

“A Dual Model for Predicting Student Dropout Risks and a Future-Oriented Career Guidance System using Machine Learning.”

Aparna Mote¹, Rugved Ande², Arhaan Shaikh³

¹Aparna Mote Computer & Zeal College Of Engineering and Research, Pune

²Rugved Ande Computer & Zeal College Of Engineering and Research, Pune

³Arhaan Shaikh Computer & Zeal College Of Engineering and Research, Pune

Abstract

Student dropout has become a significant challenge in the education sector, particularly at the secondary school level, where academic, social, and economic factors often influence a student's decision to discontinue education. At the same time, many students lack adequate guidance regarding their interests, strengths, and future career opportunities. To address these challenges, this paper presents the **Student Success and Career Navigation Management System (SSC Management System)**, an integrated web-based platform designed to support both student retention and career development. The system combines machine learning-based dropout risk prediction, psychometric interest assessment, career recommendation generation, and intervention management within a unified application. Academic performance, attendance, financial indicators, engagement levels, and personal background information are analyzed using a Random Forest classification model to identify students who are at risk of dropping out. In parallel, a RIASEC-based psychometric assessment evaluates student interests and aligns them with suitable career pathways using a career knowledge repository and labor market outlook data. The platform also enables administrators and educators to monitor student progress, schedule interventions, and maintain comprehensive academic records. Developed using Streamlit, Python, SQLite, SQLAlchemy, and Scikit-learn, the proposed system provides a scalable and user-friendly solution for educational institutions. By enabling early risk identification and personalized career guidance, the system aims to improve student engagement, reduce dropout rates, and support informed career decision-making.

making it difficult to detect early warning signs and provide timely support. At the same time, many students struggle to make informed decisions about their future careers because they lack proper guidance regarding their interests, strengths, and available opportunities.

With the advancement of educational data analytics and machine learning, it has become possible to analyze student-related data and generate meaningful insights that support decision-making. Predictive models can evaluate multiple academic and behavioral factors to identify students who may require additional attention and intervention. Similarly, psychometric assessment techniques such as the RIASEC model can help students understand their personality traits and interests, enabling them to explore career paths that align with their abilities and aspirations. These technologies provide educational institutions with effective tools for improving student outcomes and enhancing career awareness.

To address these challenges, this project proposes the **Student Success and Career Navigation Management System (SSC Management System)**, an integrated web-based platform that combines dropout risk prediction, psychometric assessment, career recommendation generation, and intervention management. The system utilizes a Random Forest machine learning model to predict dropout risk levels and a RIASEC-based assessment framework to generate personalized career suggestions. By providing educators with actionable insights and students with structured career guidance, the proposed system aims to improve student retention, promote informed career decision-making, and support overall academic success.

1. INTRODUCTION

Education plays a crucial role in shaping the future of students and contributing to the overall development of society. However, educational institutions often face challenges in identifying students who are at risk of dropping out due to factors such as poor academic performance, low attendance, financial difficulties, and lack of engagement. Traditional methods of monitoring student progress rely heavily on manual observation,

2. LITERATURE REVIEW

Several studies have explored the application of machine learning techniques in education to predict student performance and identify potential dropout risks. Researchers have utilized algorithms such as Decision Trees, Random Forests, Support Vector Machines, and Logistic Regression to analyze academic records, attendance patterns, socio-economic factors, and behavioral indicators. These studies demonstrate that

predictive analytics can help educational institutions identify at-risk students at an early stage and implement timely interventions to improve retention rates. Educational Data Mining (EDM) has emerged as an important field that enables data-driven decision-making and enhances student success through predictive modeling.

In addition to dropout prediction, career guidance systems have gained significant attention due to their ability to assist students in making informed career decisions. Many existing systems employ psychometric assessments, aptitude evaluations, and interest-based frameworks such as Holland's RIASEC model to determine suitable career pathways. Recent research has also incorporated labor market trends and skill-demand analysis to generate more relevant career recommendations. However, most existing solutions focus either on student performance monitoring or career guidance independently, resulting in fragmented support systems.

The proposed Student Success and Career Navigation Management System addresses this gap by integrating dropout prediction, psychometric assessment, career recommendation generation, and intervention management into a single platform. By combining machine learning techniques with interest-based career guidance, the system provides a comprehensive solution that supports both student retention and future career planning, thereby enhancing the overall educational experience.

3. PROPOSED METHODOLOGY

The proposed Student Success and Career Navigation Management System follows a data-driven methodology to identify students at risk of dropping out and provide personalized career guidance. The process begins with the collection of student information, including demographic details, academic performance, attendance records, fee status, extracurricular participation, engagement scores, stress levels, and psychometric assessment responses. This information is securely stored in a centralized database and made available for analysis through role-based access control.

For dropout prediction, the collected academic and behavioral data undergo preprocessing and feature preparation before being provided to a Random Forest machine learning model. The model analyzes various factors affecting student performance and generates a risk score for each student. Based on predefined thresholds, students are classified into Low, Medium, or High-risk categories. The system also generates explanatory insights highlighting the factors contributing to the predicted risk level. These predictions enable

administrators and educators to identify vulnerable students and initiate timely interventions.

For career guidance, students complete a RIASEC-based psychometric assessment designed to evaluate their interests across six personality dimensions: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional. The assessment results are combined with academic readiness and labor market outlook information to identify suitable career pathways. A career matching engine compares student profiles with predefined career profiles and generates personalized recommendations, including required skills, certifications, educational roadmaps, and future growth opportunities. The generated reports assist students in making informed academic and career decisions.

Finally, the intervention management module allows administrators to schedule counseling sessions, monitor student progress, and record support activities for at-risk students. The integration of machine learning, psychometric analysis, career recommendation, and intervention tracking creates a comprehensive framework that supports student retention, academic success, and long-term career development.

4. SYSTEM ARCHITECTURE

System follows a layered architecture consisting of the Presentation Layer, Business Logic Layer, Machine Learning Layer, and Data Storage Layer. The Presentation Layer is developed using Streamlit and provides role-based interfaces for Super Administrators, Class Administrators, and Students. User requests and inputs are processed through the Business Logic Layer, which manages functionalities such as user management, academic data uploads, intervention tracking, psychometric assessments, and career guidance generation.

The Machine Learning Layer contains the Random Forest model used for dropout prediction and the career matching engine based on the RIASEC framework. Academic, demographic, and psychometric data are analyzed to generate risk predictions and career recommendations. All information is stored in a centralized SQLite database managed through SQLAlchemy, ensuring secure and efficient data handling. This architecture promotes modularity, scalability, maintainability, and seamless interaction between different system components.

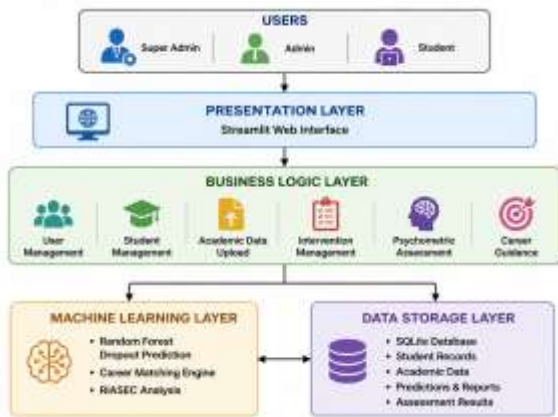


Fig. 1



Fig. 2

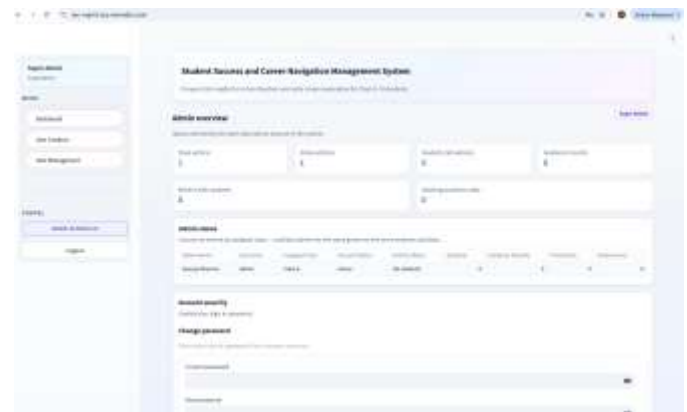


Fig. 3



Fig. 4



Fig. 5

5. IMPLEMENTATION

The Student Success and Career Navigation Management System was implemented as a web-based application using Python and Streamlit. The system follows a modular architecture comprising user interface components, business logic services, machine learning modules, and a centralized database. SQLite was used for data storage, while SQLAlchemy was utilized for database management and communication between application components. The platform provides role-based access for Super Administrators, Administrators, and Students, enabling efficient management of student records, academic data, interventions, and career guidance activities through an interactive dashboard.

The machine learning component was developed using Scikit-learn and employs a Random Forest classifier to predict student dropout risk based on academic, behavioral, and demographic attributes. The career guidance module utilizes a RIASEC-based psychometric assessment to identify student interests and generate personalized career recommendations. The system also includes intervention tracking features that allow administrators to schedule counseling sessions and monitor student progress. Screenshots of the implemented modules and system interfaces are presented below to demonstrate the functionality and workflow of the proposed system.

Then insert your screenshots immediately after this section with captions such as:

- Fig. 2 Login Page
- Fig. 3 Super Admin Dashboard
- Fig. 4 Student Management Module
- Fig. 5 Academic Data Upload
- Fig. 6 Dropout Prediction Results
- Fig. 7 Psychometric Assessment
- Fig. 8 Career Recommendation Report
- Fig. 9 Intervention Management Module

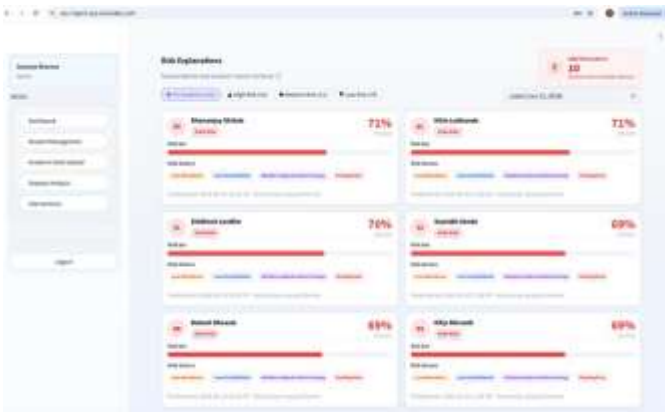


Fig. 6

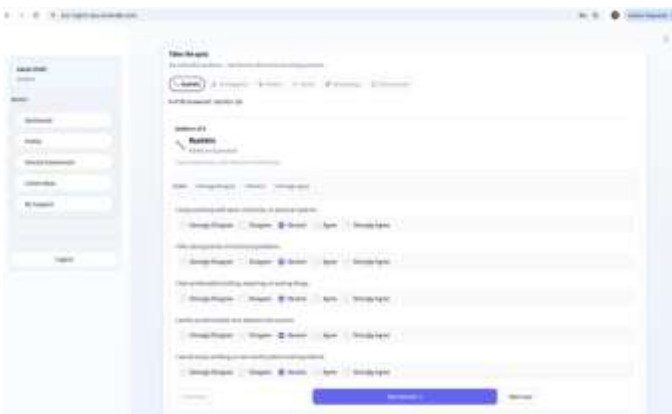


Fig. 7

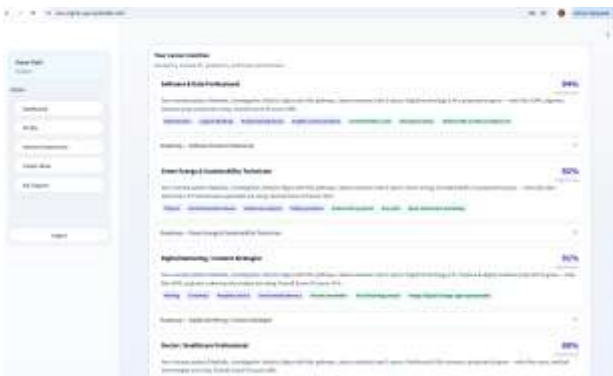


Fig. 8



Fig. 9

6. LIMITATIONS

The current version of the Student Success and Career Navigation Management System has certain limitations. The dropout prediction model is trained using a synthetic dataset and therefore may not fully capture real-world student behavior patterns. The accuracy and reliability of predictions can be further improved by training the model on large-scale historical educational data collected from actual institutions. Additionally, the system currently focuses on students from Classes 6 to 10 and does not support higher secondary education levels.

The career recommendation module primarily relies on RIASEC assessment scores, academic readiness, and predefined career datasets. Since labor market trends and career requirements change over time, the recommendation database requires periodic updates to maintain relevance and accuracy. Furthermore, the application currently uses a SQLite database, which is suitable for small to medium-scale deployments but may require migration to enterprise-grade database systems for large educational institutions with thousands of users.

7. CONCLUSION

The Student Success and Career Navigation Management System provides a comprehensive solution for addressing two critical challenges in education: student dropout prevention and career guidance. By integrating machine learning-based dropout prediction, psychometric assessment, career recommendation generation, and intervention management within a single platform, the system enables educational institutions to identify at-risk students at an early stage and take proactive measures to support their academic success.

The proposed system leverages data-driven techniques to improve decision-making for educators while helping students gain a better understanding of their interests, strengths, and future career opportunities. The implementation demonstrates the effective use of modern technologies such as Streamlit, Python, SQLite, SQLAlchemy, and Scikit-learn in developing an intelligent educational support system. Overall, the system contributes toward improving student retention, enhancing career awareness, and promoting a more personalized and supportive learning environment.

8. ACKNOWLEDGEMENT

We express our sincere gratitude to our project guide and faculty members for their valuable guidance, continuous support, and constructive suggestions throughout the development of this project. Their encouragement and expertise played a significant role in the successful completion of this work.

We would also like to thank our department and institution for providing the necessary resources and learning environment required for this project. Finally, we extend our appreciation to all individuals who directly or indirectly contributed to the development and evaluation of the Student Success and Career Navigation Management System.

9. REFERENCES

1. Marcolino, M.R., et al.: Student Dropout Prediction through Machine Learning Methodologies. Scientific Reports, Nature (2025).
2. Zhang, H., Miller, J., Weber, D.: Predicting Student Dropout Risk Through Machine Learning Analysis of Behavioral Patterns. International Journal of Engineering and Computer Science, 14(11) (2025).
3. Nandi, S., et al.: Sankalp: AI-Powered Career Guidance System. IEEE Conference Proceedings (2025).
4. Ajayi, O.O.: Predicting Student Dropout Risk in Online Learning using Stacked Ensemble Machine Learning and Explainable AI Techniques. International Journal of Computer Applications, 187(40) (2025).
5. Zaidi, S.: AI-Powered Career Coach: A Review on Digital Career Guidance Systems. AIJR Proceedings (2025).
6. Radić, G., Jovanović, L., Bacanin, N., Živković, M.: Identifying and Understanding Student Dropouts Using Metaheuristic-Optimized Classifiers and Explainable AI Techniques. ResearchGate (2024).
7. Zapata-Medina, D., et al.: Improving Automatic Detection of Dropout Risk in Middle and High School Students. ResearchGate (2024).
8. Ghuge, M., Kumari, A., Kamble, T., Raikwar, V.: Envisioning Tomorrow: AI Powered Career Counseling. IEEE, ResearchGate (2024).
9. Kumari, A., et al.: AI-Powered Career Counseling Framework Integrating Large Language Models and Labor Market Trends. (2024).
10. Birajdar, A., Jadhav, S., Shinde, P., Shinde, S., Inamdar, I.Y., Pandhare, S.D.: AI-Powered Career Counselling. International Research Journal of Engineering and Technology (IRJET) (2024).