

# Automated Assessment of Noisy Crowdsourced Free-text Answers for Hindi in Low Resource Setting

Dolly Agarwal<sup>1</sup>, Somya Gupta<sup>2</sup>, Nishant Baghel<sup>1</sup>

<sup>1</sup>{dolly.agarwal, nishant.baghel}@pratham.org, <sup>2</sup>somya.gupta1@gmail.com



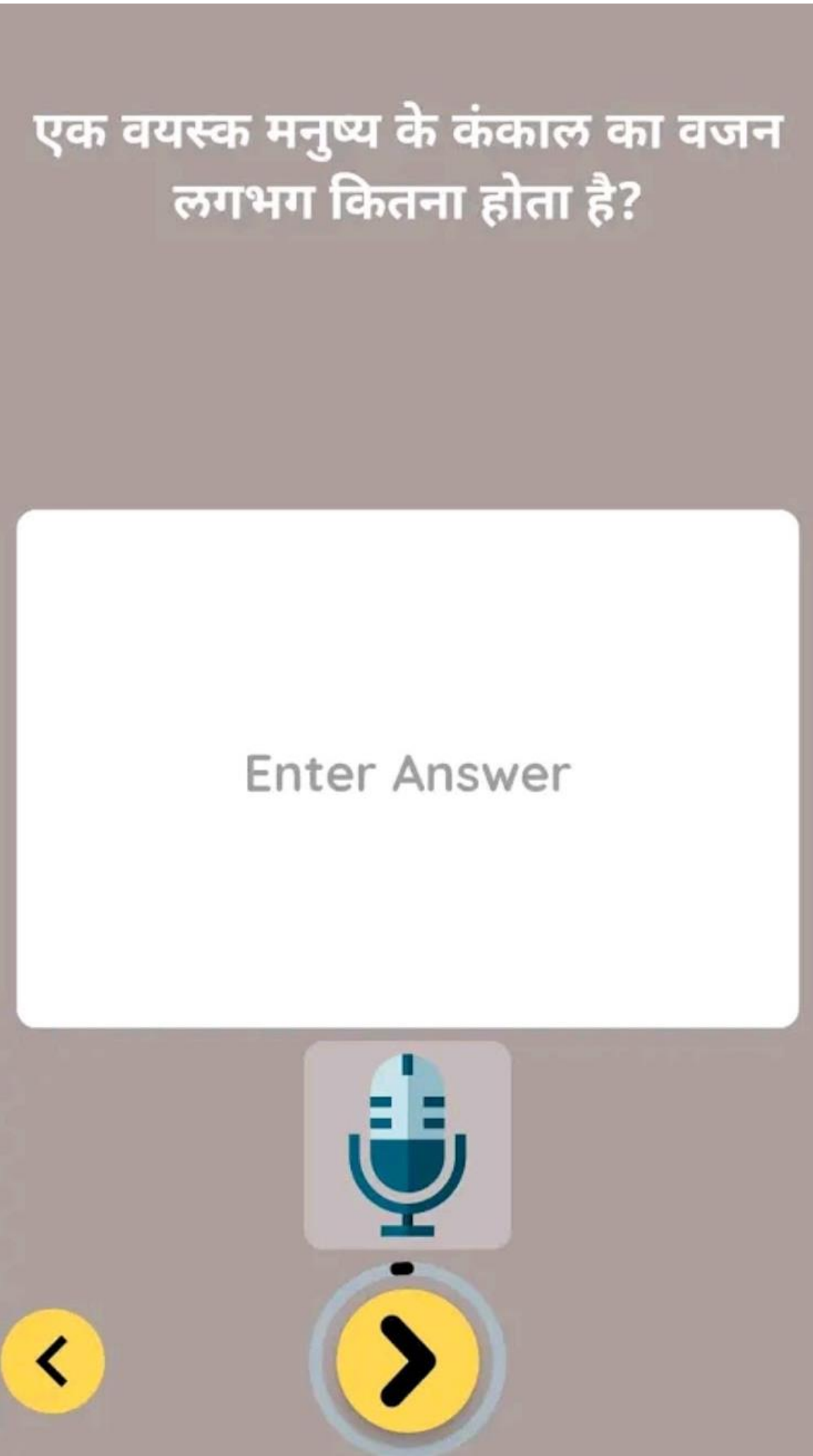
Pratham

Every Child in School and Learning Well

## Introduction

Posing and assessing **open-ended descriptive questions** to children is crucial to evaluating their learning levels and understanding of concepts.

Pratham has **deployed a speech enabled android assessment app** to assist children in their learning and crowdsource answers to free-text questions in Hindi. This program reaches more than 1000 villages and over 109,560 children



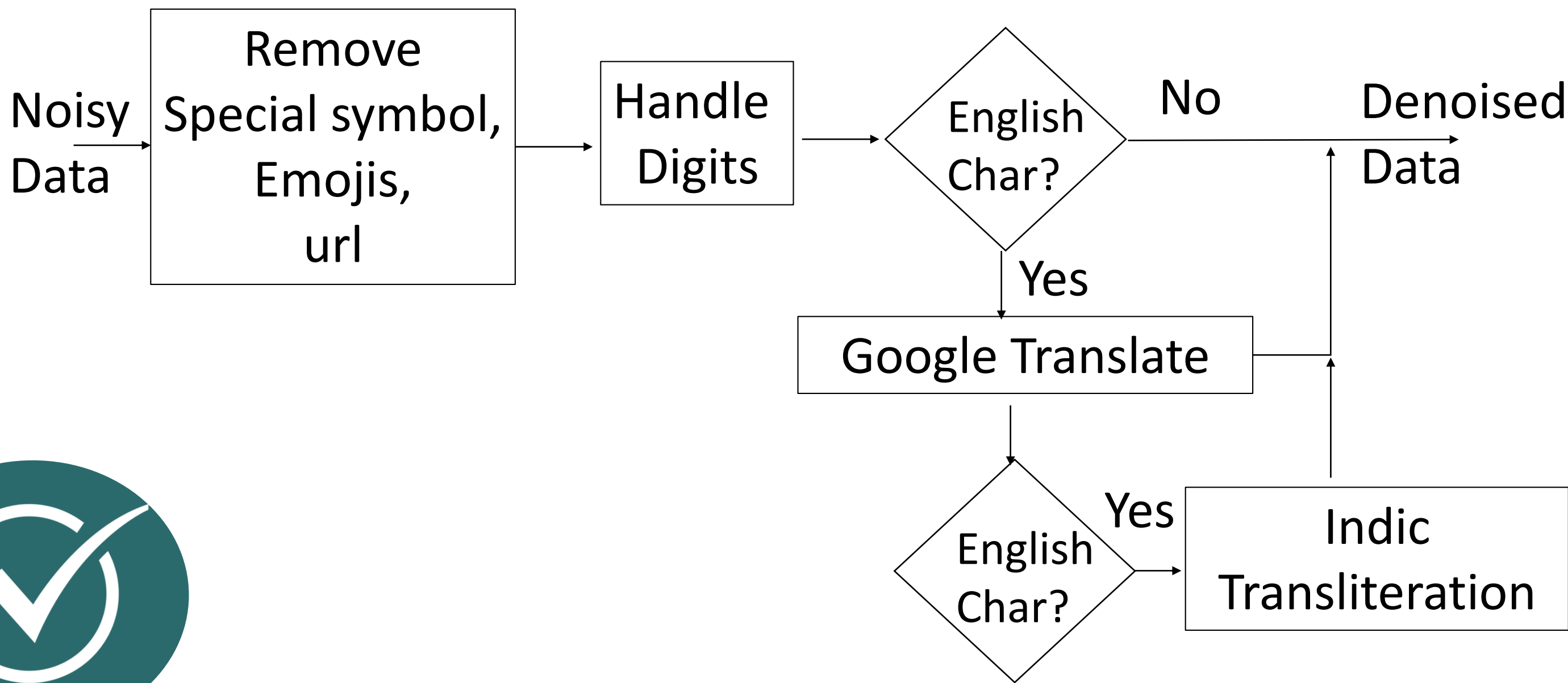
## Noise Handling

Noise Type	Example of Noisy Data	How did we process it	Denoised Text
Special Characters	10,kg \£\%£`i\$°@\$¶¶`=	Removed special character	10 kg
Emoji	😊रक्त का थक्का नहीं जमेगा	Removed emojis	रक्त का थक्का नहीं जमेगा
Translated Text	BONES	Used Google translate	हड्डियों
Transliterated Text	<i>haddi</i>	Used Indic Transliterate	हड्डी
Digit	4561	Converted to word lexemes	चार हज़ार पांच सौ इकसठ
URL	हड्डियोंसेhttps://faq.whatsapp.com/general/26000015	Removed the URL	हड्डियों से
Speech recognition errors	उड जा Reference ans: 'ऊजा'	Not processed here	

## Challenges

- Multiple versions of correct answer to the same question
- Low resource language: Hindi
- Crowdsourcing results in Noisy Data
- Low Resource Setting

## Preprocessing Pipeline



## Results

Method	ROC AUC on Noisy Data	ROC AUC on Denoised
Baseline	0.50	0.50
Jaccard	0.69	0.85
indicNLP	0.79	0.86
fastText	0.78	0.85

## Data and Attributes



8-14 years



35 science questions



39641 answers



15479 evaluated

Question (Hindi/ETL)	Answers(Original/ETL)
आंखों में धूल चली जाने पर आंसू क्यों आ जाते हैं? / Why do we tear up when dirt enters our eyes?	aansu nikalne se dhul ke kan bahar aa jate hai / dust particles get released due to tears
	आंखों से धूल बाहर निकलने के लिए / to remove dirt from eyes

## Methodology

- Ground truth using Human Evaluators
- Baseline: Mark based on coin toss
- Jaccard Similarity**  $J = \frac{(C \cap I)}{(C \cup I)}$
- We model the assessment of user answers against ideal answers as a similarity task and use word embedding based similarity measures to capture semantic similarity among user and ideal answers
  - indicNLP**
  - fastText**

## Conclusion & Future Work

- Denoising helps improve Jaccard Similarity performance, presenting a strong case for feasibility of automated assessment in low resource settings
- Embeddings for words like ऑक्सिजन (ETL: Oxygen), भारहीनता (ETL: Weightlessness) etc. are absent from fastText embeddings
- With Pratham's reach into 22 states and up to **15 million children** in India, we plan to include other regional languages

