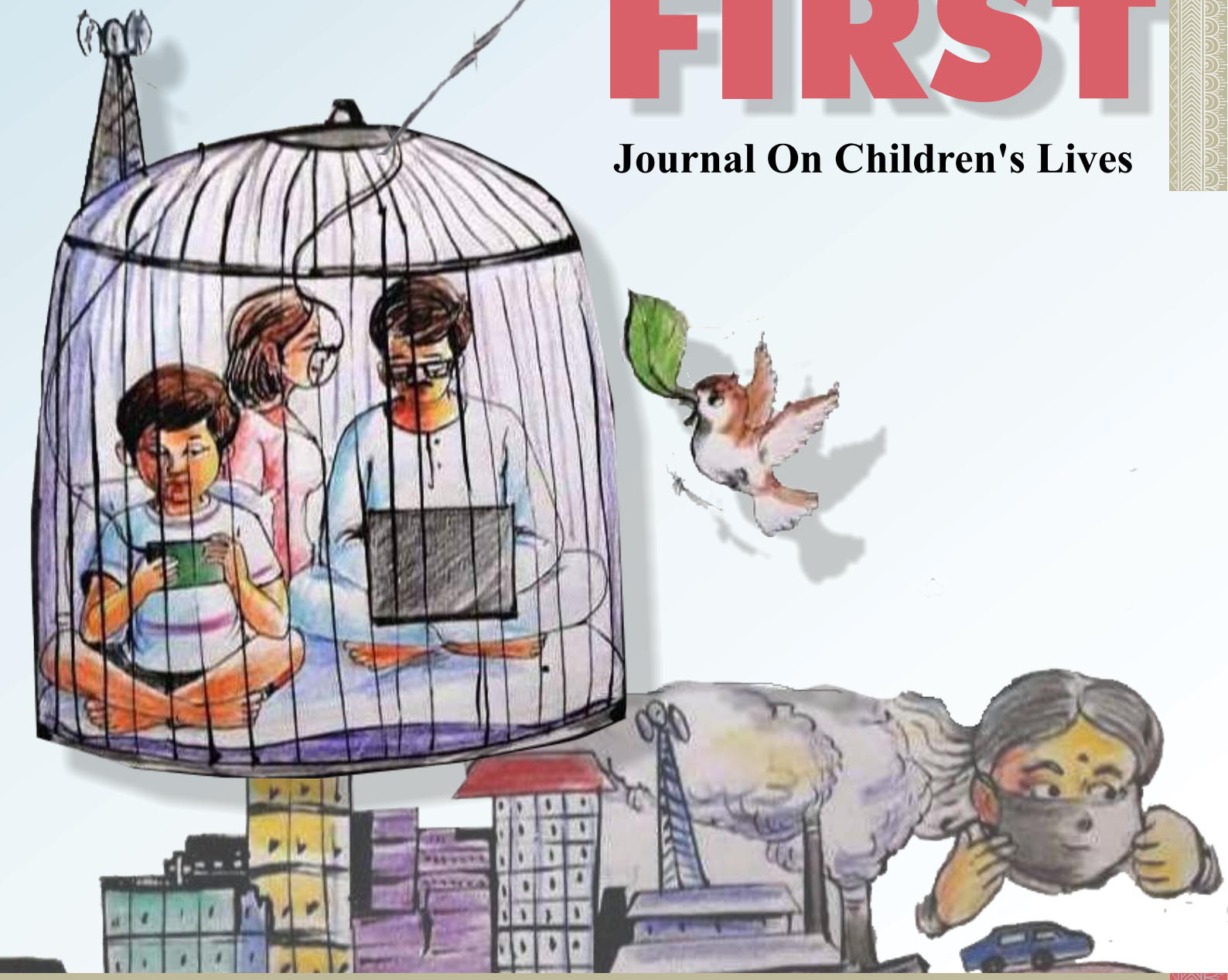
VOL. 1 ISSUE 1

CHILDREN FIRST

Journal On Children's Lives



Impact of COVID-19 Pandemic on Lives of Children

Aryanraj , 10 years

05 BEST PRACTICES

Trishit Gupta, 13 years

Using Machine Learning to Accelerate School Enrolment of India's Most Vulnerable Girls

Alison Bukhari

I. INTRODUCTION

This paper highlights the 'precision targeting' approach of Educate Girls, to solving the crisis of out-of-school girls across India. If we are to meet the SDGs by 2030, which include achieving gender equality and quality education for all, we need to move fast and accelerate the rate at which we find and enroll girls in India's most remote villages.

The reasons behind girls' exclusion from school are complex, requiring mindset change and deep community engagement. However, the first problem is also to find the girls. With the use of machine learning predictions, we can locate the majority of girls in remote villages far quicker than with traditional methods; and with limited time and resources enroll them into school, helping to secure their education and their future.

Educate Girls is a non-profit organization that mobilizes communities for girls' education in India's rural and educationally backward areas. Working in partnership with the Government, Educate Girls operates in 18,000 villages of Rajasthan, Madhya Pradesh, and Uttar Pradesh. To date and with the government, we have ensured that 950,000 out-of-school girls are back in the classroom and have improved learning outcomes for more than 1 million children.

II. THE CHALLENGE

Despite huge advances in universal access to education in India, we still have the third-largest out-of-school girl population in the world. All too often discriminatory mindsets and practices keep girls at home or working in the fields. In 'normal' times it is estimated that 4.1 million girls aged 7 to 14 are missing from India's classrooms and this is widely considered to be a conservative number. It is, however, hard to know what the numbers are, as the problem becomes invisible if the girls are invisible.

Added to this we are now facing the impact of the Coronavirus pandemic and school closures. The pandemic has been catastrophic for children, already facing a learning crisis, and now they have been kept out of school for a whole academic year. Behind the huge numbers of out-of-school children, hide the girls who have faced the shadow pandemic, bearing the burden of the lockdown and a return to old patriarchal roles in homes, sometimes facing abuse or child marriage. We are deeply concerned that when the schools reopen, a great many girls will be missing from the register.

Critical to enrolling girls into school is the local knowledge and community mobilization, but to implement enrolment strategies we first have to find the eligible girls. Villages have grown exponentially

Alison Bukhari, International Director, Educate Girls US

during the COVID-19 crisis, as huge numbers of families have returned home from the cities. The reverse migration in many areas has dramatically increased the number of children who should be in village schools, once they reopen. In one of our districts, Chitrakoot, our door-to-door survey has identified multiple villages where this has become a huge issue.

One village alone has seen 18 families return from cities across the country, with 27 children now waiting for admission in the village school. We have met drivers from Mumbai, a cook from Gujarat, all scarred by the trauma of the pandemic, lockdown, and now a lack of job security which has persuaded them to stay back in their village.

Planning for this will be an enormous challenge for the Government but, as this paper explains, Educate Girls has designed tools that can accurately predict where to find areas of the high density of out-of-school girls through a combination of community outreach and machine learning algorithms. Our key insight is that a large percentage of India's girls reside in only a few of its villages – so if we can predict where the highest out-of-school girl populations will be, we can make our work far more targeted and enroll many more girls quicker. And how we do this is through machine learning.

III. HOW WE GAINED OUR KEY INSIGHT

During our 12-year history, household visits and community outreach have been at the heart of the organization's success in getting girls back to school. However, the process has proven time-consuming and costly over the years, making our target of getting 1.5 million girls back to school by 2024 an immense challenge.

In the past, we would create a baseline before working in a district and base our decision to work there on available public data. We would then go house to house and create our door-to-door survey, meeting families and identifying girls who were out of school, or at risk of not being enrolled.

Often, we would find that the public data was not accurate enough on its own and some villages would have a great many more out-of-school girls than we had thought. Indian Census is now outdated, and DISE data is also often 2 to 3 years old when we use it, and there isn't a single field in either dataset that captures how many out-of-school girls there are. Hence there is a need to find a way to triangulate from a lot of different data sources to solve this problem.

It was during the delivery of our Development Impact Bond between 2015 and 2018 that we gained a key insight into a potential way to improve the efficiency and impact of our work. The nature of this outcomes-based funding contract demanded that we dramatically improve the way we gathered and analyzed data. We started to see that in certain areas, 50% of the out-of-school girls were found in just 10% of the villages and 23% of villages had no out-of-school girls at all.

This information led us to the conclusion that if we could accurately predict where to find the areas where most out-of-school girls were living, we could precision target our work, conduct the door-to-door survey in half the number of villages, and still find the majority of out-of-school girls. Essentially, if we could go where the need is greatest, we could enroll many more girls with a similar budget, within a shorter time frame.

IV. THE SOLUTION

Together with our partner IDinsight, we decided to build a machine-learning algorithm to help us learn from our previous

surveys that measured how many out-of-school girls were in each village where we worked. By combining this historical program data with public data (such as the 2011 census and DISE educational dataset), we could teach an algorithm to identify villages in new geographies where we can expect to find the most out-of-school girls.

The model was built by taking our existing door-to-door survey data, of around 3 million households, gathered during the earlier phase of our work, across 8,000 villages, and merging it with publicly available administrative and education data (sources that include 313 variables, i.e. indicators that look at socio-economic and socio-cultural parameters). IDinsight tested various machine learning models and gradually improved the accuracy of the predictions. Next, using these predictions, they created heat maps that identified geographic clusters of villages, or 'hotspots' with the highest density of out-of-school girls.

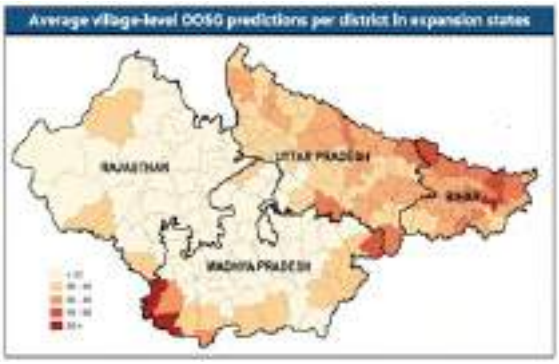


Figure 1 : *Avg. village-level OOSG predictions*

The machine learning algorithm now enables us to make the most of our field resources by selecting the high-burden clusters or 'hot spots', as a priority. Spot check surveys then allow us to assess accessibility to villages and plan for reduced travel time within clusters. Put simply, by hiring and deploying staff and volunteers to these clusters of villages, rather than working in all the villages in a given administrative district, we have been able to find villages with two and a half times as many out-of-school girls per village, compared to those found through our traditional survey techniques.

It is also important to note that of the over 2 lakh children aged 5-14 that were identified as out of school, using these predictive models, approximately 64% are from scheduled tribe (ST) communities, 9% from scheduled caste (SC) communities and 18% from other backward castes (OBC).

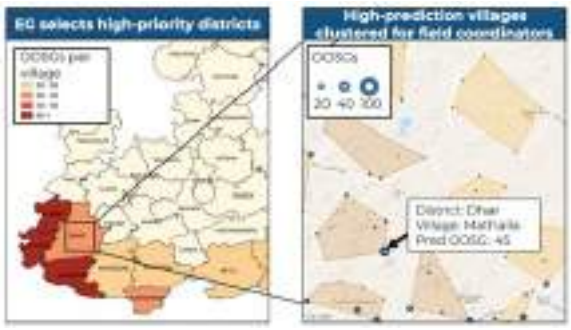


Figure 2 : *High-priority districts*

In other words, over 90% of the children are from underprivileged communities, for which the equalization of opportunity through education has not yet been fully realized.

Since the algorithm was first created in 2018, we have refined it every year to increase precision – adding new data sources and data from more villages, as our door-to-door survey has grown to 5 million households. The plan is to refine the predictive model and apply it to 212,000 villages across 4 states in the Hindi belt, namely Rajasthan, Madhya Pradesh, Uttar Pradesh, and Bihar. Educate Girls will then be able to look at the data at a district level and prioritize the top 15-20% of villages per district, enrolling more girls into school faster. Over the next 5 or more years, the total impact of this algorithm will amount to more than 500,000 additional out-of-school girls enrolled in school than would otherwise have been possible.

V. WHERE TECHNOLOGY MEETS COMMUNITY

What is so pioneering about this approach is that it combines sophisticated technology such as machine learning and geo-tagged surveys, using smartphones, with readily available village resources (young, energetic, mission-aligned men and women) and deep community knowledge.

The use of technology has dramatically improved our targeting and efficiency, but it has to go hand in hand with the the team of village gender champions or Team Balika working on the ground, currently numbering 18,000 and a figure that will go up to 35,000 as we expand our work over the next 5 years.

Team Balika volunteers receive intensive training, ongoing mentorship, and support to help them identify the best strategies for enrolling the girls, who are currently working or kept at home. They are overseen by a cadre of field staff, again locally hired, to manage the teams and adapt on an ongoing basis the implementation plans, according to the village-level realities.

Strategies involve sitting with families and village elders, challenging damaging social norms, and presenting the benefits of and right to an education. Before the pandemic, we would then work in and alongside the school through classroom-based remedial education and support to school management committees. And now with the school closures as a result of COVID-19, we have a community-based curriculum delivered in safe spaces in target villages.

At the end of the day, it is the behaviour change communications that tips the balance and changes mindsets, such that families are then willing or happy to send their girls to school. This can be a long and engaged process, but with the right precision targeting, we can work in the areas of most need and accelerate our ability to reach a large number of girls quicker and with fewer resources.

VI. RECOMMENDATIONS FOR WORKING WITH THE GOVERNMENT

The pandemic and resulting lockdown continue to create incredibly difficult times for our communities but also learning for us as an organization. As the pandemic hit, the value of our data and our ability to identify those most in need through predictive analytics, became even clearer to us. When, due to the lockdown, we had to abandon our door-to-door survey, we were still able to help the government in their ration distribution, by using our predictive data to identify some of the most remote villages in need of government schemes. More often than not, as we experienced, the government had resources, but remote districts were not able to access them because of a lack of information. The predictions we can make as a result of the machine learning algorithm, are enabling Educate Girls and the Government to optimize operational inputs and locate those most in need.

And what is more, an out-of-school girl is likely to be much more vulnerable to child trafficking, child labour, malnutrition – both stunting and wasting, health issues linked to a lack of immunization, child and maternal mortality. The list goes on. Based on our experience of building an algorithm to help us identify out-of-school girls, and then our experience using our data to support the government during COVID-19, we feel confident that this approach could apply to other social issues and for wider geographical targeting. Cities could use this approach to identify the most vulnerable neighbourhoods, districts, and state administrations along with NGOs can accelerate their ability to meet child-centric Sustainable Development Goals through the

PC: Department of Education, GNCTD

ability to locate hotspots and deliver interventions quicker.

VII. CONCLUSION

If a problem is not quantified and located, it can become invisible. We cannot let another generation of girls become invisible and lose them to illiteracy and poverty and we have to find every last girl. There is an economic, health, and climate imperative to educate as many girls as we can, as quickly as possible, and what we need is the focus. If we can find the hotspots where there is the highest number of girls who are out of school and go there first, we can tackle large numbers of out-of-school girls, quicker and at a lower cost.

Educate Girls' five-year strategy is based on the prediction that 40% of India's out-of-school girls, c.1.56 million girls are highly likely to be found in just 5% of its 650,000 villages. While the numbers could have greatly increased post the pandemic, we know we can solve the problem much faster if we use our predictive analytics and target our resources to the villages with the greatest need.

Our next issue:

We invite authors from across the globe to write for our journal's next issue to be published in April 2022 ! The theme is

"Second Year of COVID-19 Pandemic- Disrupted Childhoods, Disrupted Education"

COVID-19 has not yet ended and nor has our curiosity and need to understand the monumental impact it has left on young children. Children may not be the face of this pandemic, but the corona crisis has created significant disruption in their learning and development -- emotionally, mentally and educationally. To document, analyse and address its continued impact on children in the Indian context, *Children First - Journal on Children's Lives*, a peer reviewed, bi -annual publication, invites submissions for its second issue.

Authors can submit their papers under five categories, namely, research, voices from the field, critique and commentary, best practices and book reviews. We invite papers that will further understanding of the extent of the impact on children, suggest strategies to combat or mitigate those effects and help government authorities and other stakeholders to come forward to assist children to come out of these traumatic experiences.

To know more about the application process for the second issue, kindly visit our page at www.dcpcr.delhi.gov.in

Delhi Commission For Protection Of Child Rights (DCPCR)
Govt. of NCT of Delhi
5th Floor, ISBT Building, Kashmere Gate,
New Delhi- 110006 Contact- 011-23862685

Nandini Chanana, 17 years