**Automatically detects language and then transcribes**

| **Platform** | **Total Languages** | **Points** |
| --- | --- | --- |
| ElevenLabs  <https://elevenlabs.io/speech-to-text> | 99+ languages | * Assamese, Bengali, Gujarati, Hindi, Kannada, Malayalam, Marathi, Nepali, Punjabi, Sindhi, Tamil, Telugu. * Supports real-time streaming for English only. * Not clearly stated whether it supports Automatic Language Detection. |
| AssemblyAI  [https://www.assemblyai.com](https://www.assemblyai.com/) | 99+ languages | * Assamese, Bengali, Gujarati, Hindi, Kannada, Malayalam, Marathi, Nepali, Punjabi, Tamil. * Supports real-time streaming, but Indic support limited. * Supports Automatic Language Detection followed by Transcription. |
| GoogleSpeech-to-Text  <https://cloud.google.com/speech-to-text> | 125+ languages | * Hindi,Gujarati,Kannada,Malayalam,Marathi,Punjabi,Tamil,Telugu * Supports real-time streaming with these Indic languages * Supports Automatic Language Detection followed by Transcription |

| **Feature** | **Bhashini**  https://bhashini.gov.in/ulca/model/explore-models | **Sarvam**  <https://www.sarvam.ai/apis/speech-to-text> | Our Model |
| --- | --- | --- | --- |
| Speech Recognition | 13 languages | 11 languages (Saarika) | 23 languages |
| Input Language Parameter | Requires language input | Requires language input | Doesn’t require language input |
| Core Languages | Hindi, English, Bengali, Tamil, Telugu, Gujarati, Kannada, Malayalam, Marathi, Punjabi, Odia, Urdu, Assamese | Hindi, Bengali, Tamil, Telugu, Gujarati, Kannada, Malayalam, Marathi, Punjabi, Odia, English | Assamese, Bengali, Bodo, Dogri, Gujarati, Hindi, Kannada, Kashmiri, Konkani, Maithili, Malayalam, Manipuri, Marathi, Nepali, Odia, Punjabi, Sanskrit, Santali, Sindhi, Tamil, Telugu, Urdu, English |
| Audio Duration and Processing Time | Lengthy audios are taking more than 4 minutes processing time. (tested with audio of size 3 minutes) | Audio files longer than 30 seconds are not supported; the batch API should be used for longer audios. Lengthy audios cannot be processed using sarvam.ai. | Can process lengthy audios and the processing time is 10-15 seconds. |
| Interface  Functionality | Transcription happens **after recording stops** | Transcription happens **after recording stops** | Supports **real-time streaming transcription** + additional features (Semantic Q&A search). |

| **No.** | **Language** | **Sarvam** | **Our Model** |
| --- | --- | --- | --- |
| 1 | Hindi | प्रश्न जुआड़ में दाना इतना हल्का और कम क्यों होता है? उत्तर खेत में नमी की कमी और उर्वरक कम पड़ने से वजह से यह होता रहता है। आप एक काम कीजिए फूल आते टाइम यूरिया को पानी में घोलकर छिड़कते रहिए। | प्रश्न: झुआर में दाना इतना हल्का और कम क्यों होता है? उत्तर: खेत में नमी की कमी और उर्वरक कम पड़ने से वजह से ये होता रहता है। आप एक काम कीजिए। फूल आते टाइम यूरिया को पानी में घोल कर छिड़कते रहिए। |
| 2 | Punjabi | ਮੀਕਾ ਸਿੰਘ ਨਾਲ ਉਹ ਚੁੰਮਣ ਵਾਦ ਵਿਵਾਦ ਵਿੱਚ ਫਸੀ ਰਹੀ। | ਮੀਕਾ ਸਿੰਘ ਨਾਲੋਂ ਚੁੰਮਣਵਾਦ ਵਿਵਾਦ' ਚ ਫਸੀ ਰਹੀ. |
| 3 | English | With this transition to the big time, the band shortened their name to Stay Below. | With this transition to the big time, the band shortened their name to stay below. |
| 4 | Malayalam | പതിമുഖം. \nപതിമുഖത്തെ സാധാരണമായി സപ്പാൻ മരൻ അല്ലെങ്കിൽ കിഴക്കൻ ഇന്ത്യൻ ചുവന്ന മരം എന്നാണ് വിളിക്കുന്നത്. \nതമിഴ്നാട്, കേരളം, കർണാടക, ആന്ധ്രപ്രദേശ്, പശ്ചിമബംഗാൾ എന്നീ സംസ്ഥാനങ്ങളിൽ കൃഷി ചെയ്യുന്ന ഒരു ഔഷധ വൃക്ഷമാണിത്. | പതിമുഖം പതിമുഖത്തെ സാധാരണമായി സപ്പാൻ മരം അല്ലെങ്കിൽ കിഴക്കൻ ഇന്ത്യൻ ചുവന്ന മരം എന്നാണ് വിളിക്കുന്നത്. തമിഴ്നാട്, കേരളം, കർണാടക, ആന്ധ്രാപ്രദേശ്, പശ്ചിമബംഗാൾ എന്നീ സംസ്ഥാനങ്ങളിൽ കൃഷി ചെയ്യുന്ന ഒരു ഔഷധവൃക്ഷമാണിത്. |
| 5 | Tamil | அற முறுக்கினால் அற்றுப் போகும் | அற முறுக்கினால் அற்று போகும். |
| 6 | Kannada | ಯಾಕಂದ್ರೆ ಅದು ಎಕ್ಸಾಮ್ ನಾವು ಕಂಪ್ಲೀಟ್ ಮಾಡ್ಕೊಂಡ್ರೆ ಮಾತ್ರ ನಾವು ಆ ಸೆಕ್ಷನ್ ಮುಗಿಸಿಕೊಳ್ಳೋಕೆ ಆಗೋದು. ಮತ್ತೆ ವಯಸ್ಸಾಗಿರೋರು ಇರ್ತಾರಲ್ವಾ ಮನೇಲಿ ಅವರು ನ್ಯೂಸ್ ಗಳನ್ನ ನೋಡೋದು, ವಾರ್ತೆಗಳನ್ನ ಮತ್ತೆ ಮತ್ತೆ | ಯಾಕಂದರೆ ಅದು ಎಕ್ಸಾಮ್ ನಾವು ಕಂಪ್ಲೀಟ್ ಮಾಡ್ಕೊಂಡ್ರೆ ಮಾತ್ರ ನಾವು ಆ ಸೆಕ್ಷನ್ನ ಮುಗಿಸ್ಕೊಳೋಕ್ಕಾಗೋದು. ಮತ್ತು ವಯಸ್ಸಾಗಿರೋ ಇರ್ತಾರಲ್ವ ಮನೇಲಿ ಅವರು ನ್ಯೂಸ್ಗಳನ್ನ ನೋಡೋದು, ವಾರ್ತೆಗಳನ್ನ ಅನ್ ಮತ್ತೆ ಮತ್ತು. |
| 7 | Telugu | బ్రతికినన్ని నాళ్ళు ఫలము విచ్చుట కాదు. | బ్రతికినని నాళ్లు ఫలము విచ్చుటగాదు. |
| 8 | Gujarati | ઉપરાંત દંતકથા પ્રમાણે તેમણે તમિલ ભાષાની રચના કરી છે. | ઉપરાંત દંતકથા પ્રમાણે તેમણે તમિલ ભાષાની રચના કરી છે. |
| 9 | Marathi | उंदीर खेळताना आनंदाने हसण्याचा आवाज काढतात. | उंदीर खेळताना आनंदाने हसण्याचा आवाज काढतात। |
| 10 | Odia | ସେ ଏକୋଇଶିରୁ ଅଧିକ ବର୍ଷ ଧରି ଏ ଦାୟିତ୍ୱ ନିର୍ବାହ କରିଥିଲେ। | ସେ ଏକୋଇଶରୁ ଅଧିକ ବର୍ଷ ଧରି ଏ ଦାୟିତ୍ୱ ନିର୍ବାହ କରିଥିଲେ। |
| 11 | Bengali | পদার্থবিজ্ঞানের নিয়মগুলো অন্তরজ সমীকরণ হিসেবে লেখা হয়। | পদার্থবিজ্ঞানের নিয়মগুলো অন্তরজ সমীকরণ হিসেবে লেখা হয়। |
| 12 | Assamese | Not Supported | সেই শাৰীকেইটাৰ আকাৰৰ বস্তু সিমানখিনি দূৰত্বৰ পৰা দেখিবলৈ বন্ধ কৰি পেলালে। |
| 13 | Bodo | Not Supported |  |
| 14 | Dogri | Not Supported | मुड़े दे कुड़ी दे टिबड़े मलोंदे ते जि'यां टिबड़ा उं'दा मिलदा फिर ओह्दे बाद कुड़ी मुड़ा करदे मलाटी |
| 15 | Kashmiri | Not Supported | تُہۍ کَتہِ پٮ۪ٹھ چھِو |
| 16 | Konkani | Not Supported | बरें बरें म्हळ्यार दुदाचें तुमचें जातलें दूद म्हशींचें जाल्यार सादारण साठ रुपया लिटर आसा |
| 17 | Maithili | Not Supported | बहुत ही नीक लागैए बहुत अच्छा लागैए खाना बनाबैमे खाना पीना बनेलहुँ अपना बाल बच्चा हमर सब हम अगर सोचबै तऽ हमर बाल बच्चा हमरा जरूर ओ करऽ देतै कि हँ तू माँ कर नीक हेतौ बाल बच्चा सब सब साथ देतै |
| 18 | Manipuri | Not Supported | ꯌꯨ ꯑꯦ ꯑꯦꯟ ꯅꯝꯕꯔ ꯃꯔꯤ ꯃꯔꯤ ꯇꯔꯦꯠ ꯑꯃ ꯃꯥꯄꯜ ꯃꯉꯥ ꯃꯉꯥ ꯇꯔꯦꯠ ꯇꯔꯦꯠ ꯅꯤꯄꯥꯜ ꯃꯥꯄꯜ ꯅꯤꯄꯥꯜ ꯁꯤꯖꯤꯟꯅꯗꯨꯅ ꯑꯩꯒꯤ ꯄ꯭ꯔꯣꯕꯤꯗꯦꯟ ꯐꯟ ꯑꯦꯀꯥꯎꯟꯀꯤ ꯏ ꯄꯤ ꯑꯦꯐ ꯑꯣ ꯄꯥꯁꯕꯨꯛ ꯕꯦꯂꯦꯟꯁ ꯆꯦꯛ ꯇꯧꯕ ꯄꯥꯝꯃꯤ |
| 19 | Nepali | Not Supported | चोकटा खान गएको बुढी जोरमा डुबेर मरि। |
| 20 | Sanskrit | Not Supported | म प्रशासनेन केन्द्रीय कर्मकरेभ्य सेवोत्तरवेतनभोगिभ्यश्च जनेभ्यो महार्घताभृतौ प्रतिशतं पचाशत्मिता वृद्धि समुद्घोषित |
| 21 | Santali | Not Supported | ᱟᱨ ᱥᱮᱨᱢᱟ ᱨᱮᱭᱟᱝ ᱡᱚᱛᱚ ᱠᱷᱚᱱ ᱢᱟᱨᱟᱝ ᱯᱚᱨᱚᱵᱽ ᱫᱚ ᱥᱚᱨᱦᱟᱭ ᱥᱮ ᱵᱟᱦᱟ |
| 22 | Sindhi | Not Supported | ब्लेड जे मयार जी, साराह छो न कई। |
| 23 | Urdu | Not Supported | برسات کے موسم میں بوندوں کا افسوس کرنا  خوبصورتی ہوتا ہے۔ |