Deep Learning for Predicting School Dropout in Marginalized Communities

Tresna Dey

Abstract

In the narrow, winding corridors of Dhaka's urban slums and the flood-prone char lands of northern Bangladesh, a quiet but profound crisis is unfolding. While the country has achieved near-universal primary school enrollment, the journey through secondary education remains a leaky pipeline, with a significant number of students failing to complete their schooling.

As of 2026, national statistics indicate a secondary school dropout rate of approximately 32.85%, yet in marginalized communities, this figure often exceeds 50%, reflecting the interplay of economic, social, and environmental vulnerabilities that constrain educational continuity. The challenge today is not merely identifying who has left school but anticipating which students are at risk before the tipping point is reached. Through advances in Deep Learning (DL) and Social Risk Modeling, we are building a "Digital Safeguard," a predictive framework that detects early signals of disengagement from education and triggers timely interventions.

Education in Bangladesh functions not only as a pathway to knowledge but also as a structural determinant of health, social mobility, and long-term economic stability. Research demonstrates that children who drop out of school are 22% more likely to suffer from chronic malnutrition, indicated by underweight BMI, compared to peers who remain in the education system. Moreover, school discontinuation perpetuates intergenerational poverty, limiting employment prospects and reinforcing cycles of vulnerability. Deep data mining of school records, attendance logs, household surveys, and socio-environmental factors reveals that dropout is a multidimensional phenomenon influenced by a cluster of financial, familial, and geographic determinants.

Income remains the most decisive factor. Families earning less than 14,000 BDT per month exhibit a dropout risk more than tenfold higher than wealthier households, underscoring the direct impact of financial

fragility on a child's ability to stay in school. Beyond household income, maternal education emerges as the most significant non-financial predictor of continued schooling. Children of literate mothers are nearly three times more likely to complete secondary education, reflecting the influence of parental literacy on learning support, value attribution to education, and household decision-making. Geographic accessibility further compounds the risk. In rural areas, every additional kilometer of distance to school increases the likelihood of dropout by 15%, a figure that disproportionately affects girls due to concerns around safety, mobility, and cultural norms. Seasonal climate shocks, such as monsoon floods that disrupt transportation and access, exacerbate these vulnerabilities, particularly in low-lying char districts.

While traditional machine learning models like Random Forests and Gradient Boosting have long provided insights into educational outcomes, Deep Learning architectures—including Artificial Neural Networks (ANN) and Recurrent Neural Networks (RNN)—offer a transformative advantage. These models excel at capturing complex, non-linear relationships in data, enabling the identification of subtle, interrelated patterns that precede dropout events. For example, a minor decline in attendance, when coupled with household financial strain or exposure to a climate-related displacement event, can exponentially increase dropout risk—a signal that conventional models may fail to detect. Comparative studies conducted in Bangladesh and similar developing contexts in 2025–2026 indicate that DL models achieve an F1-score of 0.988, outperforming Gradient Boosting models, which reach 0.940, illustrating the significant improvement in predictive precision.

Importantly, the 2026 generation of predictive models integrates fairness constraints to prevent the amplification of existing inequalities. "Fairness Layers" audit the neural networks to ensure that predictions are not biased by gender, geographic location, or socio-cultural factors. This is particularly critical in contexts where girls in remote char areas or students in marginalized urban settlements are at disproportionately high risk of educational discontinuation. By embedding bias mitigation directly into the model architecture, these systems avoid reinforcing structural disparities and instead enable equitable targeting of interventions.

Deep Learning-based predictive modeling transforms dropout detection into actionable Social Risk Modeling. The purpose is not merely surveillance but prevention. Predictive scores allow for the targeted allocation of resources, optimizing both cost-effectiveness and social impact. For instance, AI-driven stipend programs can identify the top 5% of students most vulnerable to financial shocks and provide them with "Precision Stipends," ensuring that limited resources reach those with the greatest need. Similarly, early warning systems can monitor attendance data through RNNs, automatically triggering alerts to parents, teachers, and community workers when a student exhibits a decaying attendance pattern. These alerts can prompt timely home visits, health checks, or counseling, addressing the root causes of disengagement, including illness, labor responsibilities, or household challenges.

The integration of environmental and climate data into DL models further enhances their utility. In flood-prone char regions, predictive systems can forecast localized dropout surges due to school inaccessibility caused by seasonal inundation. With these insights, educational authorities can deploy mobile “Boat Schools” or temporary learning centers to maintain continuity, minimizing learning loss and preventing disengagement. This climate-resilient approach ensures that environmental shocks, which disproportionately affect marginalized communities, do not permanently disrupt educational attainment.

The long-term benefits of reducing dropout through predictive interventions extend beyond education. Retention in secondary school contributes to higher national productivity, with estimates suggesting that closing the secondary completion gap could increase GDP by 3.5%. Health outcomes also improve: educated individuals are more likely to access formal healthcare, adhere to preventive practices, and adopt healthier lifestyles, reducing the prevalence of infectious diseases and non-communicable conditions. Education thereby functions as a multiplier of social and economic development, with the potential to break cycles of poverty and vulnerability.

Moreover, predictive analytics enables proactive policy design. By continuously analyzing dropout signals across diverse geographies and demographics, policymakers can anticipate regional hotspots of disengagement and implement targeted interventions. For instance, rural upazilas with low school density and high economic fragility can be prioritized for new school construction, transportation subsidies, or community-based learning initiatives. Urban slums exhibiting early warning patterns may benefit from after-school programs, nutrition support, or safe transportation services. Deep Learning thus transforms education planning from reactive problem-solving to anticipatory governance, where decisions are data-informed and evidence-based.

The social implications of predictive dropout modeling are equally profound. By transforming “silent struggle” into a visible signal, these systems empower communities to act before crises manifest. Parents gain clarity on their child’s risk status, enabling early behavioral and financial interventions. Teachers and school administrators receive actionable intelligence to provide mentorship, remedial support, or referral to social services. Local governments and NGOs can integrate these insights into broader social protection schemes, aligning educational support with healthcare, nutrition, and climate resilience programs. This holistic approach ensures that dropout prevention is not merely a school-level concern but a cross-sectoral strategy that addresses the complex web of socio-economic determinants.

In the era of “Smart Bangladesh,” Deep Learning becomes a shield for the vulnerable, converting raw educational data into a protective mechanism that safeguards the potential of marginalized students. By embedding fairness, climate awareness, and social sensitivity into predictive models, we ensure that interventions are equitable, timely, and contextually relevant. These models do not replace human judgment but augment it, providing school administrators, policymakers, and community workers with actionable

intelligence to intervene before disengagement becomes irreversible.

In conclusion, school dropout in marginalized communities is not an inevitable outcome but a measurable, preventable phenomenon. Through the integration of Deep Learning, Social Risk Modeling, and predictive analytics, Bangladesh is pioneering a data-driven approach that converts silent educational struggles into actionable insights. In 2026, this framework represents more than just a technological advancement; it is a social instrument that protects opportunity, fosters equity, and strengthens the nation’s human capital. By detecting the warning signals early, allocating resources precisely, and responding with tailored interventions, we ensure that every child—regardless of gender, geography, or socioeconomic background—has the opportunity to complete secondary education. The classroom becomes a pathway to empowerment, not a door that closes too soon, and predictive algorithms emerge as allies in shaping a future where education remains both a right and a bridge to social mobility.