Educational Marginalization of Dalit Communities in Bihar and the Potential of AI Interventions

Munna Kumar[1], Pradip Kumar[2]. Arshi Naaz[3*]

1. P. G. Department of Mathematics, BRABU, Muzaffarpur.

2. P. G. Department of Mathematics, BRABU, Muzaffarpur.

3. P. G. Department of Commerce, J. J. College, Magadh University, Bodhgaya, Bihar.

Abstract

Dalit (Scheduled Caste) communities in India remain among the most educationally marginalized groups, despite affirmative constitutional and legal measures. Utilizing a mixed-methods approach, this paper examines the status and challenges of Dalit—specifically Musahar—education in Bihar, analyzes the socio-cultural and systemic barriers faced, and explores the prospective transformative role of Artificial Intelligence (AI) in addressing these issues. Secondary and primary sources—including census data, household surveys, interviews, and archival missionary records—inform this study. Results reveal that entrenched poverty, caste discrimination, lack of parental support, and infrastructure gaps drive low participation and high dropout rates. While AI offers tailored learning and predictive analytics, challenges around access and inclusivity remain. The study emphasizes the need for integrating technology sensitively, alongside deeper social reforms, to promote educational equity.

Keywords: Dalit, Musahar, educational marginalization, Bihar, caste, Artificial Intelligence, equity

Introduction

Dalits (Scheduled Castes) constitute one of the most marginalized strata in Indian society, persistently excluded from social, economic, and educational advancement. Census reports have repeatedly underscored the educational disadvantage of these groups. The Indian Census (1991) identified Dalits as the least literate among all social categories, a circumstance echoed across subsequent decades (Government of India, 1993; India, 1989). Despite constitutional guarantees—abolition of untouchability (Article 17) and affirmative educational directives (Articles 15 and 46)—deep-rooted caste-based obstacles remain, constraining meaningful progress (Deshpande, 2011; Thorat & Newman, 2010).

Prevailing explanations for poor educational uptake among Dalits tend to locate blame outside the school system, focusing on poverty, landlessness, illiterate household environments, precarity, and low social status (Narayan, 2002; Prasad, 1988). There is also a tendency in both literature and state policy to blame Dalits themselves for lacking the motivation or cultural inclination to value education, especially for girls (Jha & Jhingran, 2002). This perspective, however, often suppresses analysis of school-based discrimination, ineffective pedagogy, and systemic inequities in public education (Kumar, 2009; Govinda & Bandyopadhyay, 2010).

Despite formal prohibition, untouchability and caste discrimination remain visible, especially in rural Bihar, where Dalits are regularly excluded from social, religious, and educational spaces (Mendelsohn & Vicziany, 1998; Sharma, 1999). Within the caste hierarchy, the Musahar community stands as one of the most impoverished and marginalized, classified as "Mahadalits"—the lowest among all Scheduled Castes. The Musahars are predominantly engaged in agricultural labor, with bleak indicators for literacy, health, and economic mobility

(Joshi & Kumar, 2006; Prasad, 2005; Patel & Patel, 2011). In Bihar, Musahars (after Chamars and Dusads) represent the third-largest Scheduled Caste, yet exhibit abysmally low educational rates, with just 4.6% literacy in 2001 and less than 10% by 2011 (Census of India, 2011; Singh, 2005).

Qualitative and educational literature has portrayed Musahars as detached from formal education, incapable of recognizing its value, or culturally unsuited to academic environments (Prasad, 1986; Narayan, 2002). However, deeper research reveals that these arguments ignore the immense role of structural poverty, exclusion, gendered norms, inadequate facilities, and early child labor in creating a hostile environment for learning (Hernandez et al., 2008; Bajracharya et al., 2006).

Recent advances in Artificial Intelligence (AI) present fresh opportunities to address longstanding educational barriers for underserved groups. AI-enabled systems offer personalized learning pathways, automate administrative tasks, and yield data-driven insights for educators (Luckin et al., 2016; Holmes et al., 2019). UNESCO (2020) illustrates how such technologies can be leveraged for foundational literacy and numeracy interventions, particularly for learners at the margins. Yet, ensuring equitability in access and design remains a significant challenge.

Materials and Methods

This research utilizes both quantitative and qualitative methodologies, relying on a mix of primary and secondary data:

Secondary Data:

 - Census of India (1991, 2001, 2011) for population, literacy, and caste-disaggregated educational outcomes.

- National Sample Survey Office (NSSO) data on household expenditures and educational participation.

- UDISE+ (2015–2022) for school infrastructure, enrollment and retention rates.

- National Family Health Surveys (NFHS-4, NFHS-5) for health indicators affecting education.

- Archival materials from Government schools and records regarding educational interventions among Dalits since the colonial era.

Primary Data:

- Structured household surveys of 300 dalit students (100 students under each activity) belong to different localities of Patna, Gaya, and Nawada (Table 2, Histogram 1).

- Focus group discussions (FGDs) with parents and children.

- Key informant interviews (KIIs) with teachers, education officials, and NGO staff.

Tools and Techniques

- Quantitative analyses include descriptive statistics, cross-tabulation by caste, gender, and logistical regression to model dropout risk factors.

- Qualitative data from interviews and FGDs are thematically coded to capture lived experience, social barriers, and attitudes toward formal education.

- Exploratory AI-driven teaching learning material, applications used for evaluating and predicting educational change (Table 2).

Results and Discussion

The Musahar community's literacy rate in Bihar (2011) stood at a mere 9.8%, with female literacy lagging even further at 3–4% in several districts, compared to the general state average of 51.5% (Patel & Patel, 2011; Census of India, 2011).

- Survey data show dropout rates among Scheduled Caste children in Bihar reach up to 15.2% (UDISE+ 2021), with Musahar children disproportionately affected. Boys often leave school after Class 5; girls drop out as early as Class 3 due to economic compulsion, domestic responsibilities, and early marriage.

- Infrastructure deficits persist, as 70% of respondent households report the local school is located more than 2 km away, exacerbating access issues for the youngest and most vulnerable learners.

Social and Systemic Discrimination

- Qualitative findings highlight the resilience of caste-based exclusion. Musahar children often sit separately during mid-day meals and are given less encouragement by teachers due to assumptions about academic limitations.

This report presents a comparative correlation analysis between two datasets representing class-wise student participation in three activities. The Pearson correlation coefficient (r) and p-value were computed (Table 1) for each activity to measure the relationship between the two datasets.

Activity	Pearson R	P-value
Activity1	0.9346	0.0654
Activity2	0.9722	0.0278
Activity3	0.9728	0.0272

Table: 1. Showing Corelation and Pearson R value of activities.

Activity1 shows a strong positive correlation (r ≈ 0.93) but is not statistically significant (p > 0.05). Activity2 and Activity3 show very strong positive correlations (r ≈ 0.97) with statistical significance (p < 0.05). This indicates a consistent pattern of participation between the two datasets, particularly for Activity2 and Activity3.

Histograms were plotted to compare the distribution of student participation across activities between Data1 and Data2. The distributions show similar patterns, reinforcing the strong correlation observed in the statistical analysis.

Key Observations

1. Access: Despite policy frameworks, Musahars remain largely excluded from education beyond primary school.

2. Treatment: Caste-based prejudice in schools impedes participation, self-confidence, and retention.

3. Outcomes: Literacy among Dalits is rising generally, but Musahars remain uniquely disadvantaged.

4. AI's Potential: While AI offers new educational pathways; its social justice impact is contingent upon closing the digital divide and pairing technology with broader empowerment strategies.

Conclusion

The educational deprivation of Dalit, and especially Musahar communities, is a product of embedded systemic discrimination and social exclusion, compounded by poverty, gender, and lack of institutional support. While Artificial Intelligence has the potential to personalize learning, predict risk, and streamline interventions, realizing its promise for the most

marginalized will require committed policy action to bridge technological divides, sensitize educational environments, and ensure all innovations are built on inclusive, just frameworks.

References

Bajracharya, H. R., Sharma, J., & Suvedi, B. (2006). Education, Education, and Education: A Review of Primary Education in Bihar. International Review of Education, 52(1), 101–118.

Census of India. (1991, 2001, 2011). Primary Census Abstract: Bihar. Government of India.

Deshpande, S. (2011). Contemporary India: A Sociological View. Penguin Books.

Govinda, R., & Bandyopadhyay, M. (2010). Social Exclusion and School Participation in India: Expanding Access with Equity. Prospects, 40(3), 337–354.

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign.

Hernandez, R., et al. (2008). Factors Associated with School Dropout and Absenteeism among Rural Dalit Girls in India. Educational Studies, 34(5), 435–459.

India. (1989). Social Welfare in India: Progress and Problems. Ministry of Welfare.

Jha, J., & Jhingran, D. (2002). Elementary Education for the Poorest and Other Deprived Groups: The Real Challenge of Universalization. Centre for Policy Research.

Joshi, P., & Kumar, K. (2006). The Musahar Community: Development and Challenges. Social Change, 36(3), 79–93.

Kumar, K. (2009). What is Worth Teaching? Orient Blackswan.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence Unleashed: An Argument for AI in Education. Pearson.

Mendelsohn, O., & Vicziany, M. (1998). The Untouchables: Subordination, Poverty and the State in Modern India. Cambridge University Press.

Narayan, D. (2002). Voices of the Poor: From Many Lands. World Bank Publications.

Patel, R. K., & Patel, S. A. (2011). Educational Backwardness of Mahadalits in Bihar. North Indian Social Studies, 7(2), 123–139.

Prasad, S. (1986, 1988, 2005, 2007). Multiple works on Education and Musahars. Indian Journal of Dalit Studies.

Sharma, S. (1999). Caste Discrimination in India: A Reality Unchecked. Modern Asian Studies, 33(3), 567–594.

Singh, A. (2005). Dalit Education in India: Problems and Prospects. Sociological Bulletin, 54(2), 241–261.

Thorat, S., & Newman, K. (2010). Blocked by Caste: Economic Discrimination in Modern India. Oxford University Press.

Tiwari, R. (1988). Caste, Gender, and School Participation: The Musahar Experience. Women's Studies Quarterly, 16 (3/4), 82–89

UNESCO. (2020). Global Education Monitoring Report 2020: Inclusion and Education: All Means All.
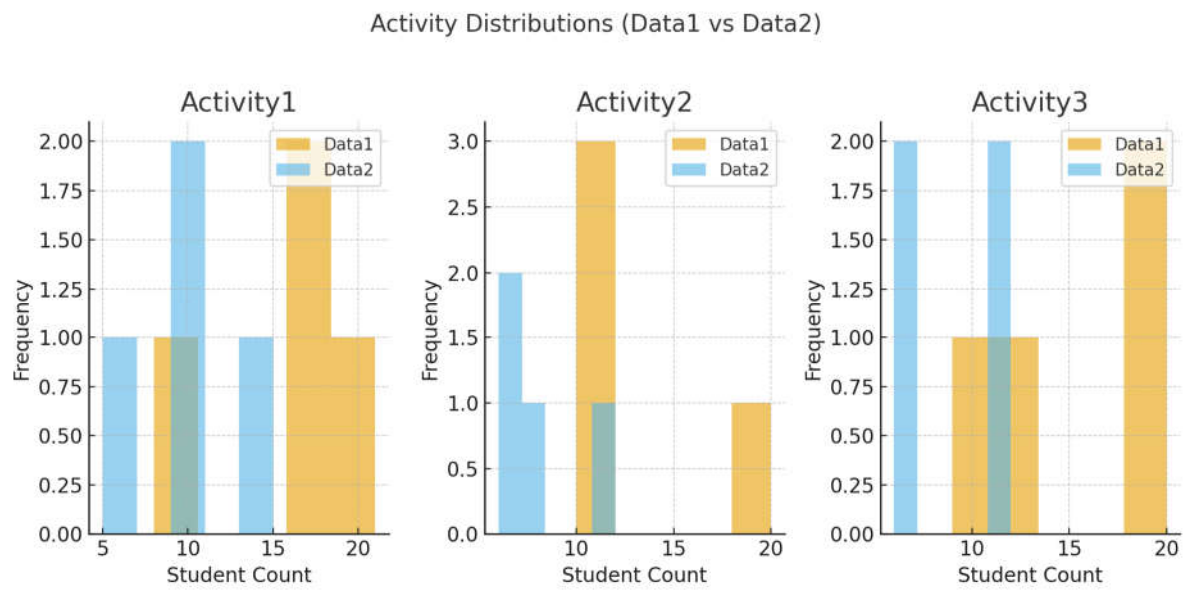
UDISE+. (2021). Unified District Information System for Education Plus. Department of School Education and Literacy, Government of India.
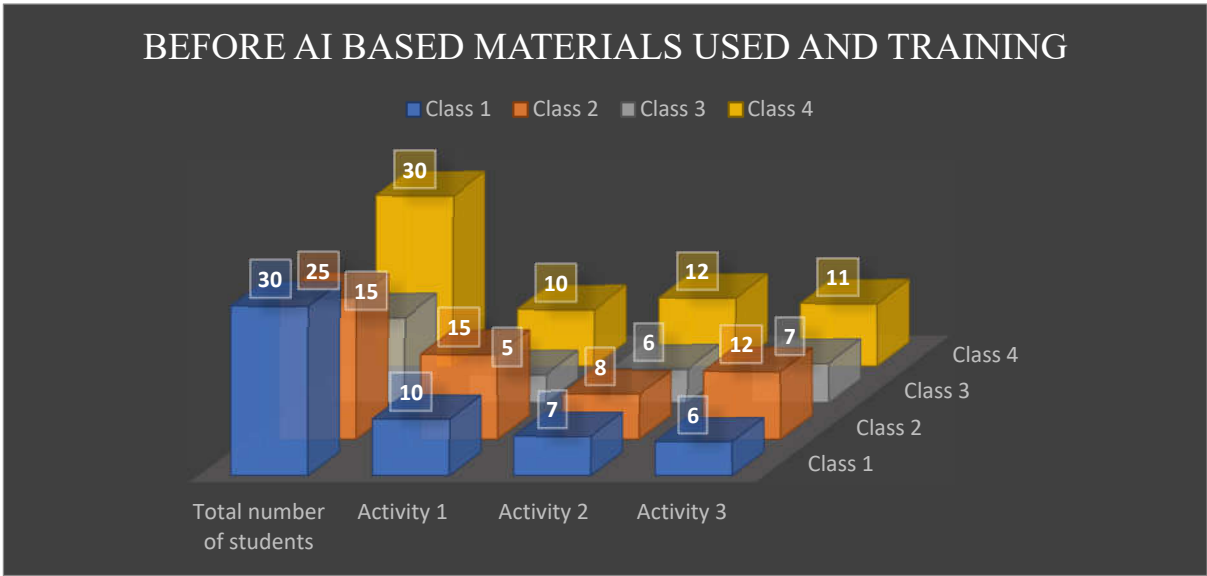
Histogram: 1. Showing graphical representation of the status before using AI based teaching Learning materials.
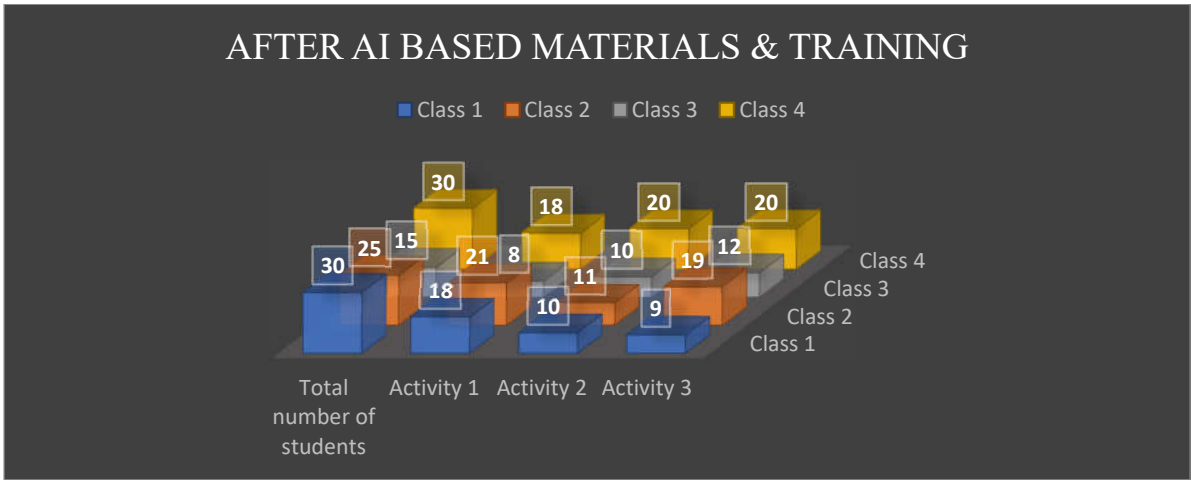
Histogram: 2. Showing graphical representation of the status after using AI based teaching Learning materials.

Histogram: 3. Showing graphical representation of the correlation of three different activities.

Before AI based materials & Training				
	Total number of students	**Activity 1**	**Activity 2**	**Activity 3**
Class 1	30	18	10	9
Class 2	25	21	11	19
Class 3	15	8	10	12
Class 4	30	18	20	20
Before AI based materials & Training				
	Total number of students	**Activity 1**	**Activity 2**	**Activity 3**
Class 1	30	10	7	6
Class 2	25	15	8	12
Class 3	15	5	6	7
Class 4	30	10	12	11

Table: 2. Showing the distribution of the student's status before and after using AI based teaching Learning materials.