Title: A Geospatial-Machine Learning Approach for Quality Improvement in School Education

Executive Summary

This paper delves into applying Geo-Artificial Intelligence (Geo-AI) technology to address the challenges faced by the education sector at the school level in Meghalaya, India. Artificial Intelligence/Machine Learning (AI/ML) technology digitally transforms education, identifying critical factors affecting education quality and enabling timely intervention to curb quality degradation. The patented tools developed by Deepspatial AI, use applied research, AI, and geospatial modeling to deduce data-driven insights. The geospatial models provide insights into the gaps in school location and accessibility, while the AI/ML models identify the factors responsible for better or poor pass percentages. Datasets used for the analysis include school infrastructure, primary and secondary surveys of teachers, students, student enrollments, student performance, the budget allocated to each school, and other relevant teaching-learning information assimilated with demographic and assets surrounding each school. Overall findings suggest planning regarding teacher vacancies, school infrastructure requirements, school locations, accessibility, education spending, and adherence to standard government norms. The inferences derived from this study have impacted the performances of 15,000 students and are set to achieve an impact on 1 million students in Meghalaya soon.

Introduction

Quality education is a powerful tool with a vast positive impact on an individual's well-being, prosperity, productivity, and quality of life. The involvement of new technology in education can achieve and manage better educational goals. Deepspatial AI is committed to enabling scientific breakthroughs in improving students' academic outcomes using technology and ensuring that AI/ML techniques serve all sectors that will enhance educational outcomes while respecting students' rights. By leveraging AI and ML insights for performance prediction, timely interventions can be made to prevent future deterioration of academic quality. Predictive surveys can help identify the factors responsible for this decline and allow for targeted interventions.

The West Garo Hills (WGHs) district of Meghalaya (India) has been experiencing low student success rates. To address this issue, DSAI patented and proprietary models are applied to historical datasets to extract information from the database about school quality, accessibility, socio-economic status, budget distribution, predictive survey data, and student performance. This information is used to create a decision support system that stakeholders and policymakers can use to improve the quality of school education for more progressive outcomes. The AI/ML models are robust with strong generalization capability with an overall accuracy of 90% and

above. Figure 1 below shows the predicted academic performance of students in the upcoming academic year.

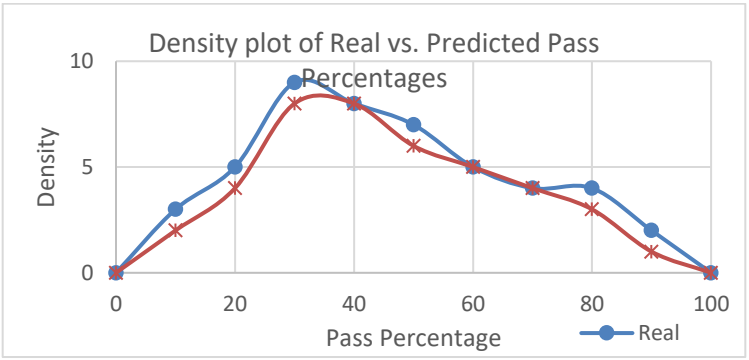


Figure 1: Real versus predicted pass percentages of students in X standard (90%+ accuracy).

Here are some outcomes attained from the objectives (Figure 2):

1. Identification of significant factors impacting students' academic performance.
2. Impact of educational indicators and environmental factors on students' academic performance.
3. Predict and forecast outcomes regarding academic performance, quality of education imparted to students, and education spending based on assimilated variables from multiple sources.
4. Track and govern changes in attributes contributing to the quality of education for students.
5. Recommend and suggest changes based on the model outcomes.
6. Provide a holistic view for a better data-centric planning process.

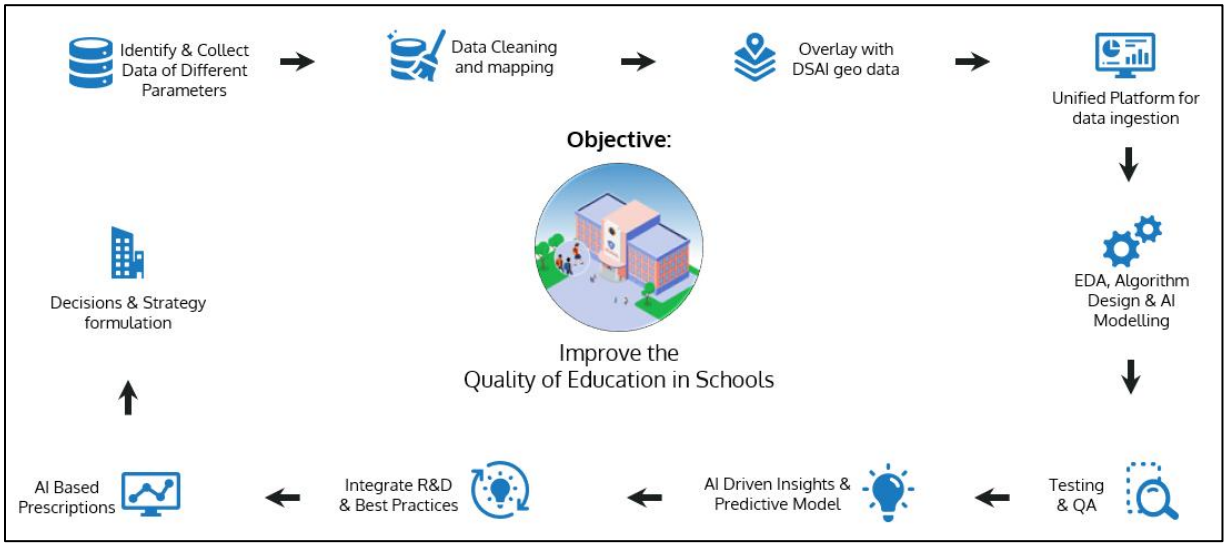


Figure 2: Geo-AI modeling flow from datasets collection, ingestion, modeling, and decision-making

The Problem Statement

In recent years, the West Garo Hills (WGHs) district of Meghalaya has faced a significant challenge due to the decreasing pass percentages and quality degradation of school students, especially at the secondary school education level. However, the disparity in pass percentage among school students between the Khasi Hills (KH) and Garo Hills (GH) regions is very prominent. The WGH district of Meghalaya (India) has been experiencing low student success rates, with an average pass rate of only 27%, across 121 government and government-aided schools. Hence, DSAI has done an investigative study using Artificial Intelligence and Machine Learning to identify factors responsible for a low success rate of students in WGHs, and insights have identified school accessibility gaps, budget allocation analysis, critical factors, and students' performance prediction.

1. **Students retaining**
    - Analysis revealed that schools that retain students from primary level to higher secondary level could educate the students better and thus result in higher pass percentages of the students.
2. **Socio-economic status**
    - Socio-economic status (SES) e.g., parents' income and travel mode, are heavily responsible for the student's success rate. The SES often prevents students from attending school, and other factors like a good home, affordability of private tuition, student resources, and attendance strongly correlate with the school's average pass percentage.
3. **Analysis**
    - Geo-AI and ML models are used to analyze and perform statistical analysis, layered geospatial approach, and assess forecasting and prediction of pass percentage of students in the upcoming academic year.
4. **Accessibility issues**
    - Geospatial analysis is carried out to assess accessibility issues vis-a-vis national standards for school accessibility at different education levels.
5. **Education spending**
    - ML models analyze the correlation between spending efficiency and student success rate. Clustering analysis and feature engineering showed that capital expenditure for school development directly correlates with the student success rate.
6. **Research outcomes**
    - The outcomes are based upon the prediction of academic performance, quality of education imparted to students, and education spending based on multivariate parameters so that one can timely intervene. It further tracks and provides a holistic view for a better data-centric planning process.

The Solution*

Deepspatial AI leverages the power of AI-ML patented models to elucidate the datasets and clarify the factors responsible for improving a school's education. Further, AI predicts and forecasts the performance and education spending by predominantly using decision trees and Random Forests regressors and correlates with the quality of education imparted to the students. The analysis provides a holistic view for better data-centric planning and governs the changes attributed to the quality of education for the students to achieve the ultimate goals. The objectives covered in the study are shown in Figure 2.

1. **Geospatial mapping**
    - Geodemographic profiling of schools for village, block, district, and state levels is carried out. Socio-economic profiling and other interdisciplinary variables are also investigated, impacting school students' outcomes along with the 3-D terrain modeling for understanding the accessibility issues.
2. **Data Monitoring**
    - The data monitoring points are primarily established using DSAI proprietary data and web harvesting. Primary and secondary data are collected from a questionnaire-based survey. Various stakeholders like school demographics, infrastructure, students' records, teachers' feedback, parents' occupation, and average students' performance for each school are collected and ingested into the ML model.
3. **Geospatial AI gap analysis – Location Assessment**

    After intensive research and analysis using against the standard norms, following gaps were found:

    At block level:
    - There are no primary schools in 6 out of 7 blocks in WGH.
    - Rongram is the only block with primary schools accessible from the border area of Gambegre block.
    - Upper primary schools are present in all seven blocks but are inaccessible from almost 50% area of all blocks according to government accessibility norms.
    - All schools have a secondary level but are still inaccessible from parts of Tikrikilla, Rongram, Daddenggri, Gambegre, Selsella, and Dalu blocks, indicating that the secondary school coverage is not 100% as per government accessibility norms.
    - 39% of students travel 0-2 km, 25% travel 2-5 km, 33% travel 5-10 km, and 3% travel 10-20 km for school trips.

    At Village level
    - The geospatial study also investigates the village-level accessibility and gaps (Figure 3) to various categories of schools and education levels.
    - The outcomes are compared with the travel distance for each category of schools (education level – primary, upper primary, secondary, higher secondary) set by government norms.

- The spatial findings help plan the city's new education and travel infrastructure to improve equal access to educational facilities.

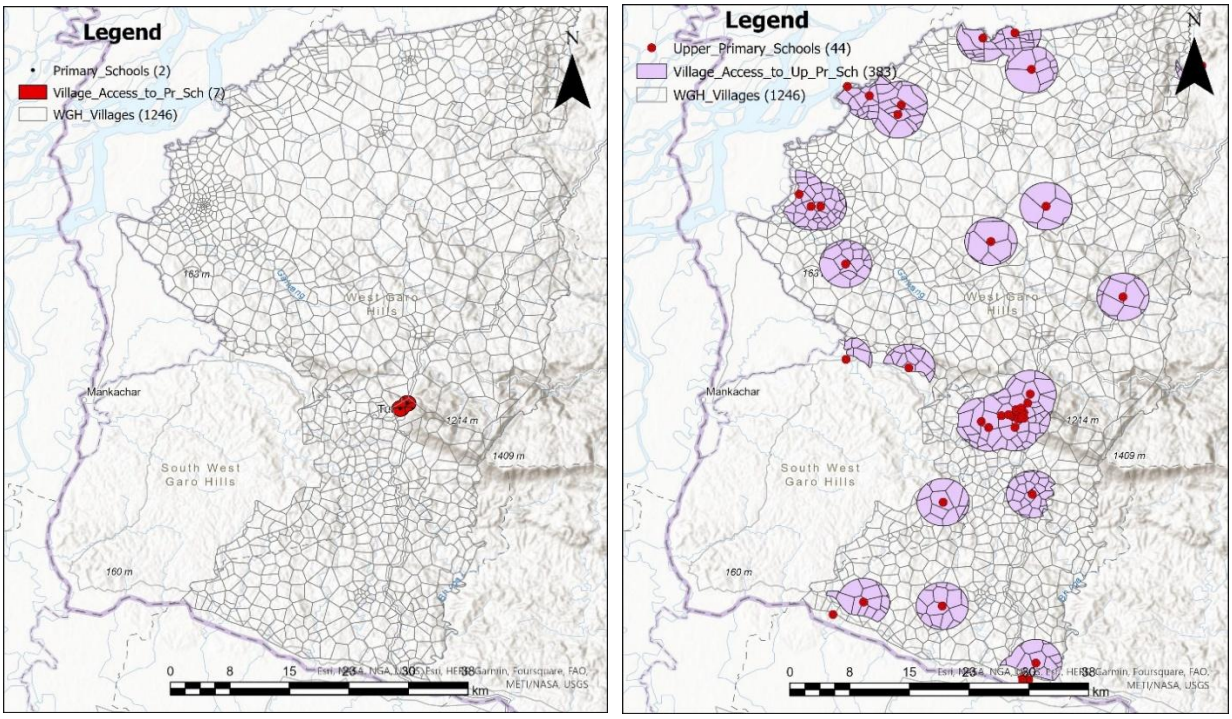


Figure 3: Gaps in the accessibility of (a) Primary schools and (b) Upper primary schools from villages

4. Other vital parameters like classroom adequacy ratio, pupil-teacher ratio, qualified teaching staff, mid-day meal facility, and availability of learning facilities are also compared and supported in easy governance.

4. Statistical Modelling

A highly sophisticated model for statistical analysis is used for dimensionality reduction, feature engineering, and subsequent geospatial AI fusion to deduce valuable insights from the entire datasets. ML algorithms help identify factors responsible for the student's performance in school education (Figure 4). Multivariate Relevance vector machine (MVRVM) model, decision tree regressor, and K-Means clustering analysis are used to obtain the objectives to conclude.

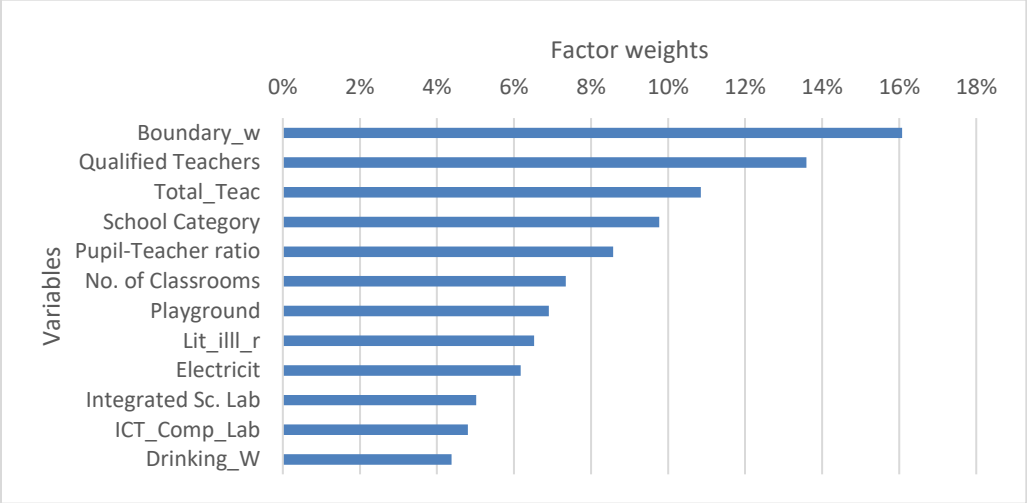


Figure 4: Artificial Intelligence (AI) models-based factors affecting the student success rate

5. Students' performance prediction

DSAI predictive algorithms coupled with alternative data forecast predict the student's success and failures at the secondary level one year in advance to generate advisories and progress on timely interventions.

6. Education spending analysis engine

Non-recurring and recurring budget analysis is carried out to correlate the students' performance with the school's expenditure and per capita analysis. Research indicates that eight schools have a maximum budget per capita but zero pass percentage, which requires immediate intervention (Figure 5).

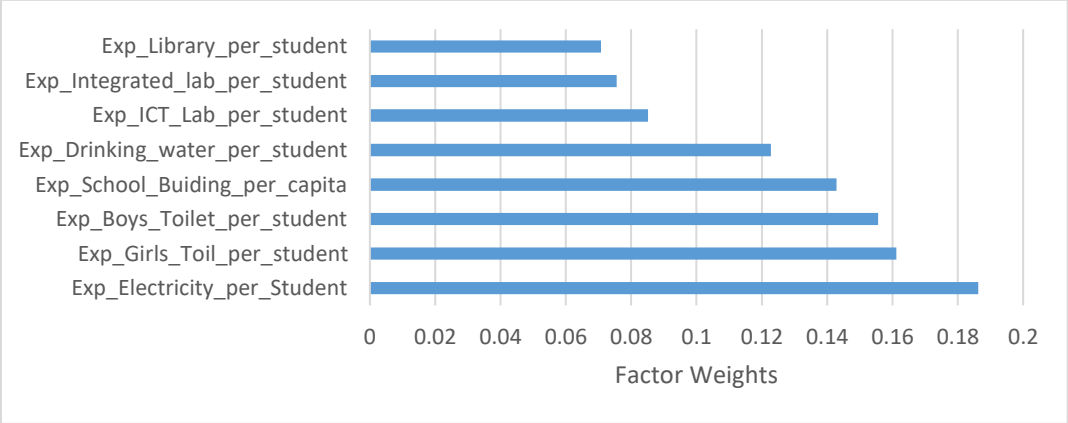


Figure 5: Feature importance of non-recurring budget per capita vis-à-vis the pass % of students

7. Governance

DSAI web platform tracks the performance in real-time and alerts the management system to help prescribe recommendations and monitor its implementation.



Figure 6: Geo-AI-enabled planning, operations, and governance

Working Methodology

The methodology involves four significant steps data aggregation, data pre-processing, feature engineering, and subsequent geospatial AI fusion for deducing valuable insights. First, statistical data cleaning is done using dimensionality reduction and spatial autocorrelation. Further, processed data is geocoded and ingested into an ML model, which is trained to deduce the most important factors affecting the student's performance and the factor weights. Subsequently, these findings provided by the analytical engine facilitate planning, managing, organizing, and enhancing education quality in schools at various levels, as shown in Figure 7.

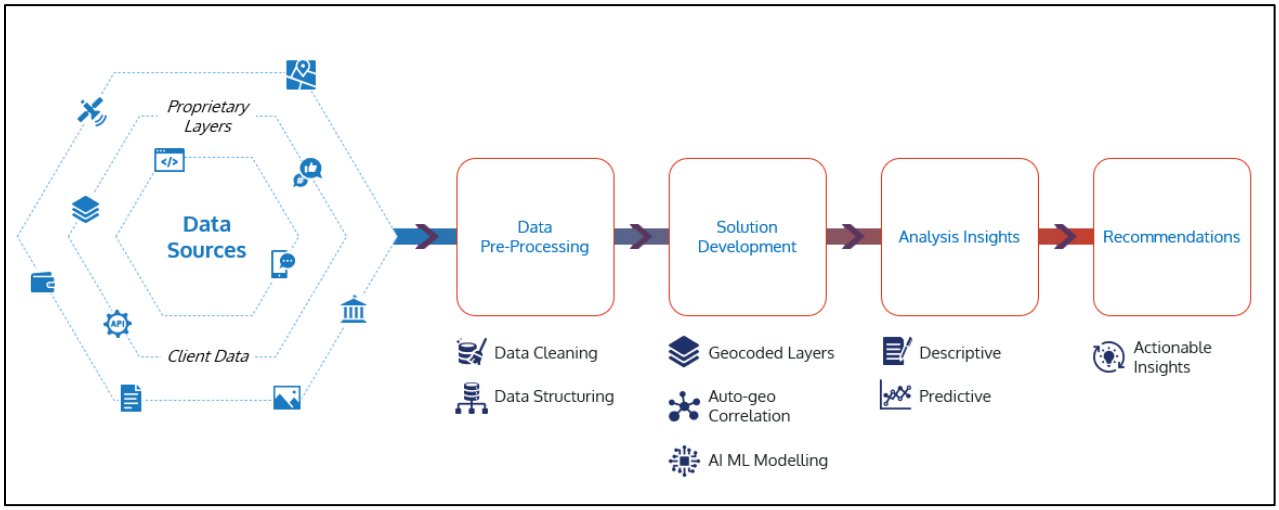


Figure 7: Working Methodology

Conclusion

Deepspatial AI tools have potential applications in transforming the education sector. Thus, low student performance in Meghalaya state (India) school education was investigated to identify the gaps and location access, critical factors affecting students' outcomes, education spending, performance forecasting, etc. Datasets and other relevant information surrounding schools and

various stakeholders were obtained from the survey. The machine learning model is trained and identifies the critical factors affecting the student's poor percentages. The geospatial AI model analyses block-wise gaps in school location and accessibility issues. Education spending engines recommend the intervention in budget allocation to schools based on the school's performance. The insights suggest in planning human resources, infrastructural requirements, school accessibility, location issues, budget expenditure, etc. The web platform supports easy governance by supplying real-time information, gaps, and ready recommendations to be implemented for improvements.

ABOUT THE CONTRIBUTORS

Dr. Rahul Kushwah is CEO at Deepspatial Inc., Canada. With over a decade of experience in the medical sciences, highly skilled innovator, entrepreneur and team leader. His experience includes positions as Federal Government Scientist, Research Officer within the HHT branch of NRC of Canada along with Professorship at the University of Ottawa. He also holds a PhD from University of Toronto.

Dr. Bushra Zaman is Chief AI Officer at Deepspatial AI, and Adjunct Professor, CEE, College of Engineering, Utah State University, USA. She has over 17 years of industry, research and teaching experience with specialties in Civil and Environmental engineering, water resources engineering, Geographical Information Systems (GIS), ERDAS Imagine, Remote sensing techniques and image processing, AI/ML Modeling, Predictive analytics, evolutionary computation etc.

Dr. Chhotu Ram is working as Associate Director in Research and Development at Deepspatial AI, India. He has around 14 years of industry, research and teaching experience with specialties in article writing, research guidance, and books in multidisciplinary areas.

Er. Aisha Sharma is Associate Director in Data Science and Engineering in Deepspatial AI, India with 10 years of experience in teaching, research and technical projects. She has worked as a hydrologist in IISc, Bangalore and thereafter as a Data Scientist in multidisciplinary areas, with expertise in GIS tools and ML modeling.

Enquiries and contact:
rahul@deepspatial.ai (Dr.Rahul Kushwah, CEO)
dev@deepspatial.ai (Mr. Debojyoti Purkayastha, Chief Business Officer)

CONTRIBUTING AUTHORS:

Dr. Rahul Kushwah

Chief Executive Officer, Deepspatial Inc., Canada

Dr. Bushra Zaman

Chief AI Officer, Deepspatial Inc., Canada

Dr. Chhotu Ram

Associate Director, Research and Development, Deepspatial Asia Pvt. Ltd. India

Er. Aisha Sharma

Associate Director, Data Science and Engineering, Deepspatial Asia Pvt. Ltd. India