REVA UNIVERSITY

CAPSTONE-BA05

Speech Analysis For Teaching A Vernacular Language

Mentor

Dr. J.B. Simha

Chief Mentor, RACE

CTO, ABIBA Systems

Participant

Taiba Naz

R19DM006

Topics Covered

- ABSTRACT
- PROBLEM STATEMENT
- OBJECTIVE
- LITERATURE SURVEY
- APPROACH(CRISP-DM)
- BUSINESS UNDERSTANDING
- DATA UNDERSTANDING
- DATA PREPARATION
- MODEL BUILDING
- EVALUATION
- DEPLOYMENT
- RESULT AND CONCLUSION
- FUTURE WORK
- REFERENCES
- Q &A

Abstract

- Mobile/Web applications being used to learn a **foreign language or to learn a second language** has attracted many users**.**
- The growth of the global market for the educational app has increased from $3.4 billion (in 2011) to $37.8 billion (2020). *
- Duolingo have expanded over 300 million registered users across the world, launched on 30th November 2011.**
- Indian languages like Hindi,Urdu,Kannada,Tamil and Malayalam are not in focus.
- Urdu learning app which **correct the pronunciation for the words, evaluates the kids performance** and capable of giving **immediate feedback** .
- This application focuses on the main aspects of language learning skills i.e. **Listening, Speaking and Reading.**
- Implementation of Data Mining and **Natural Language Processing (NLP)** techniques like text to speech conversion, Speech Recognition etc.

*https://www.statista.com/statistics/273960/global-mobile-education-market-volume/

**https://en.wikipedia.org/wiki/Duolingo

Problem Statement

Loss or Decline of Vernacular and Indigenous Languages

This is how a Class 4 student reads a Class 2 text in his mother tongue

> *A large proportion of students currently in elementary school estimated to be over 5 crores – have not attained foundational literacy and numeracy*

- Draft National Education Policy 2020

Percentage of students who failed to read even one word correctly

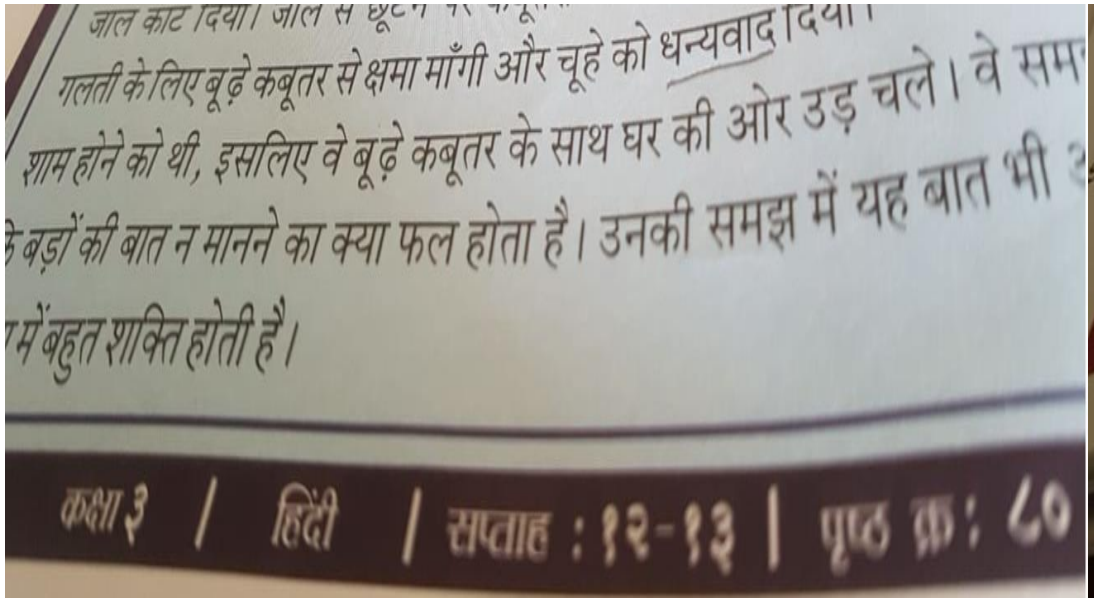
STATE	LANGUAGE	(IN %)
Uttar Pradesh	Hindi	75.9
Rajasthan	Hindi	62.5
Karnataka	Kannada	53.2
Odisha	Oriya	38.8
Uttarakhand	Hindi	30.1
Chhattisgarh	Hindi	23.2
Maharashtra	Marathi	3.8

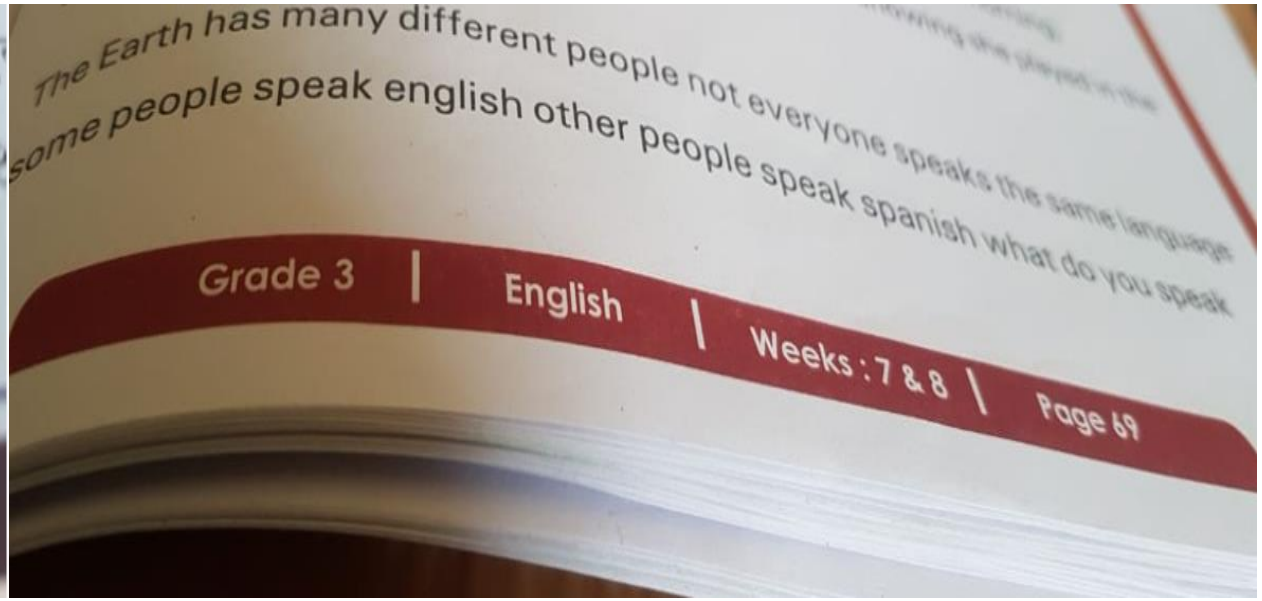
Source: 2018 USAID study conducted in seven states

> *At a potential 100 million children getting through school without basic learning, we cannot afford to underestimate the problem. It has to be a reshaping in the classroom, community and home*

- Ashok Kamath, chairperson, Akshara Foundation

Problem Statement Cont…





Hindi-15 words(2-5 letters)

English-24 words(4-9 letters)

Objectives

Design an application to help a 5 year old to learn Urdu pronunciation by focusing on the important characteristics of the language, providing equal focus to all parts of learning: Listening, speaking & reading .

The application evaluate kids pronunciation for the word in grade1 text book.

Literature Survey

REVA UNIVERSITY

AUTHORS,PAPER PUBLISHED,YEAR OF PUBLICATION	RESEARCH WORK
Catherine Regina Heil,"*A review of mobile language learning applications: trends, challenges and opportunities* " 2016.	3 major trends: apps tend to teach Vocabulary in *isolated units* rather than in relevant contexts minimally adapt *to suit the skill sets of individual learners,* and lack of explanatory *corrective feedback to learners*.
Asadullah," *Automatic Urdu Speech Recognition Using Hidden Markov Model*" 2016	An approach *for automatic speech recognition of Urdu isolated words*. Findings: *Urdu phonetics and phonology differs widely* from English language *Overall recognition accuracy of 78.2%* for 100 word with the use of deep learning models .
J. Islam, *"A speech recognition system for Bengali language using recurrent Neural network," 2019*	A Speech Recognition System for *Bengali Language using Recurrent Neural Network* , further *improved accuracy by using the CTC loss function and language model. Language* consisting of *diacritic characters are very much difficult to train in a model*

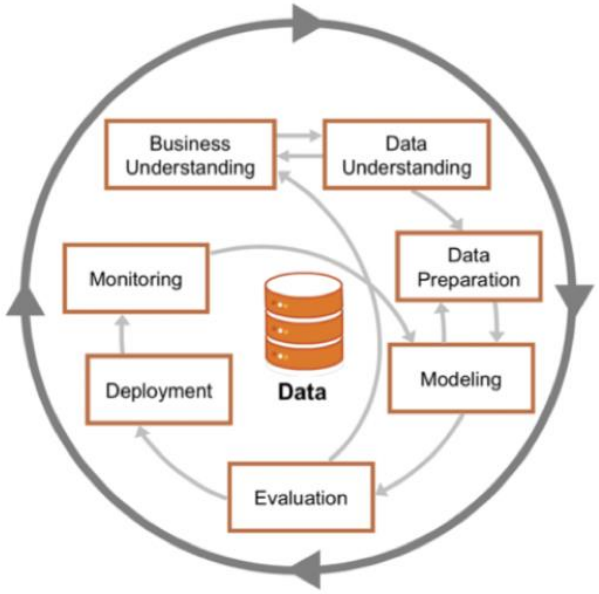
Literature Survey Cont…

REVA UNIVERSITY

AUTHORS,PAPER PUBLISHED,YEAR OF PUBLICATION	RESEARCH WORK
R. Hincks, "*Speech recognition for language teaching and evaluating: A study of existing commercial products*," *2002*	▪ Tool to practice *spoken language outside the classroom*. ▪ Software helped *students to improve fluency and confidence*. ▪ It provide *individual feedback on pronunciation*, ▪ that is often *lacking in the language classroom*. ▪ Algorithms calculate by how much a given *pronunciation has deviated* from a model, and give a *score on phonetic accuracy*.
https://www.duolingo.com/courses *Read Along (Bolo): Learn to Read with Google* *https://www.quora.com/Are-there-any-good-Urdu-educational-apps-for-kids-to-learn-Urdu*	▪ Popular language-learning *app Duolingo ,no support for Urdu language*. ▪ Google Bolo *Application focuses on a child's reading skills only.* ▪ Learn Urdu Quickly, Learn Urdu Free, and Learn Urdu kids

Approach

CRISP-DM (Cross-Industry Standard Process for Data Mining)



- ✓ **Business Understanding:** App requirement
- ✓ **Data Understanding:** Urdu words from NCERT grade1 book
- ✓ **Data Preparation:** Normalization of the data
- ✓ **Modeling:** Speech Recognition model, Sequence matcher
- ✓ **Evaluation:** Evaluated it with a kid.
- ✓ **Deployment:** Deploying model in Django framework.

Business understanding

- Shortage of teachers causing poor learning.
- One to one interactions is costly affair.
- Teacher Fatigue creates biasness and stress.
- Not a single app available in the market which correct the pronunciation for Urdu language.

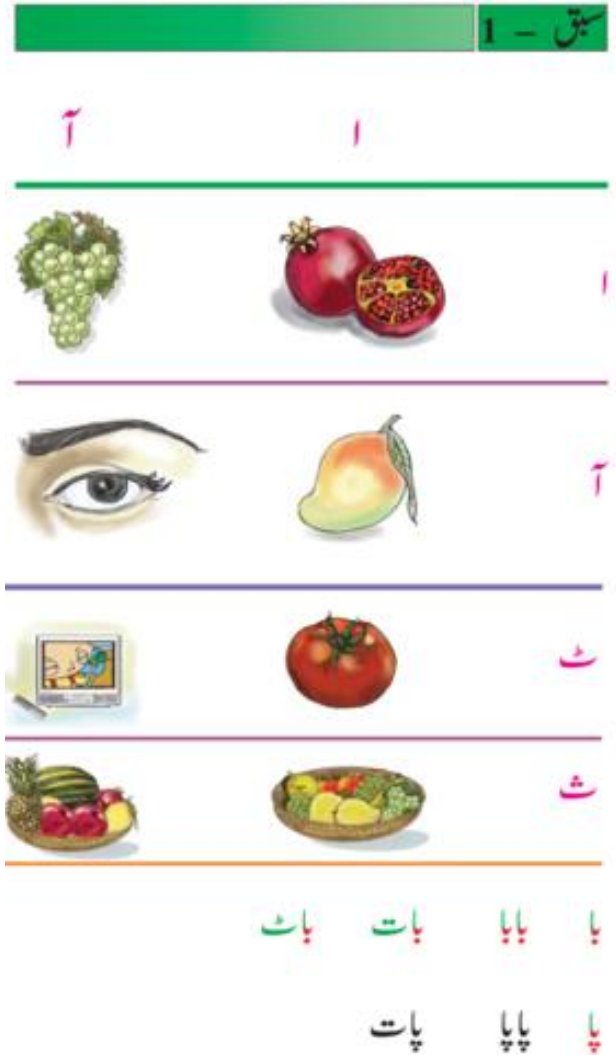
There is *a* ***need to develop an application (A virtual teacher)to assist teachers*** and these kids to ***improve quality of education.***

State of Karnataka Government Schools

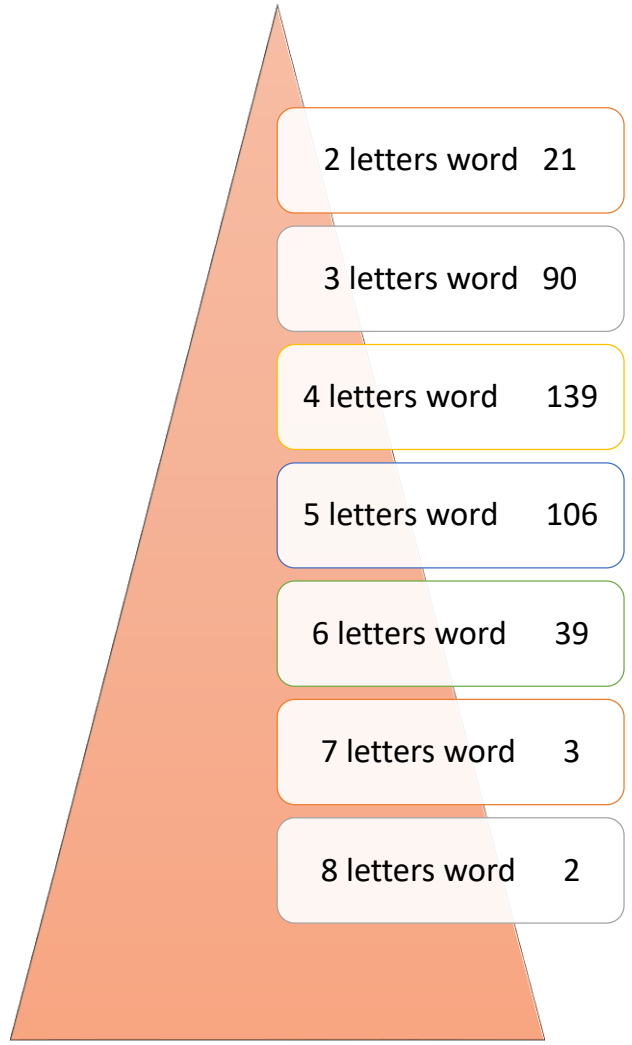
Urdu schools in a pathetic state
bangaloremirror.indiatimes.com

Government Urdu School, Yeshwant...

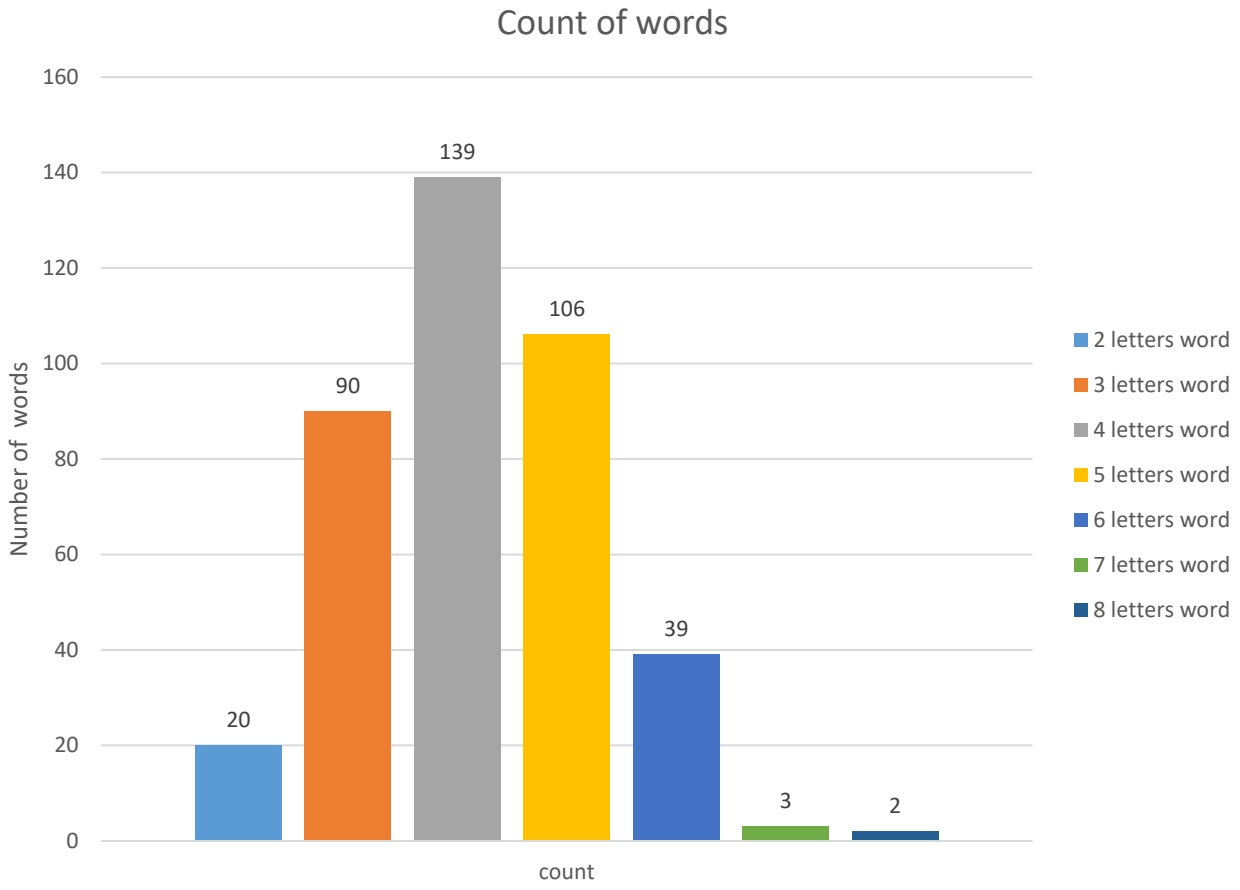
Data Understanding

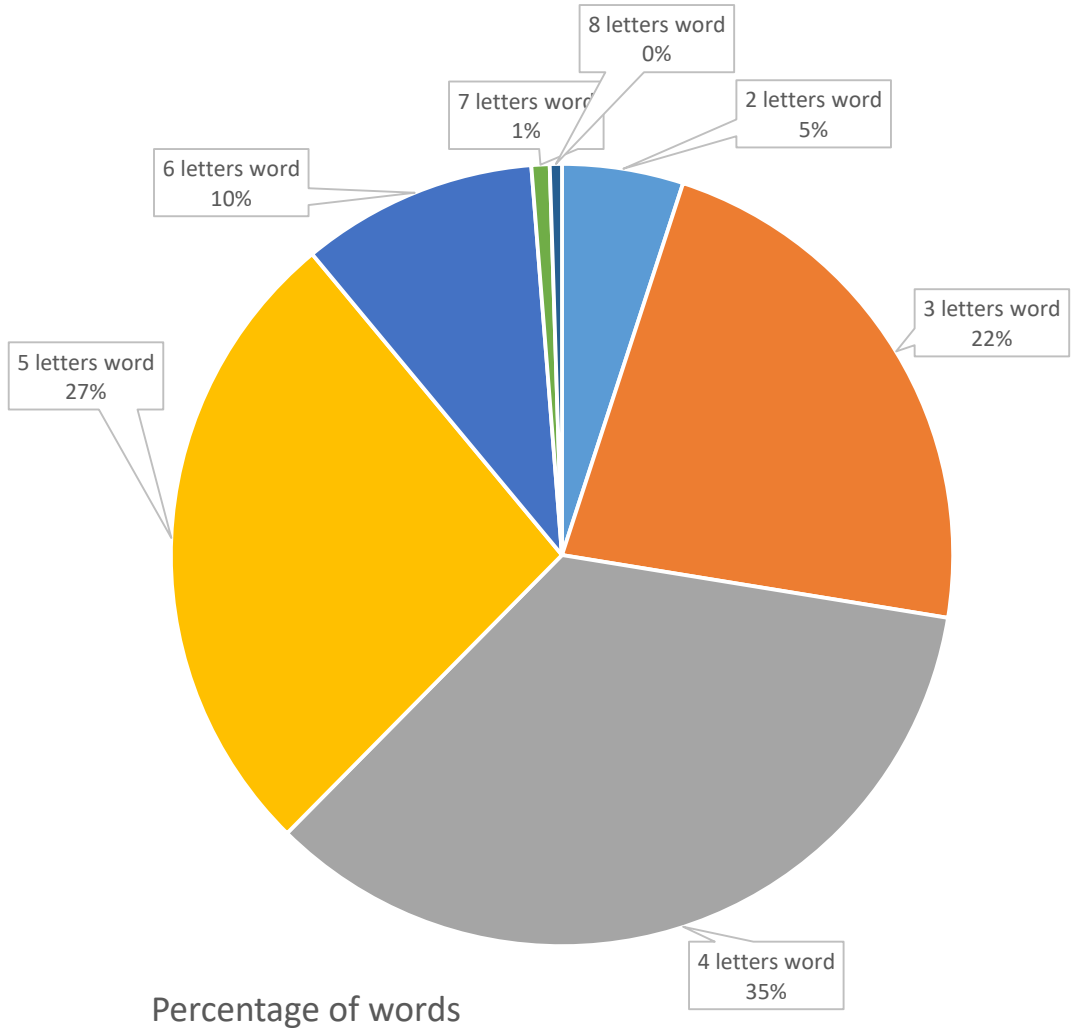


Sno.	Urdu Word
1	انار
2	انگور
3	کیلا
4	سیب
5	آپ
6	اردو
7	بس
8	دس
9	نرس
10	ابر
11	درس
12	پرس
13	تخت
14	بازار
15	سبق

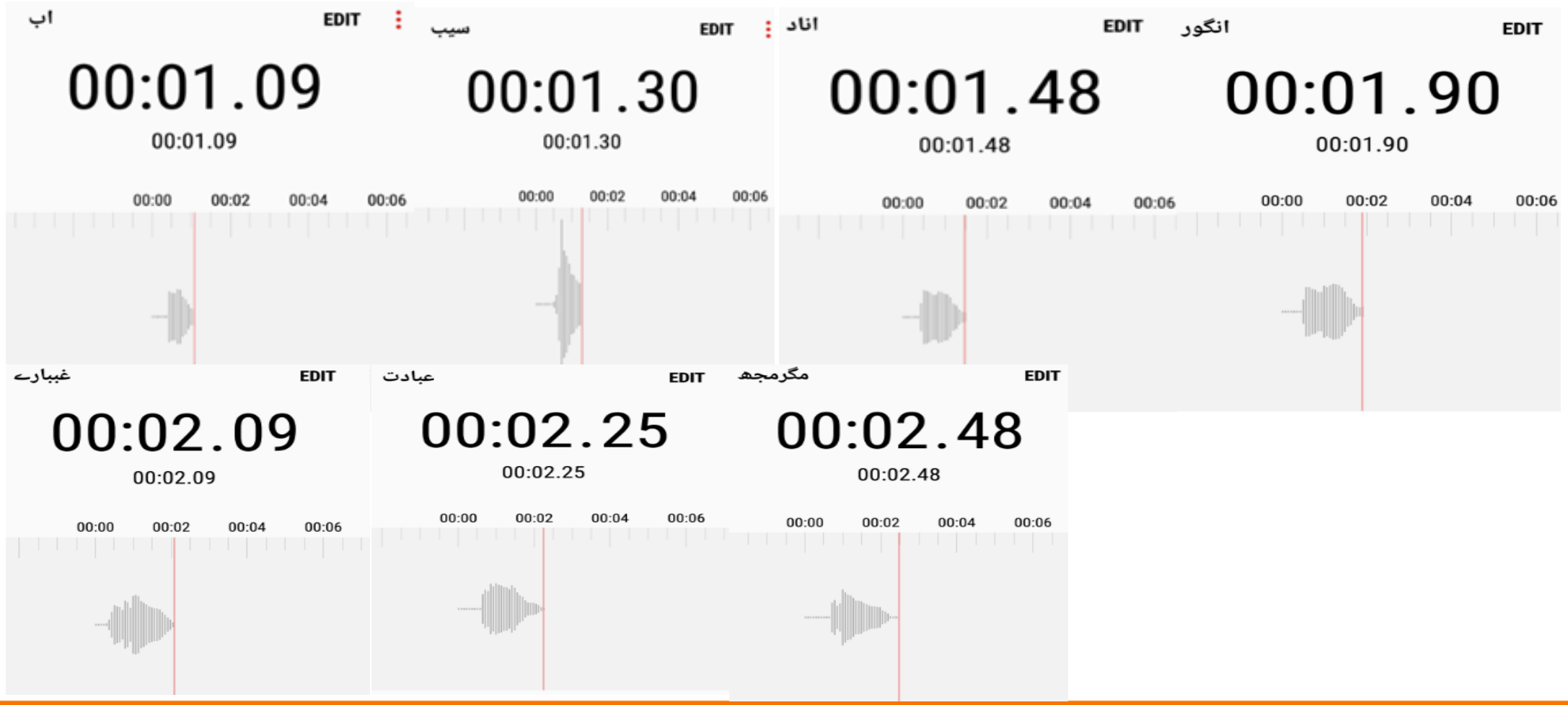


Data Understanding





Data Understanding continue...



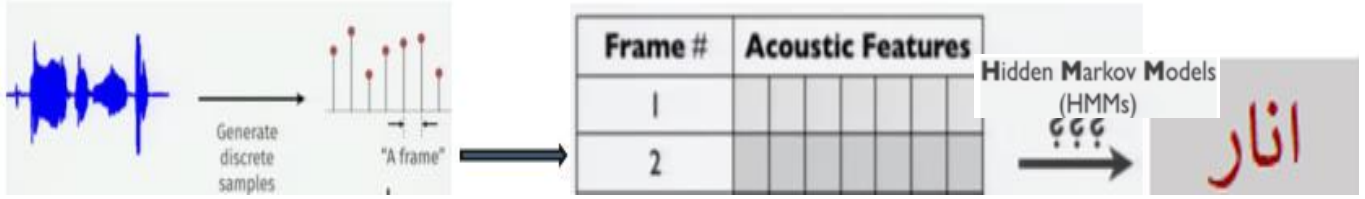
Data Preparation

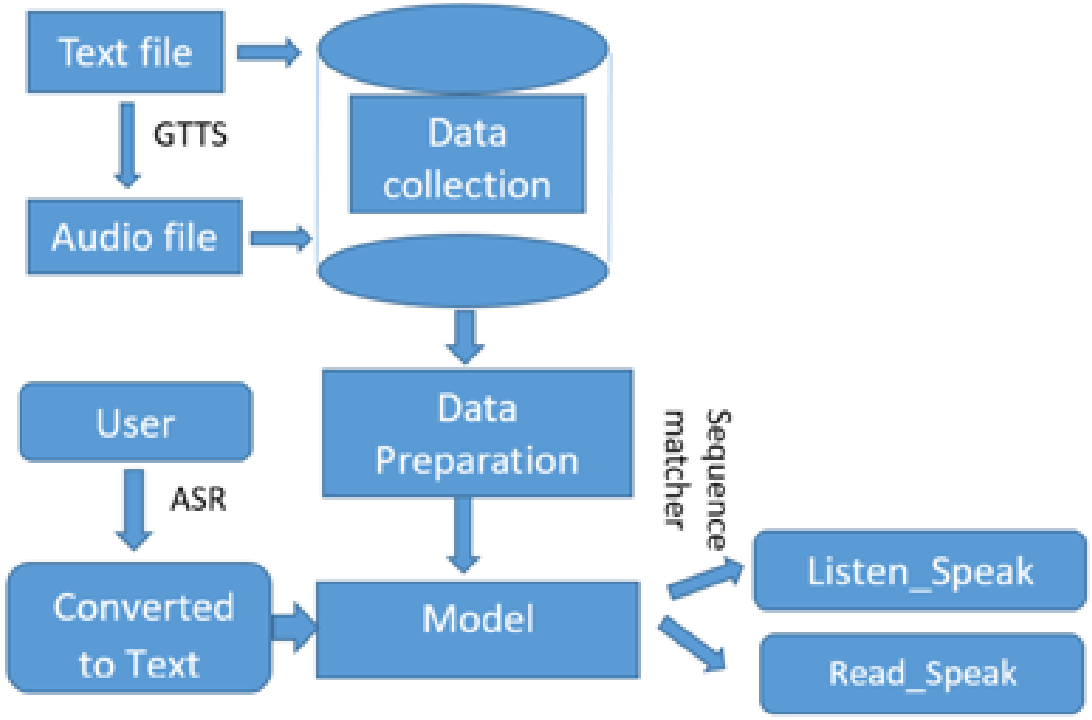
- Dataset has been created with reference to NCERT text book for grade1.
- 7-8 letters words have been removed to avoid confusion.
- Words are arranged according to its difficulty/complexity level.
- Some Conversion of text to audio using GTTS has not worked well ,so creation of those audio files done manually.
- Images has been collected with text book and internet to make the model more interactive.

id	Urdu Word	Audio_files	Image_files
1	انار	audio/1_1.mp3	image/1_1.jpg
2	انگور	audio/1_2.mp3	image/1_2.jpg
3	کیلا	audio/1_3.mp3	image/1_3.jpg
4	سیب	audio/1_4.mp3	image/1_4.jpg
5	آپ	audio/1_5.mp3	image/1_5.jpg
6	اردو	audio/1_6.mp3	image/1_6.jpg
7	بس	audio/1_7.mp3	image/1_7.jpg
8	دس	audio/1_8.mp3	image/1_8.jpg
9	نرس	audio/1_9.mp3	image/1_9.jpg
10	ابر	audio/1_10.mp3	image/1_10.jpg
11	درس	audio/1_11.mp3	image/1_11.jpg
12	پرس	audio/1_12.mp3	image/1_12.jpg
13	تخت	audio/1_13.mp3	image/1_13.jpg
14	بازار	audio/1_14.mp3	image/1_14.jpg
15	سبق	audio/1_15.mp3	image/1_15.jpg
16	جِس	audio/1_16.mp3	image/1_16.jpg
17	جوش	audio/1_17.mp3	image/1_17.jpg
18	شاباش	audio/1_18.mp3	image/1_18.jpg
19	شیشی	audio/1_19.mp3	image/1_19.jpg
20	آسان	audio/1_20.mp3	image/1_20

Model Building

GTTS: Google Text to Speech converter
ASR: Automatic Speech Recognition





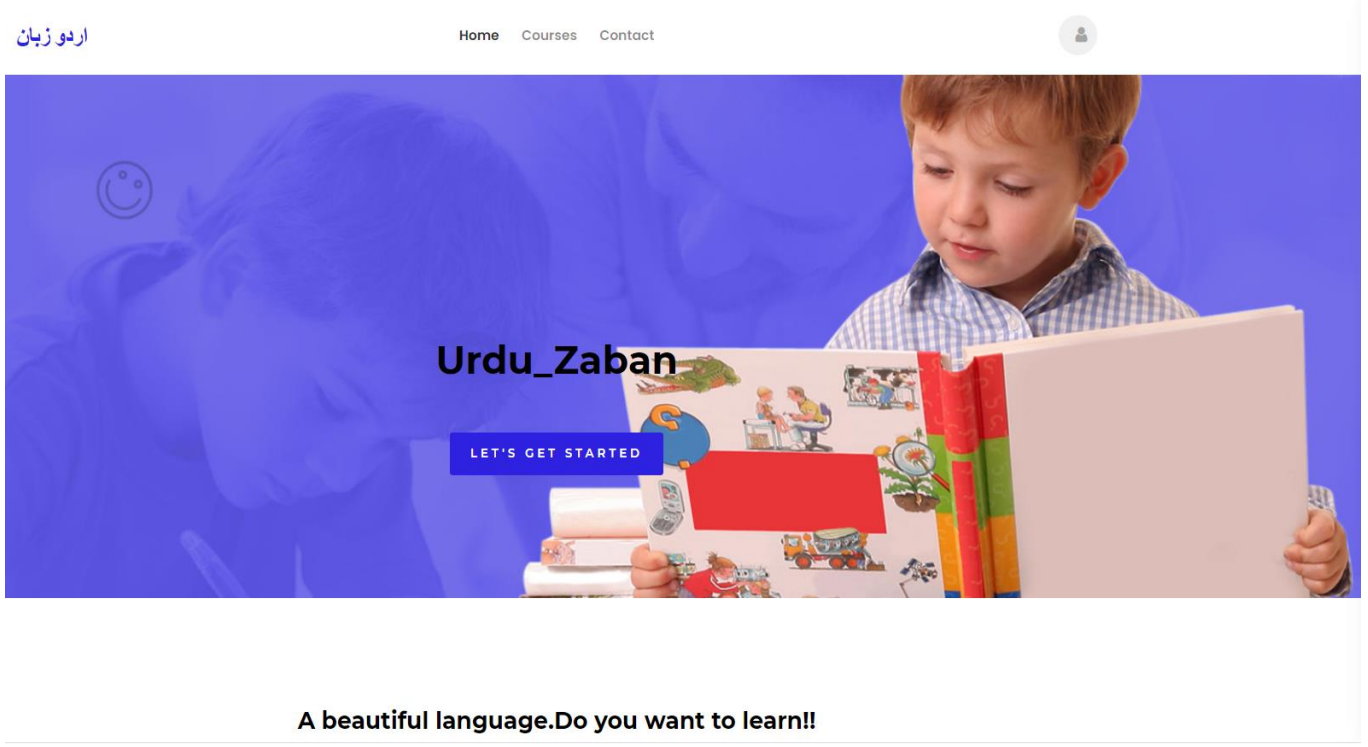
Model Evaluation

- The Error rate of the Listen_Speak model is 8.13% and the Read_Speak model is 8.75% with internal evaluation .

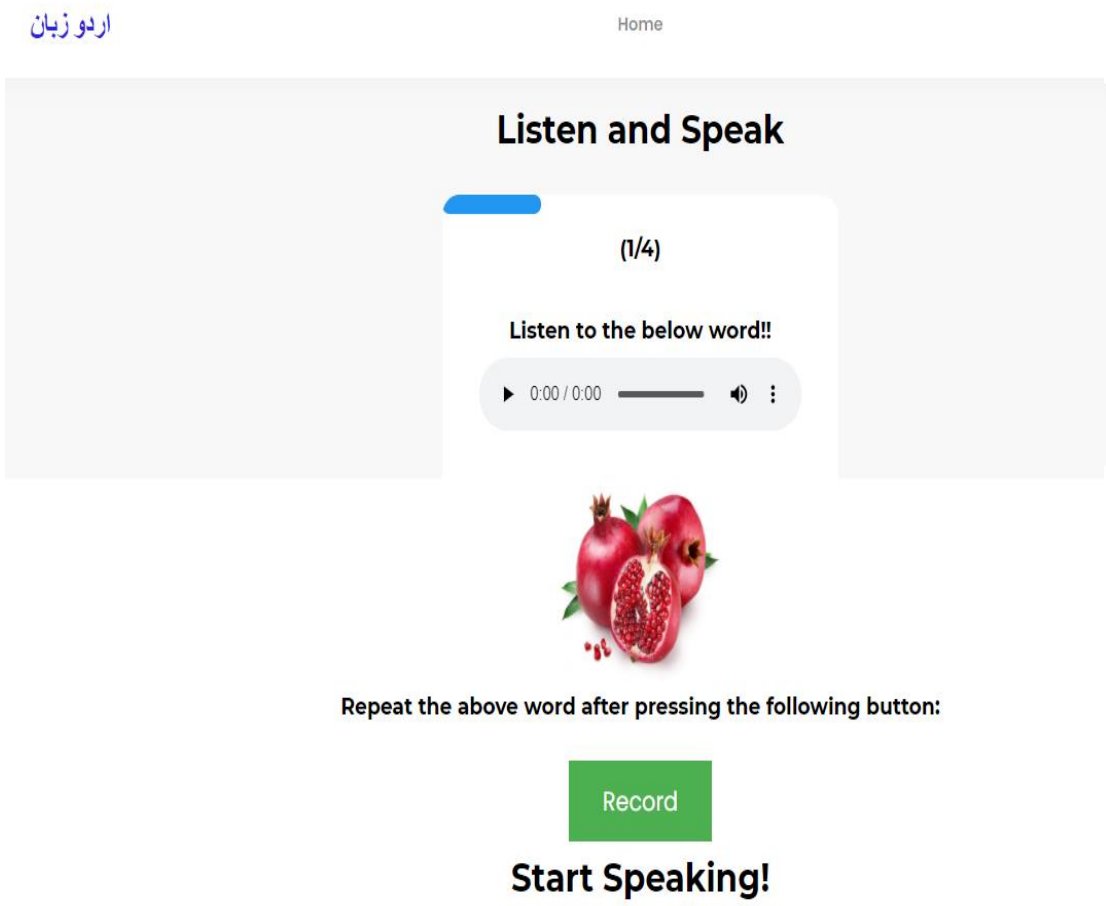
- External evaluation will be done with some government school kids.

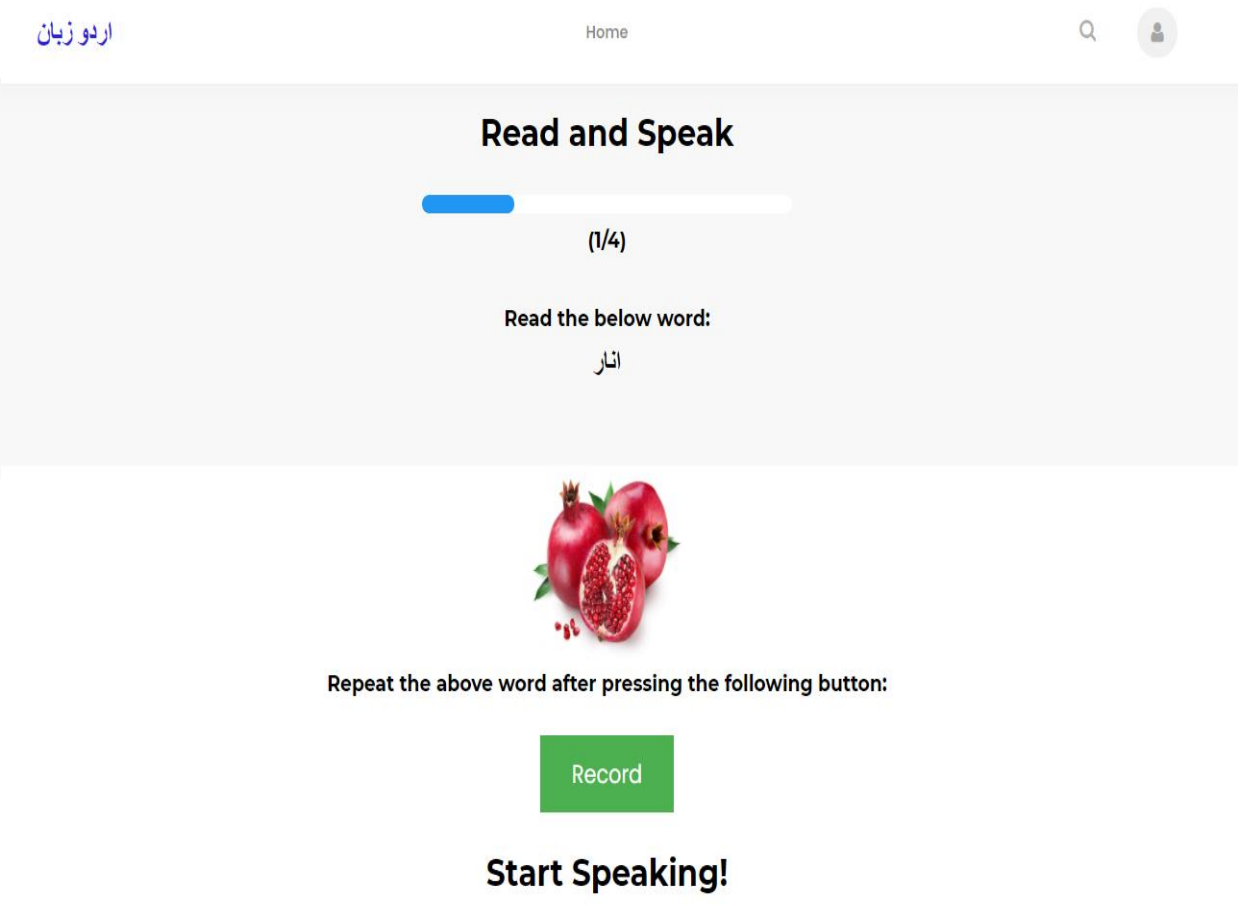
Model Type	Error Rate
Listen and Speak	8.13%
Read and Speak	8.75%

Model Deployment

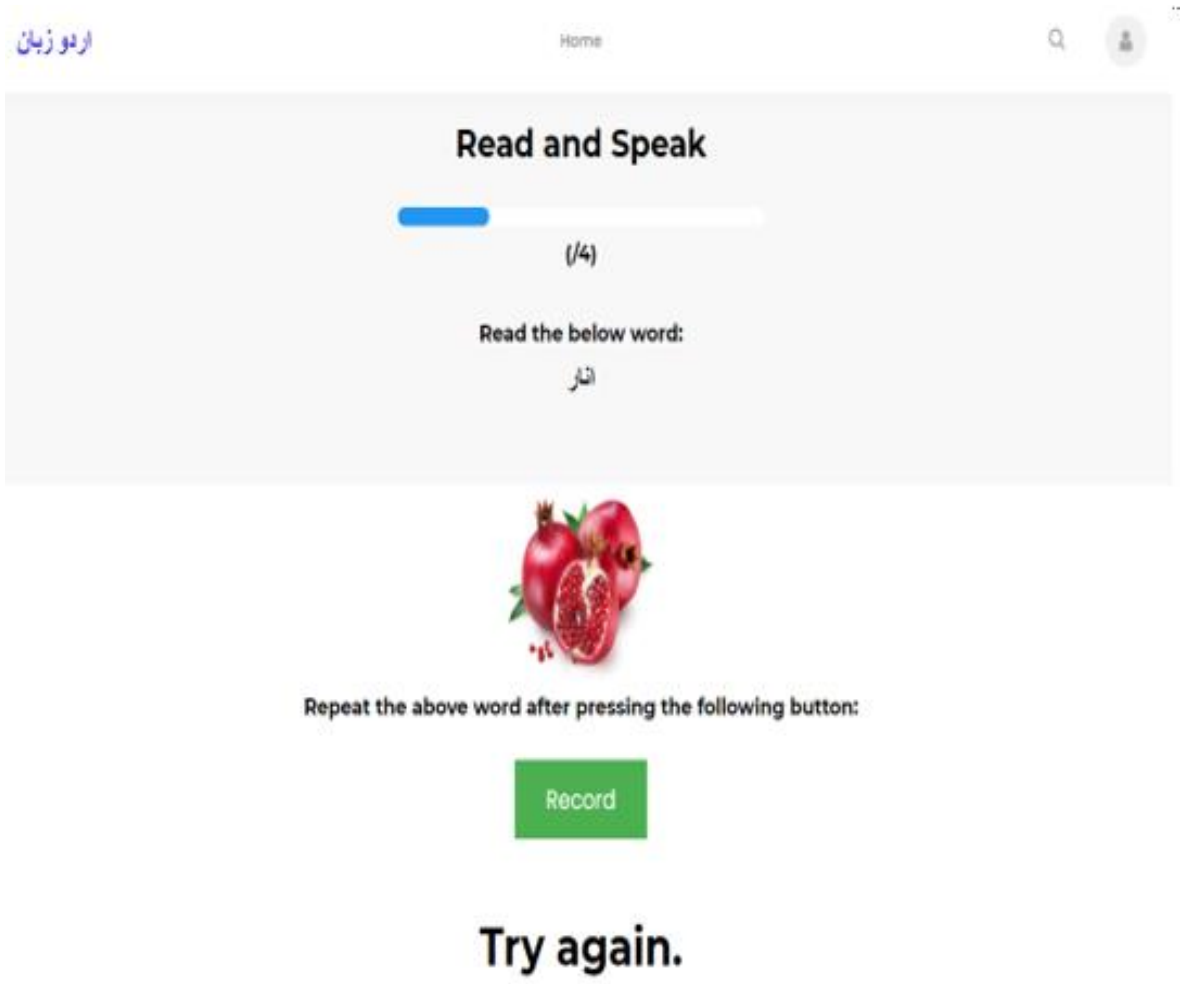


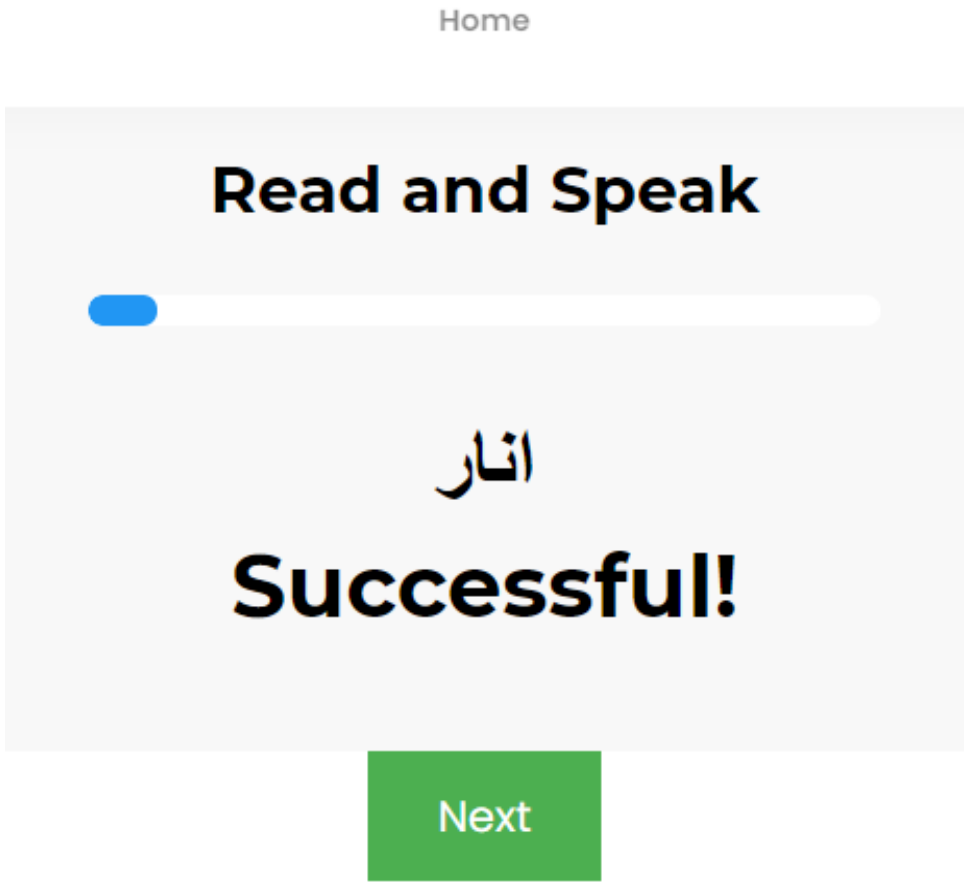
Model Deployment



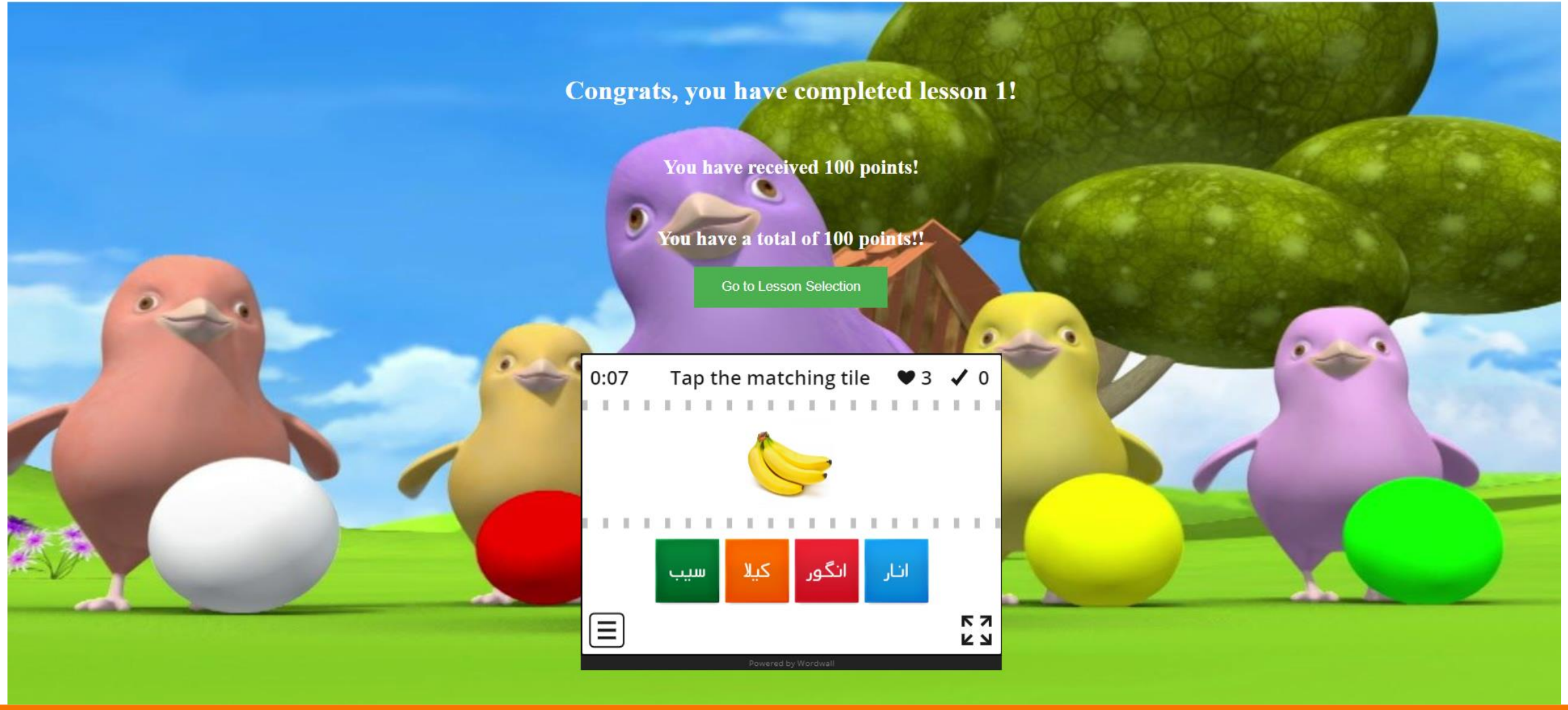


Model Performance





Puzzles in the app



Leaderboard in the Puzzle

Webapp Evaluation

Accuracy of Listen_Speak model 96.25% and Read Speak model are same 95.63% (calculated with a sample of 160 words).

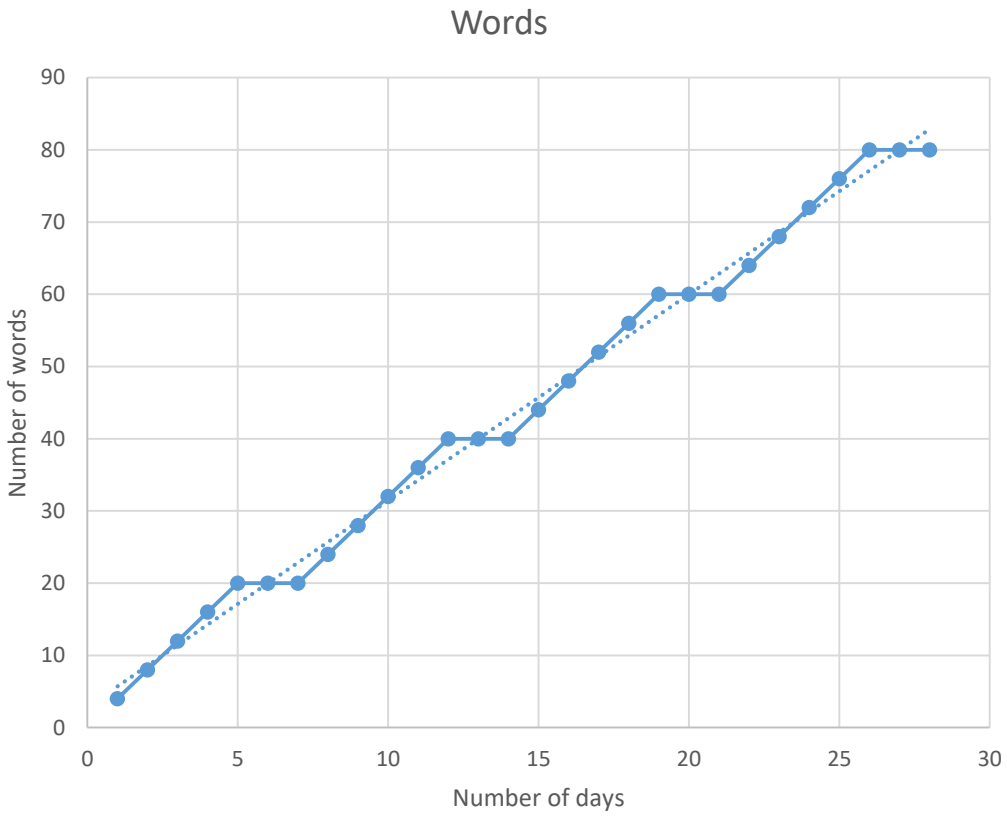
Total=160	**Predicted: Right**	**Predicted: wrong**	
Actual: Right	TP=146	FN=6	152
Actual: wrong	FP=0	TN=8	8
	146	14	

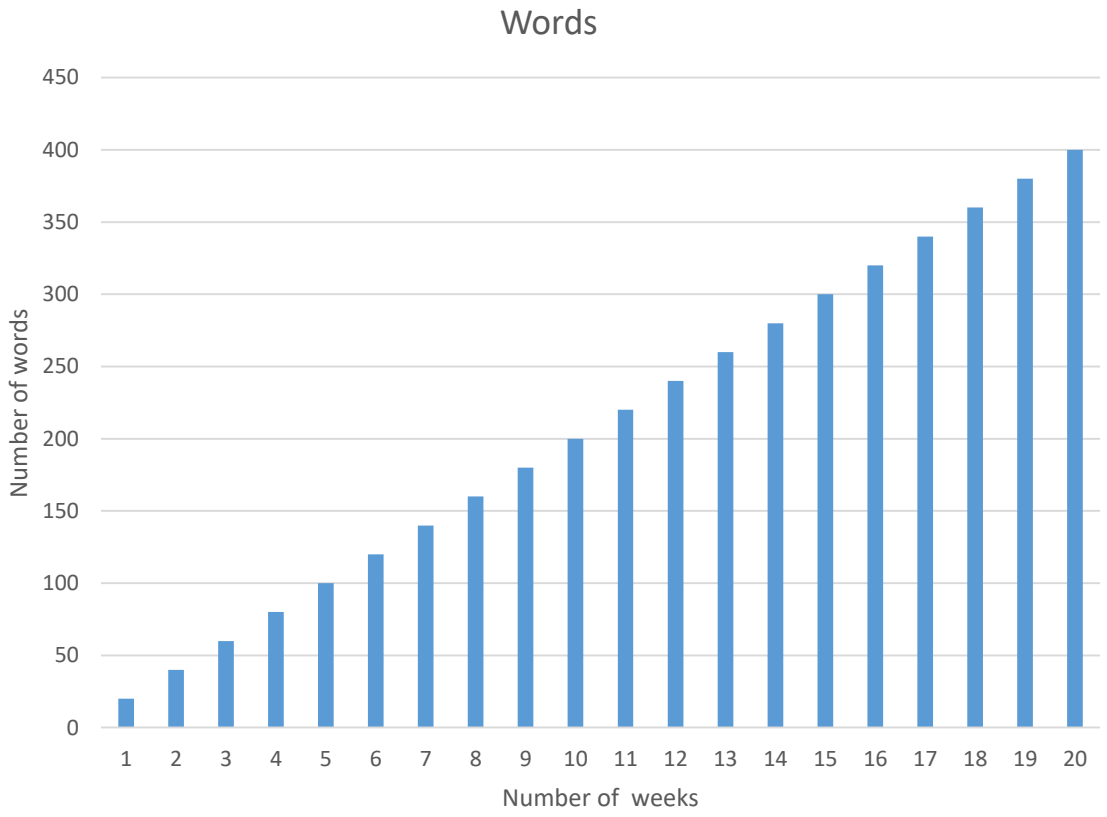
Listen_Speak Model

Total=160	**Predicted: Right**	**Predicted: wrong**	
Actual: Right	TP=143	FN=7	150
Actual: wrong	FP=0	TN=10	10
	143	17	

Read_Speak Model

Assumption with proposed plan





Result and Conclusion

- ✓All 400 words audio file(converted by ML model) tested & evaluated with the teacher.
- ✓Some words are difficult to pronounce ,so kept in the last.
- ✓To make the model more interactive some puzzles and surprizes are included at the end of every lesson(comprising only 4 words)
- ✓Error rate for Listen_Speak model and Read_Speak model are 8.13% and 8.75%.
- ✓Target is to teach 400 words in 20 weeks with gap and revision frequently(spaced repetition technique).
- ✓Reducing the time span of 52 weeks to 20 weeks only.
- ✓With internal evaluation it is found that its easy to use tool to teach and correct the pronunciation of Urdu word from the text book.

Future Work

- Plan to implement Writing evaluation in later stage of the project.
- This work extended for sentence and paragraph level also.
- Furthermore, the similar framework from text book to app can be designed for other subjects like maths, science etc.

References

Authors	Title	Year	Published In
Xin, Yang	The Design of Automatic Correction System for Spoken English Pronunciation Errors Based on Phonemes	2020	Proceedings - 2019 6th International Conferen...
Kamimura, Kohei; Takano, Kosuke	Pronunciation Error Detection in Voice Input for Correct Word Suggestion	2019	IES 2019 - International Electronics Symposium...
He, Yanzhang; Sainath, Tara N.; Prabhavalkar, Rohit; McGraw, Ian; Alvare...	Streaming End-to-end Speech Recognition for Mobile Devices	2019	ICASSP, IEEE International Conferen...
Tabibian, Behzad; Upadhyay, Utkarsh; De, Abir; Zarezade, Ali; Schölkopf, Bernhard;...	Enhancing human learning via spaced repetition optimization	2019	Proceedings of the National Academy of S...
Islam, Jahirul; Mubassira, Masiath; Islam, Md Rakibul; Das, Amit Kumar	A speech recognition system for Bengali language using recurrent Neural network	2019	2019 IEEE 4th International Conferen...
Zia, Haris Bin; Raza, Agha Ali; Athar, Awais	Pronouncur: An Urdu pronunciation lexicon generator	2019	LREC 2018 - 11th International Conferen...
Dolkar, Kesang; Tshering, Sonam; Bidha, Chencho; Thinley, Damcho; Dhungyel, Pr...	Interactive Dzongkha Learning Apps for Kids	2018	Proceedings - 2017 International Conferen...
Wihidayat, Endar Suprih; Utami, Yuniati Dwi; Budianto, Aris	Learn Arabic Language App, Mobile Based Application for Self-Directed Learning	2018	2018 4th International Conference on Educati...
Asadullah; Shaukat, Arslan; Ali, Hazrat; Akram, Usman	Automatic Urdu Speech Recognition using Hidden Markov Model	2016	2016 International Conference on Image, ...
Shirodkar, Nilkanth Shet	KONKANI SPEECH TO TEXT RECOGNITION USING HIDDEN MARKOV MODEL TOOLKIT Submitted by in partial fulfillment for the award of the degree Department of Comp...	2016	
Shaukat, Arslan; Ali, Hazrat; Akram, Usman	Mel Frequency Cepstral Coefficients (MFCCs) Model Dictionary :	2016	2016 XXI Symposium on Signal Processing, Ima...
Heil, Catherine Regina; Wu, Jason S; Lee, Joey J; Schmidt, Torben	A Review of Mobile Language Learning Applications: Trends, Challenges, and Opportunities	2016	The EuroCALL Review


Q&A

THANK YOU