



Rural Learning Through Vernacular AI

Artificial intelligence (AI) is emerging as a key enabler in addressing the learning divide in rural India, particularly by overcoming language barriers. With advances in language AI and platforms such as BHASHINI, educational content can now be delivered in multiple Indian languages, making digital learning more accessible. AI-driven vernacular EdTech solutions offer personalised, voice-enabled, and adaptive learning experiences for students. Supported by digital infrastructure and growing innovation in the EdTech ecosystem, these technologies have the potential to improve learning outcomes and expand access. However, challenges related to infrastructure, data availability, and inclusive design remain critical for scaling these solutions.

Garima Ujjainia

India's aspiration to become a knowledge-driven economy depends significantly on the inclusiveness and accessibility of its education system, particularly in rural India. With nearly 65% of the population residing in rural areas, villages continue to drive agricultural production, support traditional industries, and increasingly contribute to the country's consumption and entrepreneurial growth. In recent years, rural India has also emerged as a critical

frontier for innovation, digital adoption, and economic transformation. As India progresses towards becoming a \$5 trillion economy, the empowerment of rural communities especially through education and digital access will play a decisive role in shaping the vision of *Viksit Bharat 2047*.

Linguistic Diversity and the Rural Learning Divide

India is one of the most linguistically diverse countries in the world. The Constitution recognises 22 scheduled languages, while more than 19,000

dialects and linguistic variations are spoken across the country. For millions of rural learners, the language used in schools or digital learning platforms often does not correspond to their mother tongue, creating a significant barrier to effective learning.

India is also home to one of the world's youngest populations, with a substantial share of its demographic dividend located in rural regions. Unlocking this potential requires more than the expansion of physical infrastructure; it necessitates sustained investment in education, digital literacy, and skill development to enable rural youth and innovators to participate meaningfully in emerging sectors of the economy.

Over the past decade, India has made considerable progress in building the foundations of a digital society. The increasing adoption of smartphones and expansion of internet penetration have enabled greater access to connected schools, digital literacy programmes, and localised educational content for rural learners.

The Government of India's flagship initiative *Digital India* has significantly accelerated connectivity, digital governance, and technology-driven public service delivery. Complementary programmes such as *BharatNet* are extending high-speed internet access to Gram Panchayats, thereby enabling villages to benefit from online education, telemedicine, financial services, and e-governance platforms.

Today, rural India is witnessing rapid growth in smartphone penetration and internet usage, bringing

millions of first-time users into the digital ecosystem and laying the foundation for inclusive digital learning.

Despite significant improvements in connectivity, access alone does not guarantee meaningful participation in the digital economy. According to the *Annual Status of Education Report (ASER)*, learning outcomes in foundational literacy and numeracy remain uneven across many rural regions.

Language barriers further compound this challenge. Although India recognises 22 scheduled languages and thousands of dialects, the medium of instruction for many rural students often differs from the language spoken at home. This linguistic mismatch can limit comprehension, reduce student engagement, and widen the learning divide between rural and urban learners.

Language AI Infrastructure and Rural Digital Inclusion

As India undergoes rapid digital transformation, there is a growing emphasis on designing technological systems that accommodate its linguistic diversity. In a country where millions of citizens interact with governance systems and educational platforms in their native languages, the integration of multilingual artificial intelligence (AI) technologies has become a critical priority.

Recognising this need, the Government of India has launched several initiatives aimed at building a robust language AI ecosystem capable of supporting inclusive digital services across sectors.



BHASHINI, developed under the *Digital India* programme, is a national digital platform designed to overcome language barriers by enabling AI-driven translation, speech recognition, and multilingual digital interaction across Indian languages. The initiative provides open digital infrastructure, datasets, and application programming interfaces (APIs), allowing developers, startups, and public institutions to build language-enabled digital services.

The *BHASHINI* ecosystem currently supports more than 36 languages in text and over 23 languages in voice, with more than 350 AI language models deployed across over 500



digital platforms and applications. The platform is also expanding its coverage to include regional dialects such as Awadhi and Braj, reflecting India's linguistic diversity and enabling deeper local engagement.

These capabilities are particularly significant for rural India, where digital inclusion often depends on whether citizens can interact with technology in their native languages. By enabling real-time speech-to-text and text-to-text translation, language AI platforms facilitate easier access to public services, educational resources, and digital governance systems.

An important step in this direction is the collaboration between the Ministry of Panchayati Raj and *BHASHINI*. Through a Memorandum of Understanding (MoU), multilingual AI tools are being integrated into digital platforms used by Panchayati Raj Institutions. This enables rural citizens and elected representatives to access governance systems in regional languages. This initiative promotes inclusive participation in local governance while improving accessibility to digital services at the grassroots level.

The integration of language AI into rural governance platforms demonstrates how technology can strengthen participatory democracy. When digital systems support local languages, citizens are more likely to engage with governance processes, access information, and

participate in decision-making. Over the long term, such innovations can enhance transparency, accountability, and trust in public institutions.

Beyond governance, the language AI ecosystem also has significant implications for the education sector. AI-driven translation and speech technologies can enable the conversion of educational content into multiple Indian languages, allowing students to access learning resources in their mother tongue. This is particularly relevant for rural learners who may face challenges when using digital platforms that primarily operate in English.

Moreover, collaborative initiatives within the language AI ecosystem are working to expand datasets and improve model accuracy across Indian languages. Efforts such as community-based data contributions, along with partnerships involving academic institutions and technology organisations, are helping to build robust datasets. These datasets, in turn, support the development of AI models capable of understanding regional dialects and conversational contexts.

Such developments are essential for enabling AI-powered EdTech solutions that deliver personalised and voice-enabled learning experiences. When educational applications communicate in familiar languages and dialects, they become significantly more accessible to first-generation digital learners.

Vernacular EdTech and Personalised Learning

The rapid advancement of artificial intelligence (AI) in India is transforming education into a more interactive and learner-centric experience. The integration of AI with vernacular EdTech platforms has enabled adaptive and personalised learning for students. Extending AI-based systems to rural India can help analyse student responses, identify learning gaps, and recommend tailored lessons or exercises based on individual learning patterns.

For instance, a student struggling with mathematics can receive additional explanations and practice exercises in their preferred language. Interactive voice-based tutors can guide students step-by-step through complex concepts, closely replicating the experience of classroom instruction.

Personalised learning tools are particularly important in rural schools, where there is often a shortage of teachers and trained educators. AI-enabled

Adi Vaani

Enabling Inclusive Rural
and Tribal Development Adi
Vaani is an AI-enabled language



platform designed to address communication barriers faced by tribal communities in remote and underserved regions. Under the Adi Karmayogi framework, it facilitates access to governance, education, and healthcare services in native tribal languages. Developed using authentic linguistic data from State Tribal Research Institutes, the platform integrates technological innovation with community knowledge. It incorporates feedback mechanisms to ensure linguistic accuracy, cultural relevance, and continuous improvement. Beyond translation, Adi Vaani supports language preservation, cultural documentation, and digital learning by digitising endangered languages and oral traditions. By enhancing linguistic inclusion in public services and supporting community empowerment, the platform exemplifies the responsible use of AI to advance inclusive rural and tribal development aligned with national priorities.

platforms can complement traditional teaching methods by providing individualised support, thereby allowing teachers to focus on broader classroom management and mentoring roles.

With the development of AI technologies tailored for Indian languages, a new generation of technology startups is emerging. These companies are building language models that capture the nuances of Indian languages and dialects, enabling more effective interaction between digital systems and users.

Sarvam AI, an Indian startup, is developing large language models (LLMs) specifically designed for Indian languages and conversational contexts. Its AI systems aim to support multilingual communication, voice-based interfaces, and contextual understanding across diverse linguistic environments.

By building foundational AI infrastructure for Indian languages, companies such as *Sarvam AI* are enabling developers to create applications that can translate educational content, generate explanations in regional languages, and provide conversational tutoring support. These technologies are particularly valuable for EdTech platforms seeking to expand their reach among rural

and non-English-speaking users.

India's EdTech sector has witnessed significant growth in recent years, driven by increased digital adoption and demand for accessible learning solutions. One of the most notable examples of this growth is *PhysicsWallah*. Initially launched as a YouTube channel offering affordable physics lessons, the platform rapidly evolved into a comprehensive digital learning company providing courses for competitive examinations such as JEE and NEET.

The company's rapid expansion led to its emergence as a unicorn, followed by its public listing marking a significant milestone for India's EdTech ecosystem. This success story demonstrates the transformative potential of digital education platforms in democratising access to high-quality learning resources and reducing regional disparities in educational opportunities.

Challenges in Scaling Vernacular AI in Education

Despite notable progress in artificial intelligence and digital learning platforms, realising the full potential of vernacular AI in rural education in India requires addressing several critical structural gaps. Strengthening rural digital infrastructure remains fundamental, as reliable internet connectivity, consistent electricity supply, and access to affordable devices continue to determine the reach of digital learning in many villages.

Expanding last-mile broadband connectivity through initiatives such as *BharatNet*, along with the creation of shared digital learning spaces in government schools, Panchayat centres, and community libraries, can significantly improve access for rural learners. At the same time, the effectiveness of AI-driven education systems depends heavily on the availability of large and diverse language datasets.

Many Indian languages and dialects remain underrepresented in digital data, limiting the accuracy of speech recognition and translation technologies. National initiatives such as **BHASHINI** are working to expand multilingual AI capabilities; however, deeper collaboration among academic institutions, technology companies, and local communities will be essential to develop robust linguistic datasets.

Equally important is the need for EdTech platforms to adopt inclusive design approaches that reflect the realities of rural users. Voice-enabled interfaces, multilingual content aligned with regional curricula, and

Adi Vaani

Enabling Inclusive Rural
and Tribal Development
Adi Vaani is an AI-enabled language



platform designed to address communication barriers faced by tribal communities in remote and underserved regions. Under the Adi Karmayogi framework, it facilitates access to governance, education, and healthcare services in native tribal languages. Developed using authentic linguistic data from State Tribal Research Institutes, the platform integrates technological innovation with community knowledge. It incorporates feedback mechanisms to ensure linguistic accuracy, cultural relevance, and continuous improvement. Beyond translation, Adi Vaani supports language preservation, cultural documentation, and digital learning by digitising endangered languages and oral traditions. By enhancing linguistic inclusion in public services and supporting community empowerment, the platform exemplifies the responsible use of AI to advance inclusive rural and tribal development aligned with national priorities.

platforms can complement traditional teaching methods by providing individualised support, thereby allowing teachers to focus on broader classroom management and mentoring roles.

With the development of AI technologies tailored for Indian languages, a new generation of technology startups is emerging. These companies are building language models that capture the nuances of Indian languages and dialects, enabling more effective interaction between digital systems and users.

Sarvam AI, an Indian startup, is developing large language models (LLMs) specifically designed for Indian languages and conversational contexts. Its AI systems aim to support multilingual communication, voice-based interfaces, and contextual understanding across diverse linguistic environments.

By building foundational AI infrastructure for Indian languages, companies such as *Sarvam AI* are enabling developers to create applications that can translate educational content, generate explanations in regional languages, and provide conversational tutoring support. These technologies are particularly valuable for EdTech platforms seeking to expand their reach among rural

and non-English-speaking users.

India's EdTech sector has witnessed significant growth in recent years, driven by increased digital adoption and demand for accessible learning solutions. One of the most notable examples of this growth is *PhysicsWallah*. Initially launched as a YouTube channel offering affordable physics lessons, the platform rapidly evolved into a comprehensive digital learning company providing courses for competitive examinations such as JEE and NEET.

The company's rapid expansion led to its emergence as a unicorn, followed by its public listing marking a significant milestone for India's EdTech ecosystem. This success story demonstrates the transformative potential of digital education platforms in democratising access to high-quality learning resources and reducing regional disparities in educational opportunities.

Challenges in Scaling Vernacular AI in Education

Despite notable progress in artificial intelligence and digital learning platforms, realising the full potential of vernacular AI in rural education in India requires addressing several critical structural gaps. Strengthening rural digital infrastructure remains fundamental, as reliable internet connectivity, consistent electricity supply, and access to affordable devices continue to determine the reach of digital learning in many villages.

Expanding last-mile broadband connectivity through initiatives such as *BharatNet*, along with the creation of shared digital learning spaces in government schools, Panchayat centres, and community libraries, can significantly improve access for rural learners. At the same time, the effectiveness of AI-driven education systems depends heavily on the availability of large and diverse language datasets.

Many Indian languages and dialects remain underrepresented in digital data, limiting the accuracy of speech recognition and translation technologies. National initiatives such as **BHASHINI** are working to expand multilingual AI capabilities; however, deeper collaboration among academic institutions, technology companies, and local communities will be essential to develop robust linguistic datasets.

Equally important is the need for EdTech platforms to adopt inclusive design approaches that reflect the realities of rural users. Voice-enabled interfaces, multilingual content aligned with regional curricula, and

applications optimised for low-bandwidth environments can significantly enhance usability for first-generation digital learners.

Addressing these interconnected challenges will be crucial to ensuring that AI-powered vernacular EdTech evolves into a practical and scalable solution for bridging the learning divide in rural India.

Building a Tech-Inclusive India

India stands at a pivotal moment in the evolution of its digital education ecosystem. The convergence of AI technologies, digital public infrastructure, and vernacular EdTech innovation presents unprecedented opportunities to transform rural learning.

To maximise the impact of these advancements, policymakers and stakeholders must prioritise strengthening digital infrastructure, promoting AI research in Indian languages, and supporting

startups that develop inclusive educational solutions. Collaboration among government institutions, academic organisations and the private sector will be critical to scaling innovations that address the needs of rural learners.

At the same time, teacher training and community engagement will remain essential to ensure that technology complements, rather than replaces, traditional educational practices.

As India continues its digital transformation, sustained investment in AI-powered vernacular education systems will be essential for unlocking the full potential of its young population. Bridging the learning divide through such innovations is not merely a technological advancement but a critical step towards building an inclusive, innovative, and future-ready nation. 