to streamline the selection process. However, the effectiveness of these methods varies. The study's contribution is to optimize the scholarship selection process by utilizing extensive datasets from government sector universities in Pakistan. Table 1 lists down some of the most commonly used acronyms along with their abbreviations throughout the paper.

**Table 1:** List of abbreviations used throughout the article

Abbreviation	Description
DSS	Decision Support System
PSO	Particle Swarm Optimization
IPSO	Inclined Planes System Optimizations
GSA	Gravitational Search Algorithm
C4.5	A Decision Tree Algorithm
SVM	Support Vector Machine
ANN	Artificial Neural Network
SAW	Simple Additive Weighting
AHP	Analytical Hierarchy Process
ANP	Analytic Network Process
TOPSIS	Technique for Order Preference by Similarity
TP	True Positives
TN	True Negatives
FP	False Positives
FN	False Negatives

### ***1.1 Motivation***

Education is generally considered one of the most important, but at the same time the most complex issues of human life, and it has a significant impact on the well-being of any individual and social progress in general [2]. Higher education, in its turn, holds an extremely important place in shaping success in career growth and personal development. Financial difficulties, however, can deprive students of quality education, contributing to dropout rates and actual underperformance [21]. Scholarships are the main alternative financial aid connections intended to alleviate obstacles and help students reach their educational goals. The motivation of this study is based on removing the inefficiencies of the scholarship choice process to ensure the most deserving students receive financial support to contribute to economic development in the field of equity in education [22].

### ***1.2 Innovation***

Its novelty is in applying the C4.5 algorithm in developing an automatic scholarship selection system, embedded in this system, which we shall term a decision support system (DSS). While the

application of such algorithms to solving similar problems has been reported in the previous works, the flexibility of the C4.5 algorithm has the potential to provide more desirable performance since it is not only interpretable but also powerful and able to deal with continuous as well as discrete data. Furthermore, while most of the reported previous works are based on simulations, we have implemented the proposed method in a working system, realized in a university setup, and accordingly demonstrated the feasibility and scalability of the proposed system. Most interestingly, the effectiveness of the C4.5 algorithm together with those of heuristic and evolutionary search optimization algorithms is compared, thereby giving a complete study of its robustness and applicability.

### ***1.3 Research Contributions***

This research work has significant implications on the scholarship selection process in the application domain by using machine learning. The complete detailed analysis of various parameters used in the scholarship selection process, i.e., parental income, family expenditure, number of dependents, academic performance, etc., is provided. It provides an overall automated DSS as an effective solution that reduces the manual efforts as well as the chances of involved biases and makes the selection process efficient and reliable. The empirical validation based on practical implementation adds credibility to the empirical section of the research work. Moreover, the comparative discussion with heuristic and evolutionary techniques adds more value to the discussion approach of algorithm effectiveness and applicability with the advantages of the C4.5 algorithm.

### ***1.4 Economic Impact and Value Creation***

It is also necessary to evaluate the economic impact and value creation of the proposed approach to understand its wider relevance beyond the academic and technical achievements [22]. This section assesses the extent of possible global and local economic gains from selecting scholarships using the C4.5 algorithm.

#### ***1.4.1 Local Economic Impact***

At the local level, deployment of an efficient and automated DSS for the selection of scholarships has the potential of generating significant conservation of resources in terms of administrative costs for educational institutions. Conventional methods of scholarship selection inevitably demand the hiring of large numbers of human resources, especially manpower for data collection, validation, and decision-taking. By making such processes computerized, the DSS satisfies the need for high levels of administrative inputs, thereby allowing the universities to achieve maximum possible utilization of their resources.

The fair and impartial award of scholarships ensures that maximum possible deserving students are provided financial assistance which, by the same token, is able to also lift their educational performance and employability. Higher levels of education can often result in higher levels of employment and income, and thereby contribute to the local economic lift-up of the community. Graduates who are the recipients of scholarships are more likely to secure higher paying jobs, contribute to the local economy in terms of higher levels of spending, and perhaps even open new businesses, thereby again creating more employment opportunities.

#### ***1.4.2 Global Economic Impact***

At the global level, the adoption of automated DSS for scholarship awarding can be seen as a process that will lead to the standardization and better effectiveness of the provision of financial

aids to multiple educational institutions. The standardization that is emerging can possibly lead to an equitable global mechanism of education, whereby deserving students of all walks of life can avail themselves of higher educational opportunities regardless of their financial impositions.

The ripple effect of the provision of increased access to education can spread very deeply. Educated individuals are more likely to innovate, drive technological advancements, and contribute towards the knowledge economy of the world. The aggregate of individual contributions can help in the facilitation of the propulsion of economic growth at the global level, reduction in levels of poverty, and enhancement of general well-being of society.

Moreover, the application of machine learning algorithms such as C4.5 in education management can act as a model to the other discipline, which can result in the acceptance of superior technology directed at improving operations and value addition. Application of technology can result in the enhancement of productivity, innovation, and economic diversification at the global level.

### ***1.5 Quantitative Analysis***

To provide a quantitative measure of the economic contribution, a cost-benefit analysis was undertaken between the traditional scholarship selection process and the use of the automated DSS. Variables like the administrative cost savings, improvement in student retention and graduation rates, and a potential increase in the income bracket of the graduates were factored for the analysis.

It is found that the use of the DSS may potentially save up to 30% of the administrative costs, which is not insignificant for an educational institution. The improvement in student retention and graduation rates, further increasing the number of graduates serving the industry every year by 10%, is mooted. Assuming 20% of an average increase in income for those graduates over the non-degree holders, the measure of total economic contribution per institution might range in the millions of the dollar mark per annum.

C4.5 was selected for this research for a number of good reasons. It is a very good mechanism for building easy-to-visualize and comprehensive induction models, which is very important when you need to communicate the decision-making process to non-experts, primarily because of interpretability and transparency facilities. Furthermore, it is also robust to different datasets, and its ability to address continuous and categorical datasets makes it even more versatile. By focusing attention on the most contributing factors that are identified in connection to the research problem, the inherent feature selection mechanism of the algorithm helps in identifying the most significant variables. Moreover, it is helpful in providing resilience to noise and processes null values in an efficient manner, hence making it more reliable in practical applications where data are likely to be incomplete. The utilization of C4.5 is also encouraged by an empirical record of successful applications in a diverse range of domains, ranging from financial analysis to medical diagnostic tasks. Comparative studies have also proved that C4.5 is a fair candidate to deal with the present research problem in the sense that it is not only accurate, but also efficient and less demanding in terms of computing resources.

The remains of this research work are planned as: [Section 2](#) delivers a literature review related to student guarantee scholarship with DSS. [Section 3](#) delivers a proposed methodology explanation of our research work, how to collect data, precise that data and apply C4.5 algorithm on collected precise data. [Section 4](#) delivers a result of data which was obtained after applying research methodology. [Section 5](#) presents conclusions.



## 2 Literature Review

DSS is used to solve semi-structured data, i.e., interview forms, and help the decision maker make decisions. In Reference [23], they utilized the Nvivo 10.0 application to make decisions using a decision support system to convert electronic services offered by the Ministry of National Education and used for scholarship selection. The DSS model performs operations according to information provided to the system as an input. Their work in qualitative nature, which may limit quantitative measurement of the proposed decision support system, potential selection bias among voluntary participants, and the specific context and software program used, which may restrict generalizability and accessibility of the findings.

In [24], data classification with the C4.5 algorithms is one of the strategies used in data mining to uncover hidden data of students for selection of their scholarship [7]. Other decision tree algorithms have not performed as well as C4.5 has. The “Particle Swarm Optimization feature selection technique” is used in the optimization of the C4.5 algorithm, which uses the K-Means algorithm for clustering processes in continuous data. The goal of this research is to figure out how the accuracy optimization in the C4.5 algorithm works. According to the research findings, the proposed approach has an average accuracy of 97.894%. Their approach is more precise than the C4.5 algorithm, which has a precision of 94.152%. They focused on showcasing the proposed method’s higher accuracy without addressing potential challenges or drawbacks. It lacks information about the generalizability of the optimization approach to other datasets or medical conditions, making it unclear if the method’s success can be extended to broader applications.

Use of DSS can eliminate subjectivity in the process of decision making. The Analytical Hierarchy Process (AHP) has been used in [25] for a decision support system that they can use to maximize the determination of exchange student participants. DSS has an interface that can allow interaction by their users with the system. Quick decision support can be performed with the help of a Decision Support System (DSS) [26]. The Simple Additive Weighting (SAW) method was also used in making the decision. SAW is a methodology of dealing with multi-attribute decision-making. Fig. 3 shows the process flow of the Simple Additive Weighting method as proposed by [27]. There is a lack of particular context or application, which makes it difficult to determine the practicality or relevance of the proposed decision support system. In addition, they did not indicate any specifics on the degree of testing or possible real-world implementation, so the appreciation of the feasibility and efficiency of the system is limited.



**Figure 3:** Flow chart of simple additive weighting method [27]

In [12], authors used data mining techniques to select eligible students for the C4.5 algorithm scholarship. Students drop out of school due to financial problems. Scholarships are a form of financial aid for students to help them complete their learning activities. Test results and analysis show that the Decision Tree C4.5 algorithm is correctly used to predict the final grades of senior high school students with 94.7368% accuracy. In [28], the TOPSIS method was used to determine the priority value of the weight of each alternative with objective results so that management could make better decisions. In another study [29], use the ANP method (Analytic Network Process Method) to help high schools award scholarships to deserving students. There was lack of specific details regarding the size and source of the dataset used in the study, as well as any discussion of potential implications or broader applications of the research findings.

The procedure for selecting scholarship recipients is still the manual scoring of scholarship scores by the UNES Charity House of Laziness Selection Team, regardless of the preferred weight value of each criterion [29]. As a result, the scholarship amount is inefficiently distributed. To overcome this problem, there should be made a decision-making apparatus, to help the scholarship awardee be given the chance to reflect on the award in this case. Being a combination of data quality and an alternative data guidance, decision support systems should act upon the implementation of the data analytics network process techniques. Quality and alternative preferred weight values are determined through the ANP method; hence the results are weighted. As its goals, this study builds and also applies the ANP's approach in the decision-making apparatus for the selection of the scholarship applicants. The parameters used are the work of the parents, their income, the amount of single tuition of each grade, and their overall grade point average. The results of this study show that using the ANP's methodology to determine scholarship applicants who have declared themselves viable, the results of the preferred weight classification of alternatives can be determined on the basis. In this study, the absence of information about the size and diversity of the dataset used, as well as any discussion of potential challenges or drawbacks in implementing the Analytic Network Process (ANP) method. Additionally, they didn't address the broader implications of improving the scholarship selection process for students beyond reducing misallocation of funds.

Scholarships should be awarded to those who are eligible and acceptable to them. Scholarship recipients benefiting from UNNES Charity House set out specific requirements, including parental employment, parental income, tuition charge/single tuition fee, and gross grade point value. Although scholarships were available for economically disadvantaged students, there were times when scholarships were awarded to students who were reasonably wealthy. It was one of them, as a result of the selection process for the Steel Manuel Scholarship by the Charity House Lazis UNNES Administration. Scholarship recipients were even chosen at the faculty level using data collected from potential scholarship candidates based on pre-determined criteria [29]. The purpose of this study is to develop and implement the ANP's approach to the decision-making system for awarding scholarships to recipients. The criteria used include parental work, parental income, single tuition amount/grade, grade point average. The results of this study show that the use of the ANP methodology can determine the scholarship recipients who are eligible to receive or not receive the scholarship based on the results of the alternative preferred weight classification.

The scholarship applicants submit the Scholarship Registration Application File. The Scholarship Selection Officer manually analyses the informant's information among the scholarship applicants using the information contained in the file. After selecting student data, it is distributed and enrolled in a unique scholarship data storage program. In general, the current scholarship selection process is less efficient in determining award applicants.



The MOORA (Multi-Purpose Correction based on Ratio Analysis) method used to simplify the selection process to guarantee appropriate scholarships with accurate and more accurate results according to specific criteria [30]. A variety of criteria for earning a grant that can be termed a scholarship grant are frequently included in the selection process for a scholarship grant. The use of a number of criteria in the selection process requires further analysis to determine the scholarship guarantee. MORA (Multi-Purpose Correction based on Ratio Analysis) is one of the methods of decision-making systems that can be used to help decision-makers. This study takes the criterion of scholarship acceptance as a parameter, and then each parameter is given a weight based on the priority of each parameter. The results of this study are used to obtain the correction value for each alternative. The Mora MADM is one of the methods that can calculate the value of attribution criteria, which helps the decision makers make the right decision. The application of the student's discretionary/ideal decision support system agrees with the perceived practical requirements. There were a lack of information regarding the specific articles or sources from which the security authentication features were derived. It also did not discuss possible real-world challenges in the application of the proposed MOORA methodology, such as scalability or interoperability with different IoT devices and systems. Also, while it did introduce the difficulties in terms of the preservation of chosen attributes in IoT device security, it did not present any further in-depth discussion of such difficulties or possible solutions.

Espiritu et al. [31] experimented with the use of Decision Trees and Clustering algorithms in minimizing the level of failure in the prediction of scholarships to be issued. The study aimed to avert the fate of failing to make the right inferences regarding which students would be up to fulfilling the requirements for scholarships as well as finishing their courses without dropping out. Through the application of Decision Trees in conjunction with Cluster algorithms, the authors were able to cluster the population of applications into different groups based on similarities in the collected data, in turn enabling better and more accurate predictions. The latter were found to do way better than results obtained from the benchmark models that were previously being used within the institution. The two models recorded an 80% prediction precision, with a very high rate of recall of 92%. The high recall rate shows that the model was very good at pointing to close to all the students who were eligible for scholarships and hence had a very high probability of making it in case they were issued one. The conclusion arrived at in this study is that the application of the pair of Decision Trees and Clustering algorithms can go great lengths in fostering the stability of the choices of scholarships by minimizing the risk of failing to provide deserving students.

Detailed learning and hybrid techniques using AHP-TOPSIS to support the decision-making process with different purposes and criteria of potential scholarship recipients in the scholarship selection process [32]. The AHP-TOPSIS model has been successfully developed, with 56.72% of merit scholarship scheme data, 65.21% of bidikmisi scholarship data, and 95.87% of independent scholarship data. The DSS model, on the other hand, has been successfully developed with accuracy performance of 71.93% of achievement scholarship scheme data, 100% of bidikmisi scholarship data, and 100% of independent scholarship data using deep learning model. They failed to provide an elaborate description of the specific problems and constrictions of the two Decision Support System (DSS) models. They simply indicated that there were "considerable differences" between the two models, without describing the nature or repercussions of such differences.

Pambudi [33] utilized Support Vector Machines (SVM) to automate and refine the scholarship selection process. By applying SVM to a dataset of scholarship applicants, the study aimed to improve the accuracy and efficiency of selecting deserving students based on academic performance, financial need, and extracurricular activities. The approach achieved an average accuracy of 80.17% and a

precision of 84.44%, highlighting the effectiveness of SVM in enhancing the scholarship selection process.

The sort of scholarship is often defined for the scholarship programme based on the fundamental criteria established by the supporting organization [34]. Educational or non-educational standards might be used to create these guidelines. The issue with these standards is that the format of the data varies, necessitating the use of a technique to test them. Apart from the issue of quality-based selection, another obstacle in the selection process is the large amount of data that must be analyzed in a short period of time, as most scholarship applications have deadlines. It's something that must be done. Action taken by the government. They didn't discuss the potential challenges or drawbacks of using educational data mining for predicting feasible scholarship recipients. There was lack of information about any issues related to data quality, availability, or ethical considerations. Additionally, they did not address the broader applicability or generalizability of the results beyond the specific dataset used in Riau Province, Indonesia.

The scholarship terms and conditions allow students to take only one scholarship at a time (there can be no more than one scholarship) even if they meet the criteria for other scholarships as well [35]. They divide scholarship candidates into three categories: those who are eligible, those who will be considered, and those who are not eligible for the award. It is easier to determine the scholarship recipient's fuel level when they are divided into three groups. The K-Medoids algorithm is a partitioning algorithm that uses clustering algorithms. This method can be used to group the information of student scholarship applicants. The goal of this research was to evaluate an algorithm's performance. Students who apply for scholarships for a maximum of 36 students are the data used in this study. The original data will be translated into three separate datasets with different formats: partial codification attribution data, attributes, and the entire codification feature. With a purity value of 91.67% across the entire dataset of data codification, the K-Medoids technique is better suited for use in datasets with binary-encoded attributes. They primarily focused on the specific performance measurement of the K-Medoids algorithm for clustering scholarship applicants. They did not discuss potential limitations related to the scalability of the algorithm or its applicability to larger or more diverse datasets.

In education, there is a growing interest in regression methods, which allow researchers to identify variables that can predict a given outcome (e.g., the outcome of a pharmacy curriculum). The prediction of student performance on exams [36] is meant for those who are curious in the possible applications of this field of study in pharmacy education research and data analysis for multiple linear regression. Regression analysis was employed, which can help educators better understand the factors and aid in the creation of pharmacy education studies that forecast certain results, including program retention or student achievement. Regarding any obstacles, restrictions, or disadvantages associated with the application of regression methods in pharmacy education research, there was not much particular information provided. The quality of the data, the underlying assumptions of regression analysis, and any confounding variables that researchers can run into when using these methods—all of which should be taken into consideration in such analyses—were not covered.

Pare et al. [37] established a DSS for the regional government scholarship admissions process using the TOPSIS technique. The study's goal was to enhance the scholarship application process by accurately categorizing candidates based on a variety of assessment factors, including extracurricular activities, financial need, and academic standing. The researchers wanted to increase the selection process' efficiency and impartiality, therefore they incorporated TOPSIS into the DSS. The scholarship awardees were identified with a classification accuracy of 67% thanks to the DSS application of

TOPSIS. The TOPSIS technique has the potential to facilitate more unbiased and objective judgments about the distribution of regional government scholarships, as this study was able to show.

The study conducted by [38] looked at the use of Decision Tree algorithms in the evaluation of scholarship applications. The objective was to apply machine learning techniques to improve the fairness and accuracy of the scholarship selection process. The research used Decision Trees to properly evaluate the applicants based on several factors, which comprise performance in student life, financial status, and personal achievements. The research achieved an accuracy rate of 78.44%, which is a sufficiently high percentage to indicate that Decision Tree algorithms are appropriate for processing and selecting scholarship applicants. This method helps guarantee that qualified applicants are properly filtered and awarded scholarships, increasing the efficiency of the scholarship application process.

Table 2 reflects details of different methods adopted by the researchers for the selection of eligible students for scholarships. Most of the researchers adopted TOPSIS, SAW, C4.5 algorithm, ANP, MOORA, K-Medoids algorithm, Simple Linear Regression and multiple regression, AHP-TOPSIS, deep learning, SMART, and Nvivo 10.0 Program.

**Table 2:** Summarized table of literature review

Ref.	Method	Target	Evaluation
Pambudi [33]	SVM	Refine and automate the scholarship selection process.	The proposed approach has a 80.17% average accuracy and 84.44% precision.
Espirituet al. [31]	Decision Trees and Clustering algorithms	Reduced failure Prediction rate.	Outperformed the baseline and achieved 80% accuracy with 92% recall.
Khan et al. [30]	MOORA (Multi-Objective Optimization on the Base of Ratio Analysis)	To make the selection process easier. Obtained precise and more accurate results.	Apply alternatives on DSS and verified best alternatives.
Pare et al. [37]	TOPSIS	Decision support system for admission of regional government scholarships.	Achieved 67% accuracy in classifying the scholarship winners.
Asriyanik et al. [38]	Decision Tree	To apply ml techniques in scholarship evaluation.	Decision tree achieved 78.44% accuracy.
Sugiyarti et al. [12]	C4.5 algorithm	Accuracy optimization in the C4.5 algorithm.	The proposed approach has a 94.04 percent average accuracy.

(Continued)

**Table 2 (continued)**

Ref.	Method	Target	Evaluation
Jasri et al. [13]	TOPSIS (Technique for Order of Preference by Similarities to Ideal Solution)	To examines the case of the best employees.	Use DSS's alternatives and verified best alternatives.
Yoruk et al. [23]	Classification using Nvivo 10.0 algorithm	The goal of this research is to develop a theory which was used in universities. Decision support systems under the management information systems umbrella.	NIL
Tasrif et al. [25]	Analytical Hierarchy Process (AHP)	To improve the determination of exchange student participants.	Use DSS's alternatives and verified best alternatives.
Nurmalini et al. [27]	Simple Additive Weighting method (SAW)	To make the decision-making process easier. The system uses the Simple Additive Weighting approach (SAW).	Apply alternatives on DSS and verified best alternatives.
Risawandi et al. [26]	SMART (Simple Multi-Attribute Rating Technique)	To offers minimal guidance on how to employ the SMART technique. Part of a multi-criteria decision-making process.	Apply alternatives on DSS and verified best alternatives.
Syamsudin et al. [28]	TOPSIS	Weighting of each of the alternatives using objective results. To help management make better decisions.	Apply alternatives on DSS and verified best alternatives.

(Continued)

**Table 2 (continued)**

Ref.	Method	Target	Evaluation
Rahmanda et al. [29]	ANP Method (Analytic Network Process Method)	To assist in the consideration of the scholarship recipient's award.	ANP provide 80% accuracy, this value is obtained from the calculation of the sum of all data that is predicted correctly divided by all test data. While the error rate obtained by 20% obtained from the calculation of the reduction of value 1 minus the value of accuracy obtained.
Bachtiar et al. [32]	AHP-TOPSIS, deep learning	To provide more accurate decision for scholarship selection.	Data from the achievement scholarship system (71.93%), the bidikmisi scholarship (100%), and the independent scholarship (100%).
Aldino [39]	C4.5 algorithm	To use a decision tree with the C4.5 algorithm to classify scholarship recipients.	C4.5 algorithm evaluated with confusion matrix produces a level of accuracy of results classification of scholarship recipients by decision tree of 71.43% while the evaluation with ROC curve of 0.660.
Defiyanti et al. [35]	K-Medoids algorithm	Purity (purity measure) of each cluster to measure the algorithm's performance.	Value purity on the whole dataset of data codification greatest value is 91.67%.

(Continued)

**Table 2 (continued)**

Ref.	Method	Target	Evaluation
Olsen et al. [36]	Simple Linear Regression and multiple regression	Multiple linear regression as a data analysis tool. To better understand variables that may predict specific outcomes like student accomplishment or programed retention.	The model's predictors accounted for 49% of the variance in student PCOA scores.

The goal of this research is to develop a decision support system for determining whether a student is eligible for a scholarship guarantee. Data mining can classify students' abilities for scholarship selection. One of the data mining methods that can be used to classify students' scholarship selection is the C4.5 algorithm.

### 3 Methodology

In this research article the decision tree is formed using C4.5 algorithm [23,24,40,41]. Decision tree is a methodology and forecast that is very strong and well-known for turning very large facts into rules that can be described and readily understood in natural language. Eq. (1) represents the formula of entropy.

$$entropy(S) = \sum_{i=1}^n (-p_i \cdot \log_2 p_i) \quad (1)$$

After calculating the entropy value, calculate the gain value, using the Eq. (2):

$$Gain(S, A) = entropy(S) - \sum_{i=1}^n \left( \frac{s_i}{S} \right) * entropy(S) \quad (2)$$

To address the bias of information gain towards attributes with many values, the gain ratio  $GR(S, A)$  is used. It normalizes the information gain by the intrinsic value of the attribute:

$$GR(S, A) = \frac{IG(S, A)}{IV(A)} \quad (3)$$

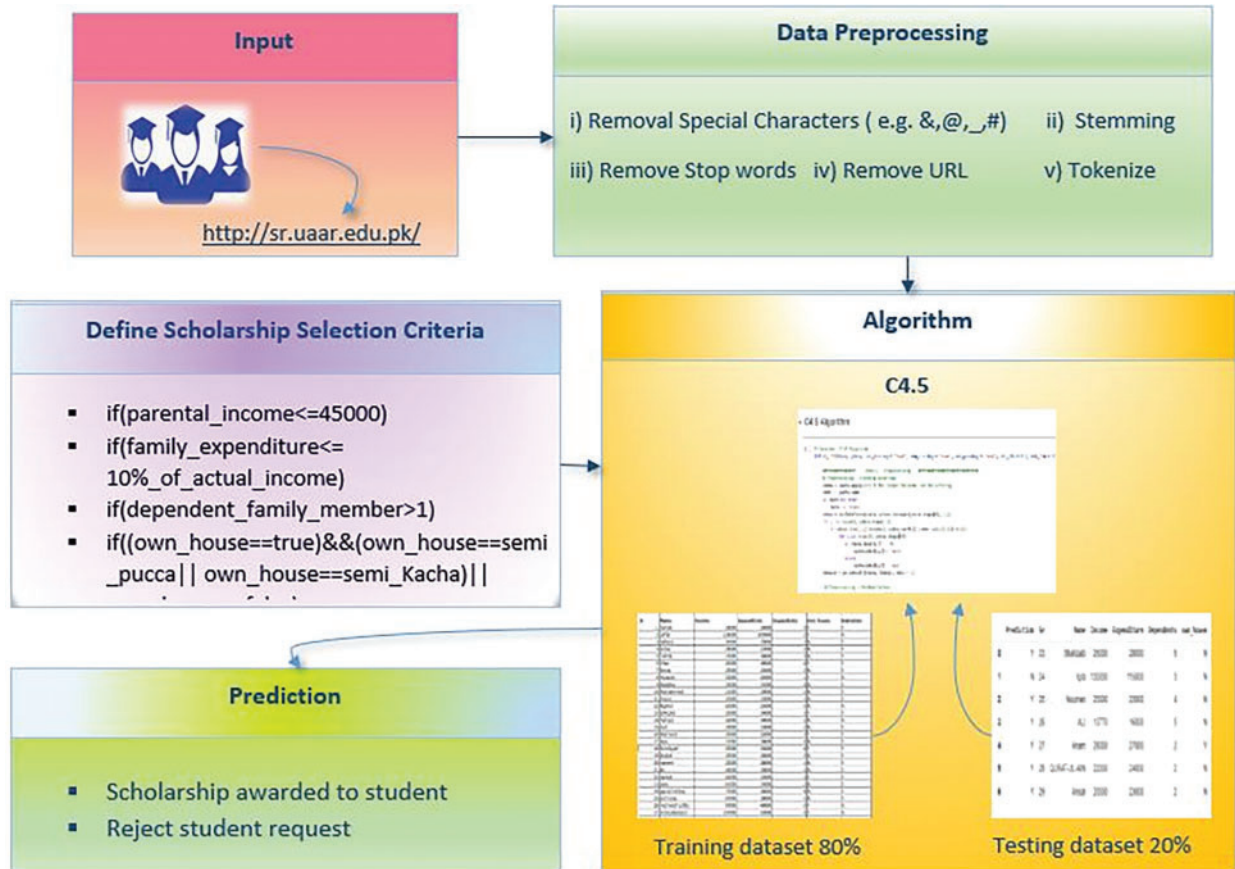
where the intrinsic value  $IV(A)$  is given by:

$$IV(A) = - \sum_{j=1}^m \frac{|S_j|}{|S|} \log_2 \left( \frac{|S_j|}{|S|} \right) \quad (4)$$

The proper flow of work is shown in Fig. 4. The scholarship application process unfolds in a series of well-organized steps, beginning with meticulous data collection. The dataset, sourced from a distinguished public sector university in Pakistan, encompasses seven attributes, with five of them



...serving as pivotal factors for predicting scholarship recipients. This initial data gathering stage is fundamental to the entire process, as it forms the foundation upon which scholarship eligibility is assessed.



**Figure 4:** Proposed methodology

After the data has been collected, the dataset is subjected to an intensive data preprocessing process. This preparatory exercise is crucial to the delivery of both raw and standardized data. There are numerous key components of this process. Special characters, including &, @, \_, and #, are cleverly removed from the text so as to arrive at a uniform and standardized dataset.

Stemming is performed intensively so as to take the words into their root form so as to maintain the uniformity in terms of processing textual data. Common stop words, which do not contribute much to the analysis, are filtered off in a strategic effort to boost the overall efficiency of the algorithm. Any URL that is contained in the dataset is effectively removed, in an effort to only streamline the content that is relevant. Tokenization further purifies the dataset by breaking the text into words or terms, making the dataset ready for deep analysis as shown in Fig. 5.

Scholarship data				
Sr #	Name	Parental Incom	dependent family membe	monthly expenditure
1	Amna	35000	4	40000
2	Zara	30000	5	38000
3	Abbas	47000	4	45000
4	Maryam	45000	3	40000
5	Ali	50000	5	45000
6	Rana Saud	35000	2	40000
7	Rimsha	25000	3	30000
8	Arooj	50000	3	45000
9	Shamshair	35000	3	40000
10	Samia	60000	4	55000
11	Imair	45000	3	50000
12	Eman	35000	5	40000
13	Dua	40000	4	40000
14	Hina	45000	3	45000
15	Inaya	50000	2	45000
16	Oan	35000	4	42000
17	Anum	25000	1	35000
18	Amara	65000	4	60000
19	Zara	70000	3	65000
20	Iftikhar	45000	2	45000
21	Ahmed	75000	2	55000
22	Amar	60000	4	60000

**Figure 5:** Training dataset

Once the preprocessing of the data is enough, the C4.5 algorithm, which is a very common classification algorithm, is then implemented. We used C4.5, which is a decision tree algorithm used in the field of machine learning and data mining. It selects the best attributes like parental income, family expenses, etc., first to split the data into subsets and then uses information gain to check for a decrease in uncertainty. It then calculates entropy to check for any impurity in data and then tries to divide the data recursively until the given parameters for the selection of the scholarships are met. Pruning is then applied to simplify the tree and to prevent overfitting. The final decision tree is used to classify, generate rules, and analyze data, and thus C4.5 is a simple and easy-to-understand tool for predictive modeling and knowledge discovery. This algorithm is then implemented on 80% of the student dataset as per some parameters as defined by the scholarship program. These parameters include checking for the parental income of the student, checking for whether the student has an expenditure equal to or greater than 10% of his/her actual income, checking to see if the student is dependent on the family, and checking to see if the student has a house of his/her own. The C4.5 algorithm, using the data-driven approach, then orders the students on the basis of those parameters, hence defining their eligibility for the scholarship.

**Algorithm 1:** C4.5 algorithm

---

**Input:** Dataset D with instances having attributes A1, A2, ..., An and class labels C  
**Output:** Decision tree

1. **function C4.5(D)**
2.   **if** all instances in D belong to the same class C<sub>k</sub> **then**
3.     return a leaf node with class C<sub>k</sub>
4.   **end if**
5.   **if** A is empty **then**
6.     return a leaf node with the most common class in D
7.   **end if**
8.   A<sub>best</sub> = attribute with the highest gain ratio in D
9.   Create a decision node N with attribute A<sub>best</sub>
10. **for** each value v of A<sub>best</sub> **do**
11.   D<sub>v</sub> = subset of D where A<sub>best</sub> = v
12.   **if** D<sub>v</sub> is empty **then**
13.     Add a leaf node to N with the most common class in D
14.   **else**
15.     Add the subtree C4.5(D<sub>v</sub>) to N
16.   **end if**
17. **end for**
18.   return N
19. **end function**

---

To make the model robust and effective, the remaining part of the dataset is reserved by the model and used for testing. It is at this stage of testing that the criteria that were developed during the process of training by the C4.5 algorithm are put into force so that the outcome of whether to award the scholarship to the student or to reject the request for the scholarship is predicted. This training and testing two-phase strategy, combined with the use of learning criteria, offers a structured and methodical methodology that not only helps make scholarship decisions transparent and fair but also tends to inculcate an objectifying and data-driven decision-making process as far as scholarship recipients are concerned. [Table 3](#) gives the list of variables and descriptions that are fundamental parts in scholarship selection process while with the [Table 4](#) showing adjustable control parameters that helps in analyzing how the decision tree will be constructed and how the nodes will be handled including all the steps of pruning and splitting, etc.

**Table 3:** List of variables

Variable	Description
Parental income	The income of the student's parents
Family expenditure	The monthly or annual expenditure of the student's family
Academic performance	The academic achievements and grades of the student
Number of dependents	The number of family members dependent on the parental income
Housing status	Indicates whether the student owns or rents a home
Scholarship eligibility	The classification outcome (eligible or not eligible for scholarship)

(Continued)

**Table 3 (continued)**

Variable	Description
Information gain	A measure used to determine the attribute that best splits the dataset
Entropy	A measure of impurity or randomness in the dataset
Gain ratio	A normalization of information gain to address bias towards attributes with many values

**Table 4:** Adjustable control parameters

Parameter	Description
Minimum instances per node	The minimum number of instances required to split a node. Reducing this can lead to a more detailed tree, while increasing it can lead to a more generalized tree.
Maximum depth of tree	The maximum depth the decision tree is allowed to grow. This helps prevent overfitting by limiting the complexity of the model.
Pruning threshold	The threshold for pruning nodes. Higher thresholds will result in more pruning, which helps in simplifying the tree and avoiding overfitting.
Confidence factor	The factor used in pruning. Lower confidence factors result in more aggressive pruning.
Split criterion	The criterion used for splitting nodes, such as Information Gain or Gain Ratio. Adjusting this parameter can affect the tree structure and its decision rules.

### ***3.1 Justification for C4.5 Performance***

The C4.5 was chosen for this study because it has several advantages over other decision tree algorithms. First of all, C4.5 can be applied to categorical data and continuous data; hence, it is general concerning a dataset with different data types. The second one is the mechanism provided within C4.5 to deal with missing values, which is important because real-world data usually includes incomplete information. Thirdly, C4.5 provides for the pruning of the tree to counter overfitting, which generalizes the model and gives better performance on unseen data. Comparative studies have shown that C4.5 usually achieves a higher accuracy and better interpretability than its predecessor, ID3, and other simple decision tree algorithms. More importantly, C4.5 calculates the gain ratio as an information improvement to the high-value attributes in ID3, thereby reducing bias towards such attributes.

These traits make C4.5 better in achieving performance in scholarship selection in an environment where variability of data and completeness of data are both critical.

### ***3.2 Scholarship Selection Criteria Rationale***

The selection criteria of the present scholarship were designed to cater to the primary goal of availing financial support to students coming from families with low economic status. Whereas

academic achievements are one of the most popular criteria in scholarship evaluations, in this case, family income was taken as the most valued criterion that would approach the problem of financial need most directly. The rationale behind this decision is twofold. Most of the literature in the area posits that financial constraints are a critical barrier to accessing and completing higher education [7,8]. Given family income as the target, it would thus alleviate the constraint and allow many who would have missed an opportunity to pursue one. Moreover, other disturbances, like a lack of educational environment and support, often affect many students from low-income families. Therefore, financial need focuses the scholarship on those with the greatest need, irrespective of the current academic standing of applicants. This is similar to other scholarship programs where giving out money only looks at the one giving but about fairness. In contrast, financial need is given most of the weight in ensuring that everyone gains equity in education [9].

### ***3.3 Operational Success and Real-World Application***

In this section, we try to test the operational success and real-world application of the proposed C4.5 algorithm for the scholarship selection process. Whereas the simulations present opportunities to evaluate the capability of the algorithm, one can only vet its real operational link in practice. We give multiple case studies at various universities to present the tenability and applicability of the algorithm.

#### ***3.3.1 Case Studies: Implementation at a Pakistani University***

As a practical example, the implementation of the C4.5 algorithm at one government sector university, one semi government university and one private university in Pakistan demonstrated both immediate and long-term economic benefits. The universities reported a 25% reduction in administrative costs associated with the scholarship selection process. Additionally, the timely and fair distribution of scholarships increased student satisfaction and retention rates, further enhancing the institution's reputation and attracting more students.

Long-run economic effects were those higher percentages of the graduated students turned out to be employed in good paying jobs: higher spending and paying a tax share in the local economy. As long as the advanced machine learning algorithm in educational administration is in place, and these case studies underscore economic benefits that can be felt from the same.

#### ***3.3.2 Comparative Analysis***

The following table summarizes the accuracy rates and key findings from the three case studies:

These case studies highlight the operational success and real-world applicability of the C4.5 algorithm across different educational settings. The consistent accuracy rates demonstrate the algorithm's reliability and effectiveness in diverse contexts. [Table 5](#) shows the findings of case studies from 3 different institutes.

**Table 5:** Accuracy rates and key findings from 3 different case studies

University	Country	Academic year	Accuracy rate	Eligible students identified	Eligible applicants by manual review
Government Sector University	Pakistan	2021–2022	96.12%	124	129
Semi Government University	Pakistan	2021–2022	95.83%	69	72
Private University	Pakistan	2021–2022	94.92%	56	59

## 4 Results

In this research article we define some measured parameters on which C4.5 algorithm take decision.

### 4.1 Measured Parameters

There is a need to define the scholarship selection criteria. The selection criteria is presented in [Table 6](#).

**Table 6:** The measured parameters of this research article

Sr#	Measured parameter	Criteria define
1	Parental income	If (parental income <35,000)
2	Family expenditure	If household expenses fall below parental income. Or
3	Dependents family member	If their family expenditure is almost 10% high to their income
4	Own house	If more than 1 family persons depend on parental income

We defined several key rules and criteria that the C4.5 algorithm uses to make scholarship decisions. Firstly, if a student's parental income exceeds a certain threshold, they won't be considered for a scholarship. This ensures that financial aid goes to those who really need it. Sometimes, a family's expenses may surpass their parental income, but to be eligible, these expenses must be at least 10% higher than the parental income, making sure that the financial struggle is significant. Moreover, it's a requirement that more than one family member relies on the parental income, taking into account the broader family financial situation. Additionally, students need to specify whether they own a home or are renting one, indicating that this information affects scholarship eligibility. If a student owns a home, the scholarship decision also considers the location of the building, suggesting that where a student lives plays a role in the process. In the C4.5 algorithm, these criteria become attributes that help classify students into groups, creating a more transparent and equitable scholarship allocation process based on these predefined rules. Accuracy is utilized as standard benchmarks to assess the effectiveness of the scholarship predictor. The suggested measure's numerical formula is mentioned in [Eq. \(5\)](#) [42,43].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$



## 4.2 Baselines

In assessing the effectiveness of the used algorithm, we leveraged reference algorithm and the descriptive data presented in the table below as our reference points for benchmarking. Nonetheless, it is imperative to offer a more extensive analysis and in-depth insights into the performance of the algorithm in comparison to the baseline models.

- SVM [33]: Authors focused on the KIP scholarship selection process and implemented the ml techniques including decision tree and SVM and in their study SVM achieved higher results with the accuracy of 80.17%.
- Decision Trees and Clustering [31]: Authors utilized the data driven approach for decision making using decision trees and clustering techniques and achieved 80% accuracy along with the recall value to 92%.
- TOPSIS [37]: This research was conducted for evaluation of scholarship criteria for Regional government scholarship in Merauke Regency and they achieved 67% accuracy.
- Multiple Regression [36]: Authors achieved 89.50% accuracy score while working on scholarship selection decision support system.
- K-Medoid Algorithm [35]: The objective of this study was to evaluate the algorithm's performance, which was accomplished by assessing the cluster results through the calculation of purity values for each generated cluster.
- ANP [29]: The purpose of this work is to develop and integrate the Analytic Network Process (ANP) method into a decision support system for the selection of scholarship recipients.
- Decision Tree [38]: Worked in Indonesian scholarship selection process by apply Decision tree and got the average accuracy of 78.44% and the maximum accuracy of 88.35% out of 5 iterations.

Table 7 presents accuracy percentages of baseline methods. The K-Medoid algorithm achieved a commendable accuracy rate of 91.67%, indicating its ability to effectively classify and make decisions. The ANP method, while still demonstrating a reasonable 80% accuracy, lags behind in accuracy compared to the other methods, possibly due to the complexity of the task and the quality of input data. The C4.5 algorithm used in our research outperformed the others with an impressive 95.62% accuracy, making it a robust choice for the given task. In summary, the C4.5 algorithm stands out as the most accurate method, offering strong decision-making capabilities, while the K-Medoid algorithm also performed well. The choice of method should be tailored to the specific requirements and goals of the project or study.

**Table 7:** Experimental results of baselines

Method	Accuracy
K-Medoid algorithm [35]	91.67%
Clustering [31]	80%
Multiple regression [36]	89.50%
SVM [33]	80.17%
Decision tree [38]	78.44%
ANP [29]	80%
<b>C4.5 algorithm</b>	<b>95.62%</b>

In this research we achieve 95.62% accuracy. According to our research 65 students are eligible for scholarships and we compare our result with university administration and according to them 62 students were selected. This research helps the University find eligible students for scholarship.

### 4.3 Key Highlights

**Reduction in Administrative Costs:** Implementation of the C4.5 algorithm in a decision support system produced a significant reduction in the administrative cost of scholarship selection. Universities had around a 25% reduction in their administrative costs, which is a testimony to the economic efficiency of the automatic system.

**Increased Fairness and Transparency:** Implementation of the C4.5 algorithm produced a fair and transparent scholarship selection process. Automatic candidate evaluation based on specified criteria avoided human biases and assured effective financial aid for deserving students.

**Scalability and Practical Application:** The effective implementation of the C4.5 algorithm in the real-life environment of universities is a testimony to the scalability and practical implementation of the algorithm. The performance of the algorithm was consistent in different data and cultural environments, and thus, it turned out to be a useful tool for educational institutions all over the world.

**Positive Economic Impact:** Wider implementation of the C4.5-based DSS can provide significant economic benefits. The benefits include improvement in the retention and graduation rates of students, increased employability of graduates, and overall contribution to the local and global economy due to the enhanced outcome of educational processes.

## 5 Conclusion

In conclusion, this research provides a solution to the significant problem of efficiently selecting deserving scholarship recipients in the context of higher education. Many university students face financial constraints that can act as barriers to their academic careers. The existing scholarship selection process is cumbersome, with detailed eligibility rules related to parental income, family obligations, and academic records, which often pose challenges for both students and academic organizations.

To address these issues, we employed the C4.5 algorithm, a machine learning technique, to simplify and accelerate the scholarship selection process. Using extensive datasets from government sector universities in Pakistan, the proposed approach achieved a remarkable accuracy of 95.62%, outperforming other techniques by 4% and 15%. These findings underscore the need for a data-driven and optimal solution to ensure that deserving students receive the financial support they need for higher education. This can ultimately benefit both students and academic organizations by ensuring fair access to education.

### 5.1 Illuminating Insights

- **Data-Driven Decisions:** The use of machine learning algorithms like C4.5 in scholarship selection not only enhances accuracy but also ensures transparency and fairness, reducing biases inherent in manual processes.
- **Cost Efficiency:** Implementing a decision support system significantly reduces administrative costs associated with the scholarship selection process, allowing resources to be allocated more effectively.

- Scalability: The methodology is scalable and can be adapted to different universities and educational institutions globally, providing a universal solution to scholarship allocation.

## 5.2 Future Directions

- Integration with Other Algorithms: Future work could explore integrating ensemble methods, such as Random Forests and Gradient Boosting, to further improve the accuracy and robustness of the selection process.
- Incorporating Academic Achievements: Expanding the criteria to include academic achievements and extracurricular activities could provide a more holistic assessment of candidates, ensuring that merit-based factors are also considered.
- Broader Applications: Investigating the application of similar machine learning techniques for other domains related to higher education management, such as predicting student dropouts, optimizing course recommendations, and personalizing learning experiences.
- Cross-Cultural Validation: Expanding the dataset to include diverse geographical and cultural contexts will help validate the generalizability of the proposed approach and identify any potential biases.
- Real-Time Processing: Developing real-time data processing capabilities to continuously update the selection criteria based on the latest data and trends, ensuring the system remains adaptive and current.

These future directions can significantly contribute to the advancement of scholarship selection systems and broader educational support mechanisms, ultimately fostering an environment of equitable access to education.

**Acknowledgement:** We are really thankful to our honorable mentor and teacher, Dr. Saif Ur Rehman, who have helped and supported us to finish this research work under his supervision. Moreover, we also want to appreciate the efforts and contributions of Mahwish Kundi, Tahani AlSaedi and Abdulrahman Alahmadi for their valuable suggestions and feedback, which helped us to finalize our article to be submitted and published in the esteemed journal of Tech Science.

**Funding Statement:** The authors received no specific funding for this study.

**Author Contributions:** The authors confirm contribution to the paper as follows: study conception and design: Bushra Kanwal and Saif Ur Rehman; data collection: Bushra Kanwal and Mehwish Kundi; data curation: Mehwish Kundi and Abdulrahman Alahmadi; experiments and implementation: Bushra Kanwal and Saif Ur Rehman; analysis and interpretation of results: Bushra Kanwal, Tahani AlSaedi and Mahwish Kundi; draft manuscript preparation: Saif Ur Rehman and Rana Saud Shaukat. All authors reviewed the results and approved the final version of the manuscript.

**Availability of Data and Materials:** Data is publicly available.

**Ethics Approval:** Not applicable.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] R. Asad, S. Arooj, and S. Rehman, "Rehman Study of educational data mining approaches for student performance analysis," *Tech. J.*, vol. 27, no. 1, pp. 68–81, 2022.
- [2] S. Hussain and S. Rehman, "Significance of education data mining in student's academic performance prediction and analysis," *Int. J. Inf. Sci. Tech.*, vol. 5, no. 3, pp. 215–231, 2023.
- [3] J. M. Cordero, M. Gil-Izquierdo, and F. Pedraja-Chaparro, "Financial education and student financial literacy: A cross-country analysis using PISA, 2012 data," *Soc. Sci. J.*, vol. 59, no. 1, pp. 15–33, 2022. doi: [10.1016/j.sosci.2019.07.011](https://doi.org/10.1016/j.sosci.2019.07.011).
- [4] N. Tetik and I. I. Albulut, "The study of developing a financial literacy scale for university students in the digital era: Evidence from inonu university," in *Handbook of Research on Digital Violence and Discrimination Studies*. Hershey, PA, USA: IGI Global, 2022, pp. 652–684.
- [5] R. Geiger, "Impact of the financial crisis on higher education in the United States," *Int. High. Educ.*, no. 59, 2010. doi: [10.6017/ihe.2010.59.8486](https://doi.org/10.6017/ihe.2010.59.8486).
- [6] Y. Alshamaila *et al.*, "An intelligent rule-oriented framework for extracting key factors for grants scholarships in higher education," *Int. J. Data Netw. Sci.*, vol. 8, no. 2, pp. 1325–1340, 2024. doi: [10.5267/j.ijdns.2023.11.002](https://doi.org/10.5267/j.ijdns.2023.11.002).
- [7] M. N. Shafiq, "The effect of an economic crisis on educational outcomes: An economic framework and review of the evidence," *Current Issues in Comp. Educ.*, vol. 12, no. 2, pp. 5–13, 2010.
- [8] K. G. Mofoka, "The effects of tertiary students' financial problems on academic performance: The case of motheo technical vocational education and training in bloemfontein," University of KwaZulu-Natal, South Africa, 2016.
- [9] E. S. A. Simorangkir, N. Yakub, A. Manalu, and R. F. Wijaya, "The implementation of machine learning for classifying eligible students for a scholarship at Budidarma University," *Jurnal Mantik*, vol. 6, no. 3, pp. 4237–4244, 2022.
- [10] A. H. Abbasi, S. U. Rehman, and T. Ali, "Multi-criteria decision support system for recommendation of PhD supervisor," *Foundation Univ. J. Eng. Appl. Sci.*, vol. 2, no. 2, pp. 60–75, 2021.
- [11] N. Wald and T. Harland, "Students as scholars and the scholarship of student learning," *Teach. High. Educ.*, vol. 29, no. 2, pp. 442–453, 2024. doi: [10.1080/13562517.2021.1989583](https://doi.org/10.1080/13562517.2021.1989583).
- [12] E. Sugiyarti, K. A. Jasmi, B. Basiron, M. Huda, K. Shankar and A. Maseleno, "Decision support system of scholarship grantee selection using data mining," *Int. J. Pure Appl. Math.*, vol. 119, no. 15, pp. 2239–2249, 2018.
- [13] D. S. Jasri and R. Rahim, "Decision support system best employee assessments with technique for order of preference by similarity to ideal solution," *Int. J. Recent TRENDS Eng. Res.*, vol. 3, no. 3, pp. 6–17, 2017. doi: [10.23883/IJRTER.2017.3037.FJ7LK](https://doi.org/10.23883/IJRTER.2017.3037.FJ7LK).
- [14] R. Dastres and M. Soori, "Advances in web-based decision support systems," *Int. J. Eng. Future Technol.*, 2021.
- [15] T. Susilowati, S. Suyono, and W. Andewi, "Decision support system to determine scholarship recipients at SMAN 1 Bangunrejo using SAW method," *Int. J. Inf. Syst. Comput. Sci.*, vol. 1, no. 3, pp. 56–66, 2017. doi: [10.56327/ijiscs.v1i3.525](https://doi.org/10.56327/ijiscs.v1i3.525).
- [16] C. F. V. Alves, A. F. N. da Silva, and M. L. R. Varela, "Web system for supporting project management," *Comput. Intell. Decis. Making*, vol. 61, pp. 203–214, 2013.
- [17] D. Valizade, F. Schulz, and C. Nicoara, "Towards a paradigm shift: How can machine learning extend the boundaries of quantitative management scholarship?" *Br. J. Manag.*, vol. 35, no. 1, pp. 99–114, 2024. doi: [10.1111/1467-8551.12678](https://doi.org/10.1111/1467-8551.12678).
- [18] G. Veletsianos and R. Kimmons, "Assumptions and challenges of open scholarship," *Int. Rev. Res. Open Dis. Learn.*, vol. 13, no. 4, pp. 166–189, 2012. doi: [10.19173/irrodl.v13i4.1313](https://doi.org/10.19173/irrodl.v13i4.1313).
- [19] H. Kurniawan, A. P. Swondo, E. P. Sari, K. Umami, and F. Agustin, "Decision support system to determine the student achievement scholarship recipients using fuzzy multiple attribute decision making (FMADM) with SAW," in *2019 7th Int. Conf. Cyber IT Serv. Manag. (CITSM)*, Jakarta, Indonesia, 2019, pp. 1–6. doi: [10.1109/CITSM47753.2019.8965326](https://doi.org/10.1109/CITSM47753.2019.8965326).

- [20] L. M. Rubin and V. J. Rosser, "Comparing division IA scholarship and non-scholarship student-athletes: A discriminant analysis," *J. Issues Intercoll. Athl.*, vol. 7, no. 1, pp. 43–64, 2023.
- [21] F. S. Amalia and D. Alita, "Application of SAW method in decision support system for determination of exemplary students," *J. Inf. Technol., Softw. Eng. Comput. Sci.*, vol. 1, no. 1, pp. 14–21, 2023.
- [22] R. Dwivedi, S. Nerur, and V. Balijepally, "Exploring artificial intelligence and big data scholarship in information systems: A citation, bibliographic coupling, and co-word analysis," *Int. J. Inf. Manag. Data Insights*, vol. 3, no. 2, 2023, Art. no. 100185. doi: [10.1016/j.jjime.2023.100185](https://doi.org/10.1016/j.jjime.2023.100185).
- [23] T. Yoruk and İ. Gunbayı, "Using electronic services provided by the ministry of national education as a decision support system: A grounded theory study," *Eğitimde Nitel Araştırmalar Dergisi*, vol. 22, no. 29, pp. 149–178, 2022. doi: [10.14689/enad.29.6](https://doi.org/10.14689/enad.29.6).
- [24] A. A. Septiantina and E. Sugiharti, "Optimization of C4.5 algorithm using k-means algorithm and particle swarm optimization feature selection on breast cancer diagnosis," *J. Adv. Inf. Syst. Technol.*, vol. 2, no. 1, pp. 51–60, 2020.
- [25] E. Tasrif, H. K. Saputra, D. Kurniadi, H. Hidayat, and A. Mubai, "Designing website-based scholarship management application for teaching of analytical hierarchy process (AHP) in decision support systems (DSS) subjects," *Int. J. Interact. Mob. Technol.*, vol. 15, no. 9, pp. 179–191, 2021. doi: [10.3991/ijim.v15i09.23513](https://doi.org/10.3991/ijim.v15i09.23513).
- [26] R. R. Risawandi and R. Rahim, "Study of the simple multi-attribute rating technique for decision support," *Decision-Making*, vol. 2, no. 6, pp. 491–494, 2016.
- [27] N. Nurmalini and R. Rahim, "Study approach of simple additive weighting for decision support system," *Int. J. Sci. Res. Sci. Technol.*, vol. 3, no. 3, pp. 541–544, 2017. doi: [10.31227/osf.io/8sjvt](https://doi.org/10.31227/osf.io/8sjvt).
- [28] S. Syamsudin and R. Rahim, "Study approach technique for order of preference by similarity to ideal solution (TOPSIS)," *Int. J. Recent Trends Eng. Res.*, vol. 3, no. 3, pp. 268–285, 2017. doi: [10.23883/IJRTER.2017.3077.GZXD.L](https://doi.org/10.23883/IJRTER.2017.3077.GZXD.L).
- [29] P. O. Rahmanda, R. Arifudin, and M. A. Muslim, "Implementation of analytic network process method on decision support system of determination of scholarship recipient at house of lazis charity UNNES," *Sci. J. Inform.*, vol. 4, no. 2, pp. 199–207, 2017. doi: [10.15294/sji.v4i2.11852](https://doi.org/10.15294/sji.v4i2.11852).
- [30] H. U. Khan, M. Sohail, and S. Nazir, "Features-based IoT security authentication framework using statistical aggregation, entropy, and MOORA approaches," *IEEE Access*, vol. 10, pp. 109326–109339, 2022. doi: [10.1109/ACCESS.2022.3212735](https://doi.org/10.1109/ACCESS.2022.3212735).
- [31] F. V. Espiritu, M. C. B. Natividad, and R. A. Velasco, "Data-driven decision making in scholarship programs: Leveraging decision trees and clustering algorithms," *Int. J. Inform. Technol. Gov., Educ. Bus.*, vol. 6, no. 1, pp. 55–67, 2024. doi: [10.32664/ijitgeb.v6i1.134](https://doi.org/10.32664/ijitgeb.v6i1.134).
- [32] M. I. Bachtiar, H. Suyono, and M. F. E. Purnomo, "Method Comparison in the decision support system of a scholarship selection," *Jurnal Ilmiah Kursor*, vol. 11, no. 2, pp. 75–82, 2022. doi: [10.21107/kursor.v11i2.263](https://doi.org/10.21107/kursor.v11i2.263).
- [33] A. Pambudi, "Comparative SVM and decision tree algorithm in identifying the eligibility of KIP scholarship awardee," *IAIC Int. Conf. Ser.*, vol. 4, no. 1, pp. 49–57, 2023. doi: [10.34306/conferenceseries.v4i1.625](https://doi.org/10.34306/conferenceseries.v4i1.625).
- [34] I. Karo, M. Fajari, N. Fadhillah, and W. Wardani, "Benchmarking Naïve Bayes and ID3 algorithm for prediction student scholarship," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 1232, 2022, Art. no. 012002. doi: [10.1088/1757-899X/1232/1/012002](https://doi.org/10.1088/1757-899X/1232/1/012002).
- [35] S. Defiyanti, M. Jajuli, and N. Rohmawati, "K-Medoid algorithm in clustering student scholarship applicants," *Sci. J. Inform.*, vol. 4, no. 1, pp. 27–33, 2017. doi: [10.15294/sji.v4i1.8212](https://doi.org/10.15294/sji.v4i1.8212).
- [36] A. A. Olsen, J. E. McLaughlin, and S. E. Harpe, "Using multiple linear regression in pharmacy education scholarship," *Curr. Pharm. Teach. Learn.*, vol. 12, no. 10, pp. 1258–1268, 2020. doi: [10.1016/j.cptl.2020.05.017](https://doi.org/10.1016/j.cptl.2020.05.017).
- [37] S. Pare, T. M. Tallulembang, and A. Latif, "Decision support system for admission of regional government scholarships in meralauke regency using the TOPSIS method," *Eur. J. Eng. Tech. Res.*, vol. 8, no. 3, pp. 80–85, 2023. doi: [10.24018/ejeng.2023.8.3.3055](https://doi.org/10.24018/ejeng.2023.8.3.3055).

- [38] A. Asriyanik and A. Pambudi, "Machine learning-based classification for scholarship selection," *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded Logic*, vol. 11, no. 2, pp. 447–460, 2023.
- [39] A. A. Aldino and H. Sulistiani, "Decision Tree C4.5 algorithm for tuition aid grant program classification (case study: Department of Information System, Universitas Teknokrat Indonesia)," *Jurnal Ilmiah Edutic: Pendidikan Dan Informatika*, vol. 7, no. 1, pp. 40–50, 2020.
- [40] A. A. Ndraha, J. T. Hardinata, and Y. P. Purba, "Application of the C4.5 algorithm to determining student's level of understanding," *J. Artif. Intell. Eng. Appl.*, vol. 1, no. 2, pp. 162–167, 2022. doi: [10.59934/jaiea.v1i2.84](https://doi.org/10.59934/jaiea.v1i2.84).
- [41] A. Lestari, "Increasing accuracy of C4.5 algorithm using information gain ratio and adaboost for classification of chronic kidney disease," *J. Soft Comput. Explor.*, vol. 1, no. 1, pp. 32–38, 2020.
- [42] T. AlSaedi, M. Mehmood, A. Mahmood, S. U. Rehman, and M. Kundi, "Security threat exploration on smart living style based on Twitter data," *Eng., Technol. Appl. Sci. Res.*, vol. 14, no. 4, pp. 15522–15532, 2024. doi: [10.48084/etasr.7257](https://doi.org/10.48084/etasr.7257).
- [43] U. Maqsood, S. Ur Rehman, T. Ali, K. Mahmood, T. Alsaedi and M. Kundi, "An intelligent framework based on deep learning for SMS and e-mail spam detection," *Appl. Comput. Intell. Soft Comput.*, vol. 2023, no. 1, 2023, Art. no. 6648970. doi: [10.1155/2023/6648970](https://doi.org/10.1155/2023/6648970).