ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206

Volume 9, Issue 4, April 2026

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

An EMIS-Driven Machine Learning Framework for Early Prediction and Intervention of School Dropout in Government Schools of Nilgiris District, Tamil Nadu

Dr. A. Vidhya[1], Shalini G[2], Subha S[3]

Assistant Professor, B. S. Abdur Rahman Crescent Institute of Science and Technology Vandalur, Chennai, Tamil Nadu, India[1]

B.Sc. Computer Science – III Year, Department of Computer Applications, B. S. Abdur Rahman Crescent Institute of Science and Technology Vandalur, Chennai, Tamil Nadu, India[2]

B.Sc. Computer Science – III Year, Department of Computer Applications, B. S. Abdur Rahman Crescent Institute of Science and Technology Vandalur, Chennai, Tamil Nadu, India[3]

ABSTRACT: Student dropout continues to pose a serious chal- lenge in rural and economically disadvantaged regions, where limited resources and socio-economic constraints affect educa- tional continuity. Identifying students at risk at an early stage is most crucial for enabling effective and timely intervention. This paper introduces a data-driven framework that is based on Education Management Information System (EMIS) data to predict potential dropout cases among students in government schools in the Nilgiris district of Tamil Nadu.

The proposed approach utilizes key student attributes, includ- ing attendance patterns, academic performances, demographic characteristics, and socio-economic indicators. A set of supervised machine learning models—namely, Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine—is applied to learn core patterns associated with dropout behavior. The framework further integrates contextual risk factors such as early marriage, child labour, migration, geographic remoteness, and challenges in post-secondary transition.

To support practical implementation, an interactive web-based dashboard has been developed for visualizing model outputs and supporting continuous monitoring by educators and administra- tors. In addition, an automated alert mechanism is incorporated to highlight high-risk students for timely intervention by relevant authorities. Model explainability is addressed through the use of SHAP (SHapley Additive exPlanations), enabling transparent understanding of feature contributions for predictions. The pro- posed system aims to support informed decision-making, promote early intervention strategies, and ultimately improve student retention outcomes in government schools.

KEYWORDS: Education Management Information System (EMIS), Machine Learning, School Dropout Prediction, Educa- tional Data Mining, Student Retention

I. INTRODUCTION

Education plays an important role in shaping individuals as well as communities, yet dropout continues to be a concern, especially in rural areas. For this reason, the Nilgiris district was chosen for this paper, as it presents a unique combination of social, economic, and geographic challenges.The region is largely rural and has a large tribal population. Many families depend on tea and coffee plantations for their livelihoods, and the nature of this work changes with the seasons. Because of this, students are not always able to attend school regularly. So, it’s gradually affecting the students’ academic progress.Reaching schools themselves can be difficult. In several cases, students have to travel long distances across hilly and uneven terrain, and also to the outskirts. Along with this, financial drawbacks at the household level often influence whether a child continues education or dropout school early.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Taken together, these conditions make the Nilgiris an im- portant setting for understanding why students drop out and what kind of support might help them stay in school instead of dropping out.

The good news is that government schools in the Nilgiris maintain detailed EMIS records. This provides a strong foundation for identifying and predicting students who are at risk of dropping out of school. By focusing on the Nilgiris region, the paper can address the unique challenges faced in hill districts and develop solutions that are specifically tailored to the local context, making retentions more effective and practical.

There are more reasons students drop out - based on money problems, having to help at home, or just not keeping up in school. When schools give attention to these students early, they can step in and help before things spin out of control, which leads to dropping out.

EMIS is a standard tool in government schools for tracking everything from attendance to grades and background information of the particular student. Machine learning gives a deep look into this data, finds patterns, and flags students who are in danger of dropping out in the early stage of students [1], [5].

II. LITERATURE REVIEW

Educational Data Mining (EDM) focuses on extracting meaningful patterns and insights from large-scale educational datasets. Here, it refers to EMIS data. In this domain, re- searchers have worked on a wide range of machine learning techniques, including Decision Trees, Random Forests, and Neural Networks, to address key challenges such as predict- ing student performance and identifying students at risk of dropping out early from the schools [2], [3].

Earlier studies indicate that multiple factors significantly influence student retention, including attendance patterns, academic performance, family income, and parental education levels. Together, these factors play a very crucial role in determining whether students continue their education or are at risk of early dropout. [4], [6].

However, a significant disadvantage of existing research is its limited focus on government schools in rural regions such as the Nilgiris, highlighting a notable research gap. In recent years, there has been a fast-growing interest in the application of AI (Artificial Intelligence) and machine learning in the field of education. Educational institutions are increasingly adopting data-driven approaches to improve student retention and academic performance. These advanced models are most capable of processing large-scale datasets and uncovering complex patterns that may not be easily or quickly detected through traditional statistical methods.

Researchers have developed predictive models to identify students at risk of dropping out by analyzing their academic performances, behavioral patterns, and demographic informa- tion of the students. Various machine learning techniques, including Na¨ıve Bayes, Support Vector Machines (SVM), and Neural Networks, have been widely used for this purpose.

There is also enlarging the emphasis on learning analytics, which involves systematically tracking and analyzing students' learning activities. These approaches enable educators or ad- ministrators to monitor engagement levels, identify potential issues at an early stage, and provide more personalized support to students.

Even with these advancements, many government schools continue to fall back on conventional data analysis methods. Leveraging machine learning techniques with EMIS datasets has the potential to transform raw educational data into ac- tionable insights, supporting timely interventions and more informed policy decisions.

Furthermore, government data from India highlights per- sistent challenges in student retention, particularly in rural and tribal regions. For example, in Eklavya Model Residential Schools (EMRS), the number of student dropouts increased significantly from 111 in 2021–22 to 552 in 2024–25, despite the presence of targeted support programs. Similarly, dropout rates remain notably high in tea garden regions such as Assam, where financial constraints often compel students to discontinue their education. Reports from the Ministry of Education further corroborate these trends, indicating that states characterized by challenging geographical conditions and socio-economic barriers tend to experience higher dropout rates, especially in rural and hilly districts.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. DATASET DESCRIPTION

We built our study around data pulled from the Education Management Information System (EMIS) records of government schools in the Nilgiris district, Tamil Nadu. EMIS acts as a central hub for all things student-related—demographics, academics, and even details about the schools themselves.

The dataset covers a good mix of factors that affect how students do in school and whether they stick with it. Here's what we looked at:

- Student ID
- Age
- Gender
- Grade Level
- Attendance Percentage
- Academic Performance
- Parent Occupation
- Family Income Level
- Distance from School
- Previous Academic Record
- Dropout Status

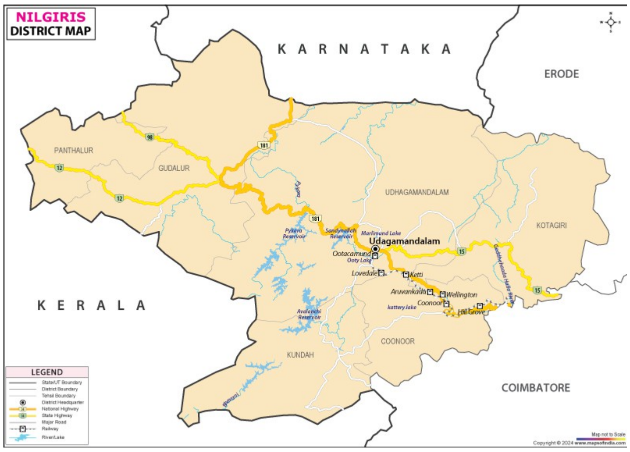

Fig. 1. Location map of Nilgiris district in Tamil Nadu, India.

Figure 1 Take a look at the map; it shows exactly where Nilgiris district sits in Tamil Nadu, which is where all of our data comes from.Before beginning the analysis, careful data preprocessing was carried out to ensure the dataset was reliable and well- structured. This included cleaning the data, handling missing values, normalizing features for consistency, and encoding variables for use in machine learning models. Irrelevant at- tributes were removed, and the most significant features were selected using statistical methods and correlation analysis.

Once the data was prepared, it was divided into training and testing sets. These processed features helped in gaining mean- ingful insights into the factors influencing student dropout and provided a strong foundation for building effective prediction models.

A. Challenges in Data Collection

Collecting data from EMIS wasn't as straightforward as you'd hope. Some files were missing key information, such as attendance rates or what the students' parents did for a living. We also ran into typos, inconsistent entries, and some underreporting—especially from the more remote or tribal schools.To fix this, we used handling missing data for gaps where it made sense and cleaned up unstructured records by double- checking them with school registers. In areas that were harder to reach, we went the extra mile—literally—with field visits and follow-ups to make sure everything was accurate and complete.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904369

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. FEATURE SELECTION

Choosing the right features can make a huge difference when it comes to training machine learning models, especially with educational data. Almost all of the information in these datasets does not really help predict which students are likely to drop out. Focusing on the features that actually matter cuts down on sound, keeps the model simple, and ensures instant accuracy.

In this paper, we examine the factors that have the greatest consequence on the prediction of student dropout. Some stand out right away—attendance percentage, academic per- formance, family income, and even how far a student lives from school all play a big role.

Attendance, for example, is a strong sign of whether a student is engaged. When students start missing a lot of classes, they are more likely to struggle and eventually drop out.

Socio-economic background matters too. Family income and parents' jobs can affect whether a student stays in school. Students from families with fewer resources are more likely to face financial problems that push them out earlier from the schools.

By narrowing things down to the most important features, the models can zero in on the patterns that really matter and deliver better, more trustworthy predictions.

V. DATA PREPROCESSING

The EMIS dataset isn't perfect—it's got missing informa- tion about the students, and some entries just don't add up. To get it ready, we struggled with a few key steps.
Data cleaning came first. We got stucked of missing values and any redundant records.

VI. PROPOSED METHODOLOGY

The system we designed moves through a few main stages: collecting the data, cleaning and pre-processing it, training models, making predictions, and finally showing the details on a dashboard.

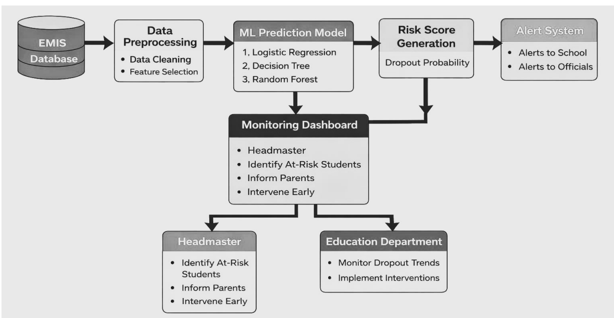

Fig. 2. Architecture of EMIS-Based School Dropout Prediction System

It all starts with pulling student information from govern- ment school EMIS databases. This gives us details on things like demographics, attendance, grades, and family background information.

Once we have the data, we clean it up and convert ev- erything into a format of a structured version that machine learning models can use—categorical stuff becomes numbers, and we make sure that features are scaled consistently with prediction results.

For training, we run several machine learning algorithms to spot the cautionary signs that a student might drop out. We use past data of the student to teach the models, then check how well they perform with different confusion metrics.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904369

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

When it's time to share the results, we use a web dashboard. It gives teachers and school staff a clear, best visual way to see which students are at risk. With these insights, they can keep an eye on students' progress and step in early to help students who need it most to sustain their education.

VII. MACHINE LEARNING MODELS

This research uses several machine learning algorithms to predict student dropout for early intervention. [5], [7]

A. Logistic Regression

Logistic Regression estimates the chance that a student will drop out. The formula looks like this:1 −(β +β x +β x +···+β x) (1)

Next, we handled categorical data like gender and parent occupation by turning them into numbers that models can work with.

B. Decision Tree

1+e0 1 1 2 2 n n

We also normalized continuous features like attendance, so everything's on the same scale. Finally, the cleaned dataset was split into training and testing sets, allocating 80 percent for model training and 20 percent for evaluation.
Decision Trees make classifications by asking a series of "yes" or "no" questions about each student. It's a straight-forward way to sort out the students based on key feature selection.

C. Random Forest

Random Forest goes a step further. It builds a bunch of different decision trees and gives the combinations of their results. This regularly gives more accurate and precise predic- tions and helps avoid overfitting.

D. Support Vector Machine

SVM tries to draw the cleanest possible line between students likely to drop out and those who aren't, based on their data.

E. Model Training and Hyperparameter Tuning

The hyperparameters for each machine learning model were carefully tuned to enhance predictive performance. Specifically:

- **Random Forest:** Configured with 100 trees and a max- imum depth of 10.
- **SVM:** Employed an RBF kernel with a regularization parameter C = 1.0.
- **Decision Tree:** Maximum depth = 8, minimum samples required per leaf = 5.
- **Logistic Regression:** Regularization parameter C=1.0 and solver set to lbfgs.

To evaluate the effectiveness of these models, we applied 5- fold cross-validation. In this process, the dataset is divided into five equal parts. Each part is used once as a test set, while the remaining four parts serve as the training set. The results from all five iterations are then averaged, providing a reliable measure of mode performance and helping to reduce the risk of overfitting.

VIII. FEATURE IMPORTANCE ANALYSIS

Explainable AI matters a lot here. It's not enough just to predict which students are at risk—teachers, school leaders, and policymakers need to know why. If they don't understand what's driving the predictions, they can't make smart decisions or take action.

So, we use feature importance analysis to spotlight which factors matter most. Looking at the Random Forest model, the top indicators are attendance percentage, academic perfor- mance, family income, and how far students live from school. The SHAP(SHapley Additive exPlanations) feature impor- tance chart (see Figure 3) makes this clear. Attendance and academic performance stand out as the biggest predictors. SHAP is especially handy—it breaks down exactly how much each feature contributes to a prediction. When you run SHAP on our dropout model, you see that low attendance, poorgrades, and low family income are the heavy hitters.

Figure 3 illustrates the SHAP feature importance values for the trained model. The analysis indicates that attendance percentage and academic performance are the most influential predictors of student dropout risk.

This kind of insight goes beyond the usual "black box" feeling you get with machine learning. Now, the predictions actually make sense to people who work in education.

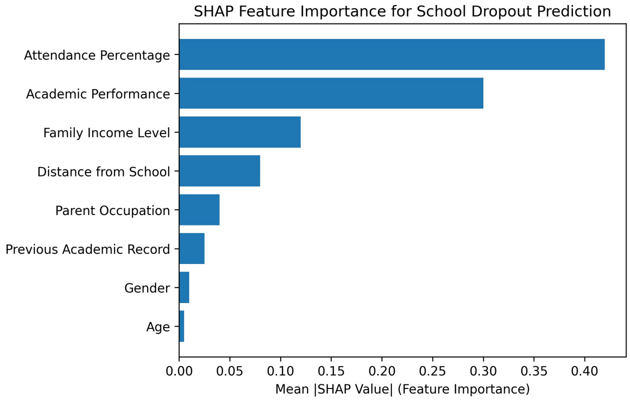

Fig. 3. SHAP Feature Importance for School Dropout Prediction

When you know which factors matter most, you can do something about them. The results show that attendance is the key—students who show up less than 70 percent of the time are much more likely to drop out. Academic performance comes next; students who keep failing or getting low marks are also at higher risk.
Family income and parental jobs matter too. Kids from low- income families often struggle to stay in school for financial reasons. And in rural places like Nilgiris, just getting to school can be a big obstacle. If the distance is too far, attendance drops.

All of these point to the need for schools to track both academic and social factors if they want to find problems early. Using explainable AI tools keeps the system open and trustworthy, so decisions aren't just based on ambiguous numerical values.

IX.SYSTEM IMPLEMENTATION AND EMIS FRAMEWORK

Figure 4 shows the whole setup. The system starts by col- lecting data through the Education Management Information System (EMIS). That data then flows through the prediction models, and the results feed into early intervention efforts.
This district-level system is designed to integrate smoothly with existing education management frameworks, helping schools and districts identify students who may struggle to continue their education and take timely action to support them before they are at risk of dropping out.

The system runs as a district-level Education Management Information System (EMIS) for government schools in the Nil- giris. It pulls everything together—student records, analytics, and even machine learning—to help the education department keep track of what's happening across all its schools.

A. Role-Based Login System

Everyone who uses the system signs in according to their role. District Education Officers get access to everything, in- cluding big-picture district analytics. Block Education Officers can check on schools in their area, which means under the taluks. Headmasters see their own school's dashboard and student details. Teachers can view student attendance and grades and enter the reasons for the absences, while Data Entry Operators update and manage the records. This setup keeps sensitive student data safe, more confidential user data, and only lets people see what they actually need.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

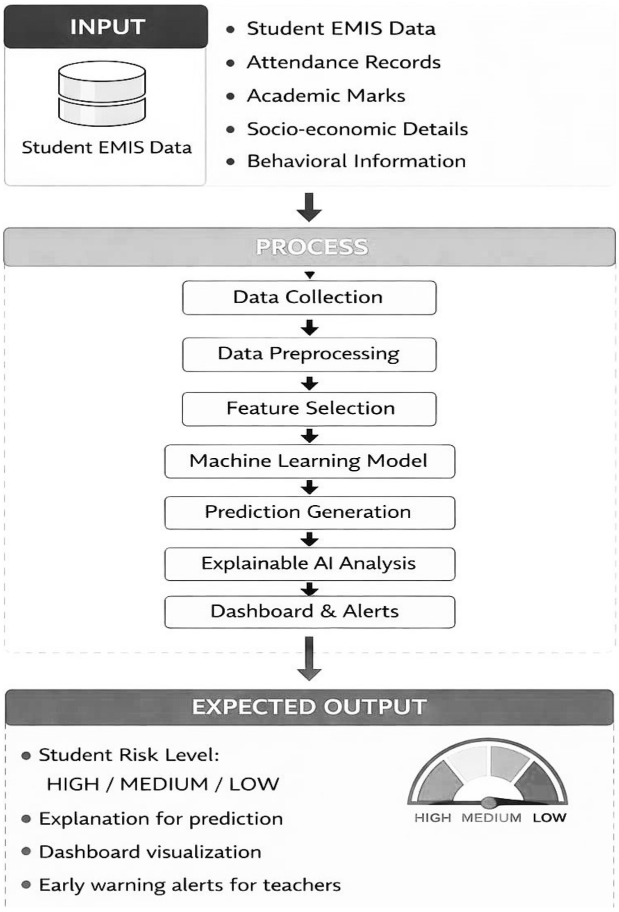

Fig. 4. Overall Workflow of the EMIS-Based School Dropout Prediction System

The primary user roles include:

- District Education Officer – Full system access and district-level analytics
- Block Education Officer – Monitoring of schools within a specific block
- Headmaster – Access to school-level student records and dashboards
- Teacher – Viewing student academic and attendance information, enter the reasons for the absences of the students.
- Data Entry Operator – Updating and maintaining student records

This role-based architecture ensures data security and con- trolled access to sensitive student information.
School Master ModuleThis module is where every school's basic details live—name, block, type, number of students and teachers, and even infrastructure info. The focus stays on government schools across Ooty, Coonoor, Gudalur, Pandalur, and Kota- giri.

B. Student Master Module
Here's where all the student info is stored. The system collects everything: demographics, academic results, attendance, and family background. Each record includes student ID, gender, age, community category, grades, attendance rates, parents' jobs, and income level. All of this feeds into the machine learning models, which use the data to spot patterns and risks.

C. Post-10th Transition Tracking
One of the most important features: the system keeps an eye on students after 10th grade. It automatically checks who finished 10th standard and whether they've signed up for higher secondary in the district. If someone hasn't enrolled, the system flags them as a potential dropout and alerts the right education officials to follow up.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

D. District Dashboard
A central dashboard pulls together key numbers for the whole district—total schools, student enrollment, percentage of high-risk students, and a comparison of dropout risks across blocks. You also get charts showing why students drop out, gender breakdowns, and risk analysis for tribal students.

E. School-Level Dashboard
Each school gets its own dashboard. Headmasters and teachers can track attendance, academic results, and see which students are marked as high risk. High-risk students stand out on the dashboard, so teachers know who needs extra support or counseling right away.

F. Early Warning System
The system's early warning feature spots students who might drop out—think low attendance, dropping grades, or family situations like seasonal migration for tea plantation work in areas like Ooty, Coonoor, and Gudalur. When the system flags a student, teachers, headmasters, and block offi- cers get an alert. They have to acknowledge it in the system, so there's a record and the response can be tracked.

From there, schools can step in with the right help—counseling, financial aid, mentoring, working with parents, or adjusting for attendance issues tied to seasonal labor. Every action and acknowledgment is logged, so administrators can see what's working and make sure students don't slip through the cracks.

In the end, this whole setup does more than just identify students at risk. It makes sure they get real support, aiming to keep more kids in school and cut down on dropouts across the Nilgiris.

G. Early Warning Dashboard
The dashboard gives district officials a clear look at dropout risks, broken down by block and taluk. This way, they can focus resources and support where students actually need them. The tool uses a simple "Traffic Light" system to show risk levels:

- **Red (High Risk):** Students with less than 60
- **Yellow (Moderate Risk):** Students whose attendance has dropped for three months in a row.
- **Green (Low Risk):** Students who keep up steady atten- dance and grades.

X. RESULTS AND DISCUSSION

The results show that ensemble learning, especially Random Forest, predicts dropout risk better than models such as logistic regression or single decision trees. That's because Random Forest pools the strengths of several decision trees, which helps to ignore overfitting and tends to make more reliable predictions.Attendance and academic performance really stand out as key factors in predicting which students are at risk. When students start skipping more classes, the model tags them as high-risk.By bringing machine learning together with EMIS data, schools move from reacting to problems to spotting trouble early. They can step in before positive students fall too far behind, giving them a better shot at staying in school.Check out Table I for how each model performed.

Table I: Performance Comparison of Machine Learning Models
TABLE I

PERFORMANCE COMPARISON OF MACHINE LEARNING MODELS

Model	Accuracy	Precision	Recall	F1 Score
Logistic Regression	82%	0.80	0.79	0.79
Decision Tree	85%	0.83	0.82	0.82
Random Forest	91%	0.90	0.89	0.89
SVM	88%	0.86	0.85	0.85

Random Forest comes out on top—it's just better at picking up on the complicated patterns in the data.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

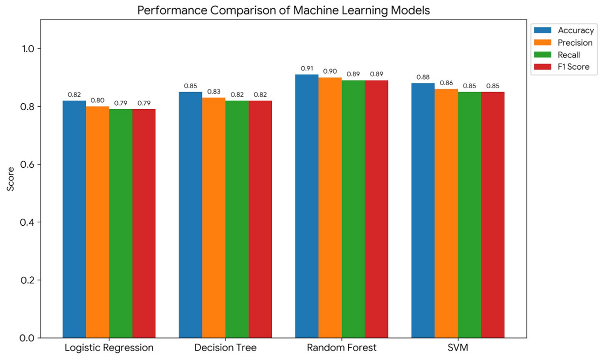

Fig. 5. Comparative Performance Analysis of Machine Learning Classifiers for Student Dropout Prediction.

XI. EVALUATION METRICS

To judge how well the models work, we look at:

- Accuracy
- Precision
- Recall
- F1 Score

A. Performance Metrics

The machine learning models were estimated utilized stan- dard criteria:

Model performance can be evaluated using several key metrics. **Accuracy** measures the overall proportion of correct predictions among all cases, giving a general sense of cor- rectness. **Precision** indicates how reliable the model's positive predictions are, showing the fraction of predicted positives that are actually correct. **Recall** (or sensitivity) reflects the model's ability to identify all true positive cases, representing the proportion of actual positives detected. The **F1 Score** combines precision and recall into a single value by calculating their harmonic mean, offering a balanced assessment of the model's performance. Together, these criteria give a good sense of both how frequently the models are right and how important you can trust their prognostications

XII. LIMITATIONS

Indeed, though the system works well, it's not perfect. The study only uses data from government seminaries in the Nilgiris quarter, so the results might not match up. Plus, the dataset focuses on structured EMIS data. It doesn't include effects like pupil gesture or cerebral factors that could play a big part in dropouts. latterly studies can get better results by adding new data sources — like how engaged scholars are in class, whether they join clubs or sports, and perceptivity from learning gesture analytics. That should help paint a fuller picture of a powerhouse threat.

SOCIO-ECONOMIC RISK ANALYSIS IN NILGIRIS DISTRICT

The Nilgiris district in Tamil Nadu is largely rural and is home to a significant tribal population. In this region, everyday living conditions and economic pressures play an important role in determining whether students continue their education or drop out. For many families facing financial difficulties, the need to earn a livelihood can lead students to leave school at an early stage.

One of the major concerns in these areas is early mar- riage, particularly among girls. In several villages, stu- dents—especially those in their mid-teens—discontinue their education due to family expectations and prevailing social norms. These situations are often accompanied by warn- ing signs such as irregular attendance, declining academic performance, and low household income. Machine learning approaches can help identify such patterns early and flag students who may be at risk.

Seasonal migration further contributes to the problem. Many families are employed in tea and coffee plantations, and duringpeak seasons, they relocate temporarily for work. This disrupts students' regular schooling and increases the

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

likelihood of dropout. Regions such as Gudalur and Pandalur are particu- larly affected, where tribal students frequently encounter these challenges.

Accessibility is another significant issue. The hilly terrain of the Nilgiris, along with limited transportation facilities, makes it difficult for students to travel to schools located in towns such as Ooty and Coonoor. Long travel distances and challenging routes often discourage regular attendance and, in some cases, lead to students discontinuing their education altogether.

Recent statistics from Tamil Nadu indicate that dropout rates remain a concern, particularly in such geographically and economically constrained areas. By analyzing these socio- economic factors using EMIS datasets, it becomes possible to identify at-risk students at an early stage. This approach enables education authorities to take timely action and pro- vide targeted support, ultimately helping to improve student retention.

INTERVENTION STRATEGIES

Once we know which scholars are at threat, seminaries can step in with many solid strategies, academic counseling, reg- ular parent-schoolteacher meetings, fiscal aid, and mentoring programs. Getting involved beforehand helps keep the students in the schools and cuts down on dropouts.

XIII. CONCLUSION

This exploration puts forward an EMIS- driven machine literacy system to flag scholars at risk of dropping out in Nilgiris. government seminaries. By pulling together educa- tional data and prophetic analytics, the system spots advising signs before it's too late. The frame uses everything from pupil backgrounds to academic performance and home life to train models that predict scholars by threat position. There's also a dashboard and an early warning system so preceptors and directors can track progress and act when demanded. It shows how smart, data- driven opinions can strengthen academy operations, keep further students in class, and help pastoral and ethnic communities make a better future.

XIV. FUTURE WORK

Looking ahead, there's room to make this system even stronger. We can add real- time data from academy attendance systems and mobile apps, and try out advanced deep literacy styles to edge the prognostications. There's an eventuality to roll this out across all government seminaries in Tamil Nadu. Tying in support like literacy, better transport, and social welfare programs would make interventions more effective. On top of that, the system could send mobile cautions to preceptors, parents, and officers when a pupil's threat goes over — so they can step in right away.

REFERENCES

[1] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann, 2012.
[2] R. Baker and K. Yacef, ”The state of educational data mining in 2009: A review and future visions,” Journal of Educational Data Mining, vol. 1, no. 1, pp. 3–17, 2009.
[3] C. Romero and S. Ventura, ”Educational data mining: A review of the state of the art,” IEEE Trans. Syst., Man, Cybern., vol. 40, no. 6, pp. 601–618, 2010.
[4] S. B. Kotsiantis, ”Use of machine learning techniques for educational data mining,” Int. J. Artif. Intell. Appl., vol. 3, no. 2, pp. 1–10, 2012.
[5] T. M. Mitchell, Machine Learning. New York, NY, USA: McGraw-Hill, 1997.
[6] I. H. Witten, E. Frank, and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, 2011.
[7] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning. New York, NY, USA: Springer, 2009.
[8] UNESCO, Global Education Monitoring Report. Paris, France: UN- ESCO Publishing, 2020.
[9] K. Pal, ”Predicting student dropout using machine learning,” IEEE Access, vol. 8, pp. 100345–100356, 2020.
[10] S. Romero, ”Educational data mining applications in student perfor- mance analysis,” Int. J. Educ. Technol., vol. 14, no. 1, pp. 45–56, 2020.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[11] S. Kumar and R. Patel, ”Educational data mining for student dropout prediction,” Int. J. Educ. Technol., vol. 15, no. 2, pp. 120–128, 2021.
[12] P. Sharma and M. Gupta, ”Machine learning approaches for educational analytics,” IEEE Access, vol. 9, pp. 45210–45218, 2021.
[13] Economic Times, ”Dropouts from EMRS schools jumped nearly five- fold in 4 years: Chhattisgarh, Odisha worst hit,” Economic Times Education News, Mar. 2025. [Online]
[14] A. Kumar and S. Singh, “Machine learning approaches for student dropout prediction in educational systemsIEEE Access, vol. 10, pp. 45678–45689, 2022.

INNO SPACE
SJIF Scientific Journal Impact Factor

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com