Chronic Absenteeism Predictor User Manual

1. Overview

This application predicts and analyzes chronic absenteeism risk among students using advanced machine learning techniques. It features five core functionalities:
- System Training - Train the predictive model on your data
- Batch Prediction - Analyze risk for multiple students at once
- Student Verification - Evaluate individual student risk
- Advanced Analytics - Explore detailed insights and visualizations
- System Settings - Configure risk thresholds and interventions

2. Getting Started

First-Time Setup:
1. Start with System Training to train the model
2. Upload historical student data (.csv or .xlsx file)
3. Choose a model type (Random Forest recommended for first-time users)
4. Train the model and click 'Enable All Features'
5. Adjust settings in System Settings as needed

Data File Requirements:
Your data file should include these columns:
- Student_ID - Unique identifier for each student
- School - School name
- Grade - Numeric grade level (1-12)
- Gender - Student gender
- Present_Days - Number of days present
- Absent_Days - Number of days absent
- Meal_Code - Free, Reduced, or Paid
- Academic_Performance - Score from 0-100

3. System Training

Purpose: The System Training module allows you to train a machine learning model on historical student data to predict chronic absenteeism risk.

Detailed Steps:

1. Prepare Your Data: Clean your data and save as .csv or .xlsx file

2. Upload Your Data: Click 'Browse files' or drag and drop your file

3. Choose Your Model Type: Random Forest (Default), Gradient Boost, Logistic Regression, or Neural Network

4. Configure Training Parameters (Optional): Adjust parameters specific to the selected model

5. Train the Model: Click 'Train Model' button

6. Review Results: Examine accuracy metrics and feature importance

7. Enable All Features: Click 'Enable All Features' button

8. Save the Model (Optional): Click 'Save Model' to download the trained model file

Tips for Successful Training:

- Data Quality: Ensure your data is clean and complete

- Data Volume: More historical data usually produces better models

- Balanced Data: Include examples of both high-risk and low-risk students

4. Using Other Modules

Batch Prediction:

- Upload current student data

- The system will generate risk scores for all students

- Filter and sort results to identify high-risk students

- Export results for further analysis

Student Verification:

- Select a student from the dropdown

- View detailed risk assessment and recommendations

- Use the Attendance Impact Simulator to model intervention effects

Advanced Analytics:

- Explore detailed reports including:

- School-level risk analysis

- Demographic patterns

- Grade-level trends

- Attendance vs. Academic Performance

- Risk Heatmap by Grade & SES

- Temporal Attendance Trends

- Intervention Cost-Benefit Analysis

- Geographic Risk Mapping

- Cohort Analysis

System Settings:
- Adjust risk thresholds
- Configure intervention strategies and their effectiveness
- Customize the system to your educational context

5. Troubleshooting

- Data Upload Issues: Ensure your file format matches requirements
- Model Performance: If accuracy is low, try different model types or parameters
- Visualization Problems: Check that your data has sufficient variety
- Unexpected Results: Verify data quality and consistency

6. Best Practices

- Retrain your model at the beginning of each school year
- Combine predictive insights with educational expertise
- Use the 'What-If' simulator to plan interventions
- Review Advanced Analytics regularly to identify patterns
- Document successful interventions for future reference

For additional support, contact your system administrator.