Data Mining for Smarter Administration of TVET Institutes

Aashir Sheeraz[1], Yasir Ali Rasheed[1], Nazia Azim[2], Asif Khan[3], Sheeraz Ahmed[1]
[1]Department of Computer Science, Iqra National University, Peshawar, Pakistan
[2]Department of Computer Science, Abdul Wali Khan University, Mardan, Pakistan
[3]Government of KP Nowshera, Pakistan
*Correspondence: aashirsheeraz50@gmail.com, yasirkhan2@hotmial.com, n.azim@awkum.edu.pk, asif.khan43@yahoo.com, sheerazahmed306@gmail.com

Citation | Sheeraz. A, Rasheed. Y. A, Azim. N, Khan. A, Ahmed. A, "Data Mining for Smarter Administration of TVET Institutes", IJIST, Vol. 07, Issue. 04 pp 2573-2580, October 2025

Received | August 27, 2025 **Revised |** September 21, 2025 **Accepted |** September 23, 2025 **Published |** October 26, 2025.

Retaining the trainees is a major problem for the TVET institutes today. The trend of TVET education in Khyber Pakhtunkhwa province has improved in recent decades. Despite its high cultural barriers, resistance in women's education, and dropout rates, on the basis of annual admission, Khyber Pakhtunkhwa holds the 2nd position among all other provinces of Pakistan. In this research, we have tried to decrease the dropout ratio by enhancing the daily attendance of the trainees and improving their results. Monthly fee slip, date sheet, and results will be shared with the parents/guardians through SMS/WhatsApp. New TVET institutes will be able to check the trainee's educational record from the previous TVET institute. The data mining for smarter administration of TVET Institutes will be a mobile and web-based application and will keep a close relationship between parents, teachers, and administration of the TVET institute.

Keywords: TVET Education, Data-Mining, Smart Administration

Introduction:

TVET institutes are facing increasing difficulty in keeping trainees enrolled. To address this, they are implementing measures to help more trainees complete their programs and lower the dropout rate. The tinto model investigated and analyzed the attributes of trainees that contributed to their failure in the institute. These characteristics included their regional, cultural, social, familial, or educational backgrounds; socioeconomic level, psychological profile, and academic advancement were among these variables. The primary reason for completion is a trainee's social and academic incorporation into the educational institution. Bharadawaj and Pal discovered reasons impacting trainee dropout as the senior secondary test, annual family income, domicile, and family status. Meanwhile, other researchers employed enrolment data to predict dropout [1].

Numerous events, such as the September 11 attacks, the COVID-19 pandemic, and the 2005 Pakistan earthquake, caused significant loss of life and property. These disasters highlight the urgent need to enhance our capacity to prevent, respond to, mitigate, and recover from both natural and man-made calamities, while also improving our ability to identify the missing or deceased. Three essential components comprise this conceptual framework: data integration, data mining, and multi-criteria decision making (MCDM). The data mining component identifies relevant patterns through data mining techniques and is a distinct service delivery procedure for pre- and post-incident information management [2].

In academics, various assessments such as beginning of term evaluations, periodic tests, mid-term exams, and end of term assessments are used to measure trainees' learning progress. These results are then aggregated to generate overall termly trainee assessment reports [3]. The Internet and advances in technology have profoundly influenced education and learning. E-learning refers to the use of computer networks to deliver information and instruction to individuals through an intranet or the Internet. However, there are a number of challenges associated with e-learning, such as different learning styles and cultural differences [4].

In this research, a method for tracking trainee progress is proposed that can be used as the foundation for e-Learning systems. The proposed methodology aims to develop a comprehensive set of feedback indicators to help trainees better understand their performance and identify ways to improve it [5]. As a result of the COVID-19 catastrophe, governments worldwide have implemented public policies that include social isolation, self-quarantine, and social distancing, resulting in unprecedented economic and emotional implications globally. The Learning Management System (LMS) is the technical term for this technology. If trainees reject the new way of learning, universities' return on investment will be diminished [6].

It retrieves the information through the use of mathematics and machine learning techniques and algorithms. Prediction is the most widely used technique in educational data mining research [7].

The authors of this paper/developers created a school management system using an Android application for managing instructional activities in schools. Teachers can monitor trainee attendance, upload assignments, and record grades online, while trainees can access their performance records, including grades and attendance [8]. Collecting, managing, processing, and analyzing such data promptly can be extremely challenging. By applying data mining techniques, valuable and meaningful patterns can be identified, unlocking the wealth of information and insights hidden within large datasets [9].

TVET Education is critical in the country's development and long-term economic prosperity. The Fig. 1 shows that the EDM process comprises five major phases, which are used to extract useful information from educational data, and it has been widely applied in research on academic achievement [10]

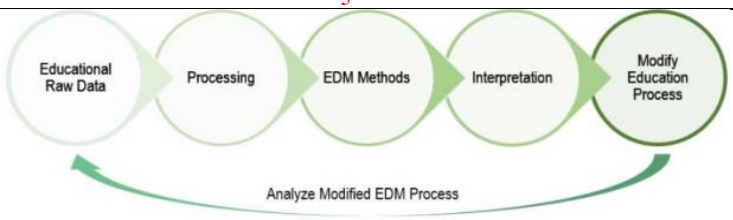


Figure 1. Outline of EDM Process (Adapted from [1]).

When applied to education, these methodologies are known as Learning Analytics (LA) and Educational Data Mining (EDM). The only difference is that EDM researchers are more interested in utilizing automated models to extract information from educational data, whereas Learning Analytics researchers are more interested in investigating educational data using human-led techniques [11].

It is frequently more crucial to consider the elements that influence the quality of instruction. Furthermore, no detailed association analysis is performed to explore the various factors influencing teaching quality ratings [12].

Education Data Mining can be carried out using a variety of techniques, including decision trees, neural networks, naive bayesand K-Nearest neighbor. Predicting trainee performance can help identify underperforming trainees and guide interventions to enhance their outcomes [13].

Objectives:

The main objectives of this research are:

The objective of the research is to develop a mobile and web-based application that will help in maintaining the close relationship between parents, teachers, and institute management.

It will maintain detailed profiles of trainees, faculty, and staff.

It will enhance attendance and improve grades by sharing daily attendance and results of the trainees with their parents.

Literature Survey:

In [14], the authors performed a study to identify at-risk trainees to provide them with appropriate mentoring. The trainees with deficiencies in all of the pre-enrolment attributes were more likely to drop out. The authors tested three hypotheses and discovered that the entrance point score and school rank had a substantial impact on first-semester course results when they were combined. The trainees with deficiencies in all of the pre-enrolment attributes were more likely to drop out.

The authors of the paper/developers have studied the Trainees' continuous academic assessment and developed a general, customizable computerized system for trainees' assessment. They analyzed, designed, and developed a web-based computerized decision support system that assisted the users in tracking the performance of trainees from time to time, viewing the statistics of the trainees' performance through graphs and numerical information for an accurate decision-making process. It is a general, customizable computerized system used by a wider number of schools, and a few of them either have limited or no internet connectivity. It is successful in easing, speeding up, and simplifying the decision-making process by the school management [3].

The Authors of this research developed an android mobile application named Bagpack, used by faculty and trainees, including features such as a material uploading portal, smart attendance management system, and group chat interface [15].

The authors in this paper studied 80 publications of different journals and indexed in the scopus database between 2010 and 2020, with a focus on the social sciences. In this study, the content analysis method is used, and the findings show that the majority of articles were published year 2019, with South Africa having the highest number of publications, the

Mediterranean Journal of Social Sciences publishing the most articles, and quantitative research methods being used more frequently than other research methods [16].

Motivation:

Educational data mining is commonly used in numerous educational institutes and is helping to compute their trainees' performance in advance. Nowadays, TVET educational institutes operate in a highly competitive environment, and retaining trainees has become a major challenge for administrators. It is critical for the management of institutes to do a thorough analysis of themselves as well as their trainees in order to build a roadmap and plan for the future and prevent failure or dropout.

Many natural and man-made events caused significant loss of life and property, underscoring the need to strengthen our ability to prevent, respond to, mitigate, and recover from these disasters, as well as to identify missing or deceased individuals. Recent developments in information technology, trainees' interest in Commercial Intelligence (CI)/Data Mining (DM) grew, resulting in a rapid increase in business and organizational databases. Information such as trends and patterns can be used to enhance decision-making and maximize outcomes.

An android-based school management system manages the educational efforts in an effective way that leads management to the efficiency and improvement benchmark, providing them with functionality to see teachers' and trainees' performance. Teachers can mark attendance online, upload assignments, and view results. Trainees can get their performance view; they can check their results and attendance [6].

Problem Statement:

Technical & Vocational Education and Training(TVET) is an important component in attaining long-term economic growth. The educational level of the Khyber Pakhtunkhwa province has improved in recent decades. However, owing to its high cultural barriers, resistance in women's education, and dropout rates, statistics place Khyber Pakhtunkhwa at the 2nd position in Pakistan. The dropout rates mostly owe to the unawareness of parents sitting at home or workplaces of their children going to institute. Results, daily attendance, and the status of monthly fee payments are not shared with parents. Additionally, the institute's leaving certificate and academic results are not available online. MIS can play an important part in the field of decision-making because it helps in monitoring any gaps available in the TVET education system. It can also assist in defining a course of action and determining how to manage the system by supporting analysis, evaluation, decision-making, and providing the necessary information for implementation. Most importantly, the institute does not maintain records of employees, including faculty and administrative staff, which prevented management from arranging makeup classes for trainees due to the absence of an online leave system.

Significance of the Research:

Total TVET institutes (both Public and Private) in Khyber Pakhtunkhwa are [17]:

Table 1. Total TVET Institutes in Khyber Pakhtunkhwa

Institutes Details	Total
Public	106
Private	580
Total	**686**

Table 2. Gender-Wise TVET Institutes in Khyber Pakhtunkhwa

Category	Public (Technical)	Private (Technical)	Vocational Total (all)	Total
Male	26	2	476	**504**
Female	4	2	135	**141**
Co-education	–	2	39	**41**
Institutes Count	30	6	650	**686**

Approximately 57,906 trainees (both male and Female) are enrolled in these institutes [17]:

Table 3. Gender-Wise Enrolment in TVET Institutes in Khyber Pakhtunkhwa

Province	Technical Enrolment	Vocational Enrolment	Total TVET Enrolment
Male	17,197	23,379	**40,576**
Female	873	16,457	**17,330**
Total	**18,070**	**39,836**	**57,906**

The main objectives of this research are to develop a Data Mining Application System that will be used to smartly administer the TVET Institutes.

Proposed Solution:

A Data mining Application to identify the trainee, as it will use the CNIC/Form-B number as a unique ID of the trainee and will remain the same throughout the educational career.

Help to implement a single national curriculum as well as Boards and TVET institutes to implement the policy that 01 Certificate/Diploma during a single session.

Help NCHD in the Census of TVET Education. It will also help to identify the HDI of different areas.

Daily attendance (Biometric) of the Trainee and Monthly Fee slips, date sheet, and Results will be shared with the parents/guardians through SMS/WhatsApp.

Trainees will be able to view their grades, attendance, apply for Leave, as well as the Institute Leaving Certificate online.

Details of the faculty and staff will also be available in the system.

They will be able to apply for the leave, Experience, Job Leaving, and No Objection Certificates.

Research Methodology:

During the course of this research, the following roadmap was followed, which is explained with the help of the flowcharts:

Flowchart A (For Trainees):

In the First step, the Candidate applies for admission in the institute, and if the age is less than 15, then the application will be rejected.

The candidate submits the form with all the relevant documents.

If Form-B/CNIC (unique number of trainee) is available candidate's interview is conducted.

After a successful interview, admission is granted.

After admission, automatically, class and section are assigned to the trainee as the total number of trainees per section is already defined in the system by the Administration.

Trainees' fingerprints are registered in the biometric device, which is connected to the online system.

The system automatically forwards the daily attendance to parents/guardians via SMS or WhatsApp.

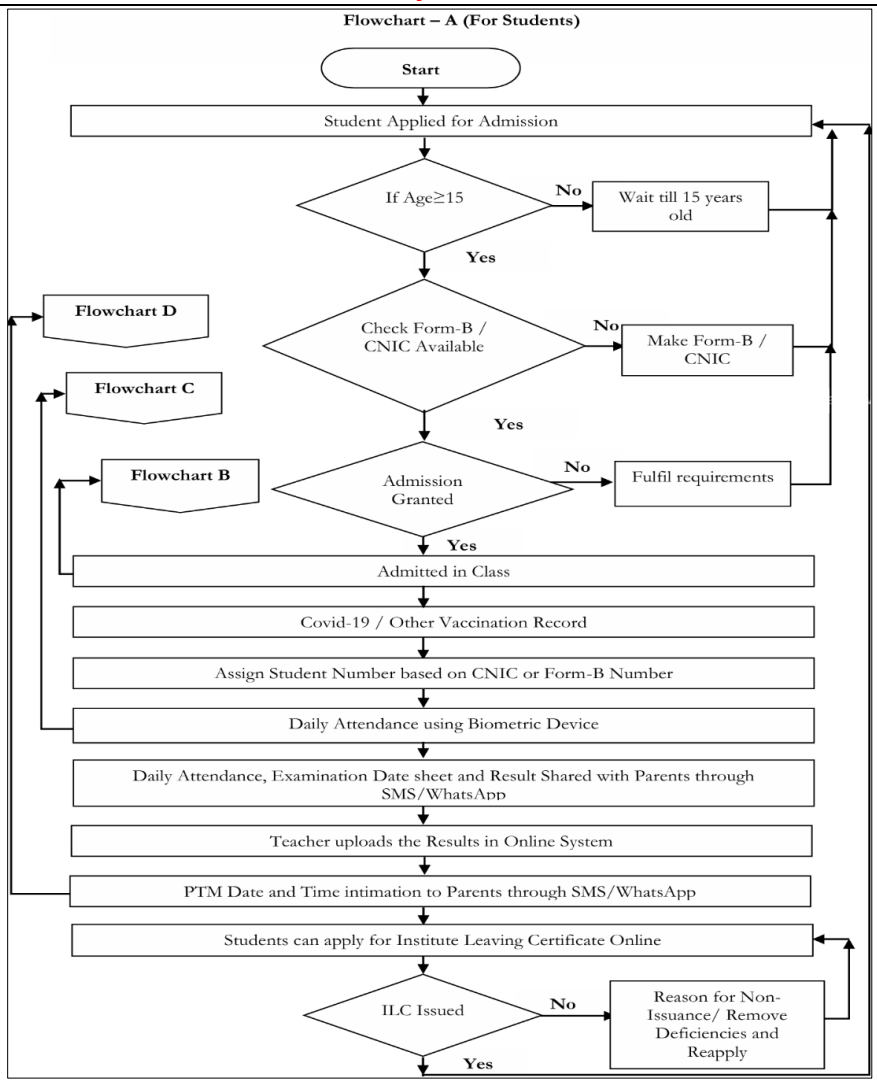
Examination dates and the results of the trainee are communicated to parents/guardians through SMS or WhatsApp, and are also accessible online for the trainee.

If the trainee passes the examination, he/she automatically promoted to the next class.

If a trainee wants to leave the institute, he/she apply for the Institute Leaving Certificate (ILC) through the online system.

The trainee can now get admission in any other institute.

The administration of the new institute is able to access the trainee's previous academic and vaccination records through the online system.



Flowchart B (For Management):

In the first step, the candidate applies for a management job at the institute.

If he/she does not qualify, the application is rejected; otherwise, a job is offered, provided that he/she possesses a valid CNIC (unique employment number).

Employees' fingerprints are registered in the biometric device, which is connected to the online system to mark daily attendance.

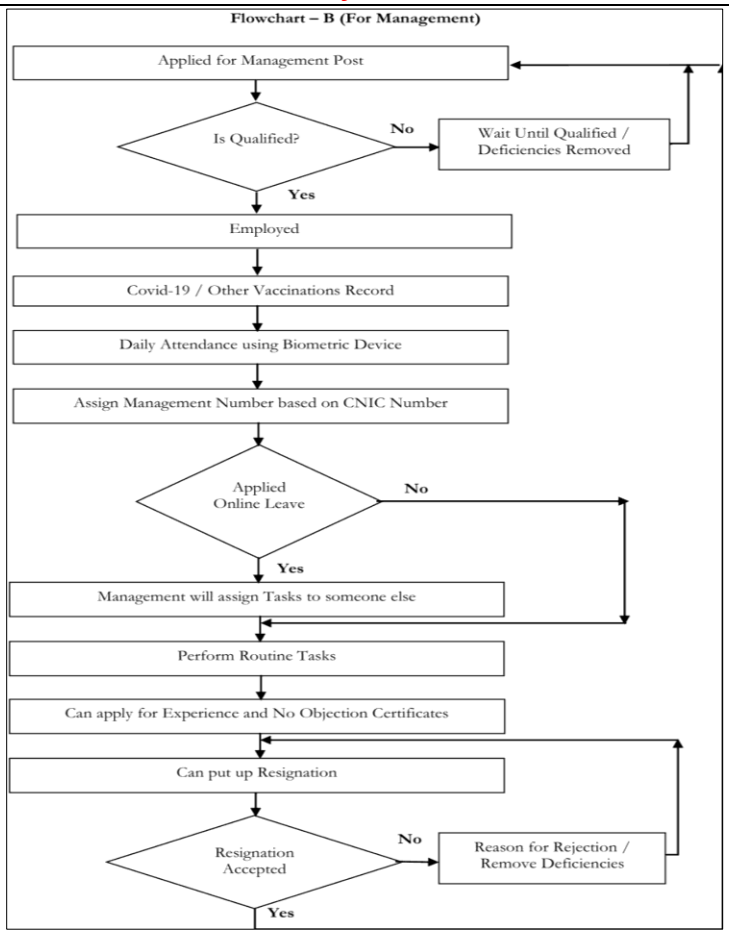
Employees apply for leave using this online system, which allows management to assign his/her tasks to other employees.

Employee is able to apply for the experience certificate and no objection certificate through this online system, and his/her application is processed and the same is issued to him/her.

If an employee wishes to leave the job, he/she is able to put up an online resignation, which is processed by the management, and if there is no pending liability on the part of the employee, the resignation is accepted; otherwise, the reason for its rejection is communicated to the employee.

Tools of Research:

The Data Mining for Smarter Administration of TVET institutes' systems will be a mobile and web-based application. It will keep a close relationship between parents, Teachers, and the Institute Management. It will save the time of instructors, administrators, and trainees.



Conclusion:

The data mining for smarter administration of TVET institutes system will streamline and simplify the management of (TVET) institutes. The system will help decision-making become more effective and increase transparency. It will also act as a link between the stakeholders, i.e., management of the institute and the parents/guardians of the trainees. The system will share the information regarding attendance, performance, and general advancement with parents/guardians. In addition, the system will also have a complete database of the trainees, and therefore, institutes will easily keep a record of students who are either newly enrolled, transferring within the institute, or transferring to a different institute. This intelligent administrative system will eventually enhance accountability, help in the census of the TVET training, increase the efficiency of institutions, and more interconnected and data-driven TVET ecosystem.

References:

[1] N. D. N. and W. Z. W. H. W F Wan Yaacob, N Mohd Sobri, S A Md Nasir, W F Wan Yaacob, "Predicting Student Drop-Out in Higher Institution Using Data Mining Techniques," *J. Phys. Conf. Ser.*, vol. 1496, 2020, doi: 10.1088/1742-6596/1496/1/012005.

[2] S. L. Yi Peng, Yong Zhang, Yu Tang, "An incident information management framework based on data integration, data mining, and multi-criteria decision making," *Decis. Support Syst.*, vol. 51, no. 2, pp. 316–327, 2011, doi: https://doi.org/10.1016/j.dss.2010.11.025.

[3] A. L. Moses, Byaruhanga, "A Flexible Trainee Assessment System for Continuous Performance Monitoring," *Int. J. Technol. Manag.*, vol. 5, no. 2, pp. 1–22, 2020.

[4] A. S. MohammadNoor Injadat, Abdallah Moubayed, Ali Bou Nassif, "Systematic ensemble model selection approach for educational data mining," *Knowledge-Based Syst.*, vol. 200, p. 105992, 2020, doi: https://doi.org/10.1016/j.knosys.2020.105992.

[5] F. Yang and F. W. B. Li, "Study on student performance estimation, student progress analysis, and student potential prediction based on data mining," *Comput. Educ.*, vol. 123, pp. 97–108, Aug. 2018, doi: 10.1016/J.COMPEDU.2018.04.006.

[6] W. Q. Syed A. Raza, "Social Isolation and Acceptance of the Learning Management System (LMS) in the time of COVID-19 Pandemic: An Expansion of the UTAUT Model," *J. Educ. Comput. Res.*, vol. 59, no. 2, 2021, doi: https://doi.org/10.1177/0735633120960421.

[7] M. Y. O. Delima, Allemar Jhone P.Arroyo, Jan Carlo T.Elape, Michelle C., "The Use of Schoology as Learning Management System in the College of Computing Education: A Response Assessment using Data Mining Techniques," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 3, 2020, doi: https://www.warse.org/IJATCSE/static/pdf/file/ijatcse169932020.pdf.

[8] Muhammad Anwar Shahid, Akmal Rehan, Salman Afsar Awan, Ahmed Mateen, Raim Odenaev, "Improving Education With Enhanced Machine Learning Technique for Android Applications," *Int. J. Adv. Res. Eng. Technol.*, vol. 11, no. 9, pp. 582–590, 2020, [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3712873

[9] B. Wibawa, J. S. Siregar, D. A. Asrorie, and H. Syakdiyah, "Learning analytic and educational data mining for learning science and technology," *AIP Conf. Proc.*, vol. 2331, no. 1, Apr. 2021, doi: 10.1063/5.0041844.

[10] T. O. Phauk Sokkhey, "Developing Web-based Support Systems for Predicting Poor-performing Students using Educational Data Mining Techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, 2020, doi: 10.14569/IJACSA.2020.0110704.

[11] Ginika Mahajan, Bhavna Saini, "Educational Data Mining: A state-of-the-art survey on tools and techniques used in EDM," *Int. J. Comput. Appl. Inf. Technol.*, vol. 12, no. 1, 2020, [Online]. Available: https://www.researchgate.net/publication/340983783_Educational_Data_Mining_A_state-of-the-art_survey_on_tools_and_techniques_used_in_EDM

[12] H. Li, "Application of Classification Mining Technology Based on Decision Tree in Student Resource Management," *Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng. LNICST*, vol. 465 LNICST, pp. 149–160, 2023, doi: 10.1007/978-3-031-23950-2_17/COVER.

[13] A. U. K. and Harwati, "A Comparative Study to Predict Student's Performance Using Educational Data Mining Techniques," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 215, 2017, doi: 10.1088/1757-899X/215/1/012036.

[14] Á. S. Erika B. Varga, "Detecting at-risk students on Computer Science bachelor programs based on pre-enrollment characteristics," *Hungarian Educ. Res. J.*, 2021, doi: https://doi.org/10.1556/063.2021.00017.

[15] G Dhanush, Dr KG Mohan, "Integrated Adaptable Mobile Based Platform For University/College Management Using Firebase," *Int. J. Eng. Appl. Sci. Technol.*, vol. 5, no. 1, pp. 373–378, 2020, [Online]. Available: https://www.ijeast.com/papers/373-378,Tesma501,IJEAST.pdf

[16] R. N. Tangiri, I. Adamu, and F. Köprülü, "Content Analysis On School Management And Academic Performance," *Near East Univ. Online J. Educ.*, vol. 3, no. 2, pp. 42–52, Sep. 2020, doi: 10.32955/NEUJE.V3I2.246.

[17] Farah Nadeem, Muhammad Muzaffar Iqbal, "Annual School Census Report 2020-2021," *ResearchGate*, 2021, doi: 10.13140/RG.2.2.31015.27043.