# ShikshaLens – Dropout Prediction & Early Intervention System

## Objective

ShikshaLens is a data-driven platform designed to predict school dropout risks among Below Poverty Line (BPL) students and trigger early interventions using academic, demographic, and behavioral data. The goal is to empower teachers, NGOs, and guardians to act before students fall through the cracks.

## Key Features

- \*\*Predictive ML Engine\*\*

Trained on real-world data (attendance, grades, socio-economic status) to detect early signs of dropout.

- \*\*SHAP-based Explainability\*\*

Provides transparency into model decisions so educators and NGOs understand the "why" behind each risk alert.

- \*\*Intuitive Dashboard\*\*

Helps visualize risk levels across students, schools, and regions – enabling proactive interventions.

- \*\*SMS & WhatsApp Alerts\*\*

Notifies guardians and mentors when a student is at risk, with actionable suggestions.

## Data Sources

- [DIKSHA](https://diksha.gov.in/) – Student learning engagement metrics

- [UDISE+](https://udiseplus.gov.in/) – Student & school demographics

- Local NGO & survey data – Income, parental education level, behavioral feedback

## Tech Stack

| Layer | Technology |

|---------------|------------------------------|

| ML Modeling | Python, Scikit-learn, XGBoost, SHAP |

| Frontend | React / Streamlit |

| Messaging | Twilio SMS, WhatsApp API |

| Hosting | Render / Railway / Heroku |

## Project Workflow

This section explains how ShikshaLens operates from start to finish.

### Data Collection

We aggregate data from multiple sources:

- \*\*DIKSHA\*\*: Learning engagement metrics

- \*\*UDISE+\*\*: Demographics, school attendance, academic data

- \*\*NGO/Survey Data\*\*: Income levels, parent education, behavioral feedback

All data is unified into a consistent format.

### Data Preprocessing

We clean and prepare the dataset:

- Handle missing values

- Encode categorical variables

- Scale/normalize features

- Engineer new indicators (e.g., attendance trends)

Ensures data is model-ready and reliable.

### Model Training

We train ML models such as:

- Logistic Regression

- Random Forest

- XGBoost

Models are evaluated using accuracy, F1-score, and recall for imbalanced classification.

### Model Explainability

We use \*\*SHAP (SHapley Additive exPlanations)\*\* to:

- Highlight the top features contributing to risk

- Provide understandable reasoning behind each prediction

This boosts trust and interpretability.

### Risk Categorization

Based on model output, students are tagged as:

- Low Risk

- Medium Risk

- High Risk

Each risk level maps to different interventions.

### Intervention Engine

Automated actions based on risk level:

- Dashboard alerts for teachers/NGOs

- WhatsApp/SMS notifications to guardians

- Trackable follow-up plans

Helps take timely, data-backed actions.

### Interactive Dashboard

An intuitive web dashboard with:

- Student-level risk profiles

- School-wise heatmaps

- Cluster & regional views for NGOs

Built with React or Streamlit.

### Deployment

Hosted and deployed using:

- \*\*Flask/FastAPI\*\* backend

- \*\*Heroku/Railway\*\* for cloud hosting

- \*\*PostgreSQL/Firebase\*\* for data storage

- \*\*Twilio/WhatsApp API\*\* for communication

Ready for demo or real-world implementation.

## Risk Categories & Interventions

| Risk Level | Action Taken |

|------------|----------------------------------------------------|

| Low | Teacher alert via dashboard |

| Medium | Notify mentor + motivational content to guardian |

| High | NGO escalation + personalized outreach via SMS |

## Future Enhancements

- Mobile app for field mentors

- Gamified attendance boosters for students

- AI-based intervention recommendations

- Integrate with Aptos blockchain for tamper-proof intervention logs

For questions, partnerships, or demo requests:

📧 [rjeevansai2116@gmail.com]

🔗 [https://github.com/jeevan2110/hackthon]

> "Every child deserves a fair shot at education — let’s catch them before they fall."