# AKSHARAM

An Educational Platform for Malayalam Language

**Guide Name**
Ms. Philo Sumi
Assistant Professor
Department of AIDS

**Team 21:**
Marc George - Roll No : 43
Merlin Sarah Jiju - Roll No : 45
Timon.K.John - Roll No : 67

# Contents

# Problem Statement

**Problem**

Traditional textbook-based learning methods for concepts, subjects, and languages have become increasingly unappealing to modern learners. With the average adult attention span now at just 8.25 seconds, there is a growing demand for faster, more engaging, and cost-effective approaches to language acquisition. In our state, Malayalam remains predominantly taught in conventional classroom settings, limiting accessibility and failing to align with evolving learning preferences.

**Solution**

**Aksharam, AI Powered Malayalam Learning Platform**

# Objective

- The objective of this project is to develop an AI-powered Malayalam learning platform that enhances language acquisition through interactive experiences.
- The system integrates a LeNet-based handwritten character recognition model to assist users in learning how to write Malayalam characters with real-time feedback.
- It also provides pronunciation assistance, contextual learning, and conversation-based examples.
- The platform employs OCR and a translation model to extract and translate Malayalam text from images.

# Existing system

- **Malayalam Alphabets 360**
  - This is a language learning tool that focuses on teaching Malayalam alphabets. It provides interactive lessons, pronunciation guides, and practice exercises to help users learn to read and write Malayalam characters.
- **Ling learn**
  - Ling Learn is a language learning app that likely offers courses in multiple languages, including possibly Malayalam. It features interactive lessons, conversation practice and pronunciation.
- **Malayalam 101 - Learn to Write**
  - Malayalam 101 is an app with basic language learning functionalities with a focus on vocabulary acquisition and simple phrases. It provides flashcards with audio pronunciation by native speakers.
- **Enjuvadi**
  - Enjuvadi focuses on teaching Malayalam alphabet through interactive exercises and visual learning. The application provides letter tracing activities where users can practice writing Malayalam characters by following on-screen patterns and includes audio pronunciation guides.

# Proposed System

- The proposed system aims to assist users in learning Malayalam characters, words, and sentences through AI-powered handwritten character recognition and language translation.
- Utilizes computer vision and optical character recognition technologies to enhance language comprehension by accurately extracting and analyzing Malayalam text and translating it so users can understand.
- Focused on usability, accuracy, and accessibility to help learners master Malayalam in an interactive and enjoyable manner.
- Suitable for students, language learners, and non-native speakers.
- Scalable to incorporate advanced features such as voice recognition and AI-driven personalized learning paths for a more immersive and adaptive learning experience.

## A. Data Flow Diagrams

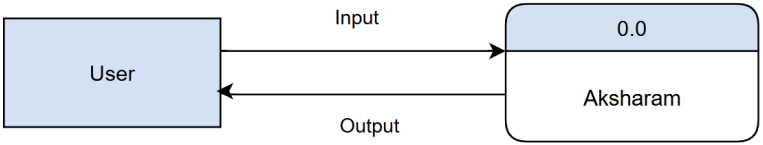



Fig. 1: Level 0 Data Flow Diagram

# Proposed System

## A. Data Flow Diagrams

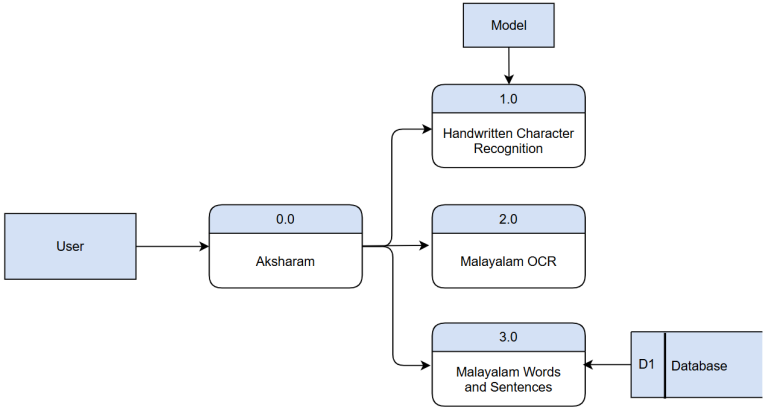



Fig. 2: Level 1 Data Flow Diagram

# Proposed System

## A. Data Flow Diagrams

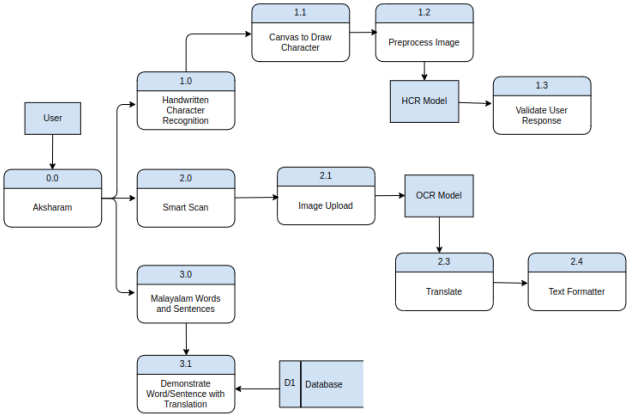



Fig. 3: Level 2 Data Flow Diagram

# Proposed System

## B. Architecture Diagram

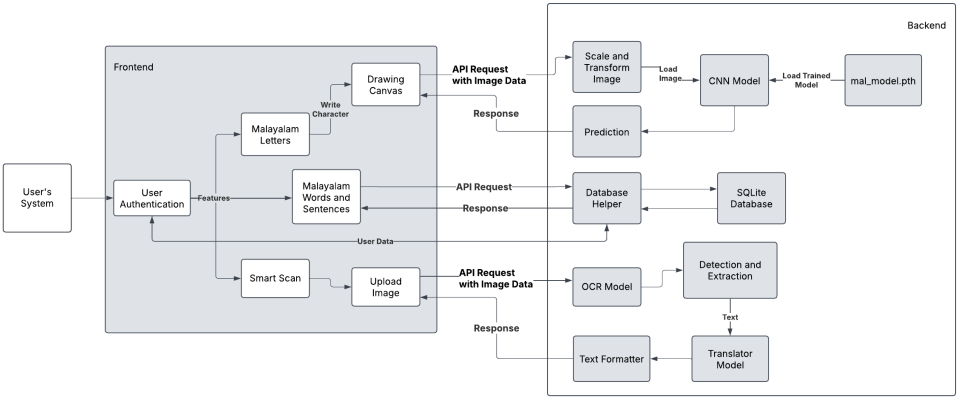



Fig. 4: System Architecture Diagram

**C. Module Details**

- **1. Letter Classification Module:**
- The user is shown how to write the character and then their written input is captured in the canvas provided. The image is pre-processed and sent to a classification model, which determines the correct alphabet and decides whether the user can proceed to the next letter.
- **Components of Letter Classification Module:**
- Demo: Demonstrates how to write the letter.
- Canvas: The user then writes the letter on the given space.
- Classification Model: The written letter is then inputted to the model where it is classified.
- Result: Determines whether the user can move on to the next character or not.

**C. Module Details**

- **2. Smart Scan Module:**
- The Image-Based Learning Module helps users learn Malayalam by extracting and translating text from images. Using an OCR model, the system detects Malayalam words in images, while a translation model converts them into the user's preferred language. This feature enhances learning by providing real-world examples, making language acquisition more interactive and engaging.
- **Components of Smart Scan Module:**
- Image Input: Users upload or capture an image containing Malayalam text.
- OCR Model: Extracts Malayalam words from the image.
- Translation Model: Translates the extracted text into the desired language.
- Result: Shows the original and translated text for learning.

**C. Module Details**

- **3. Words and Sentences Learning Module:**
- Helps users build their Malayalam vocabulary and conversational skills using a structured database. It teaches Malayalam words by displaying their meanings, pronunciation. Additionally, it introduces basic Malayalam sentences providing example sentences along with its English translations to enhance understanding and fluency.
- **Components of Words and Sentences Learning Module:**
- Demonstrates the Word/Sentence
- Provides English translation for understanding

# Novelty of Proposed System

- **Handwritten Character Recognition**
  - Users can write characters in their own handwriting and get real-time accuracy feedback.
- **Smart Scanner for Malayalam Text**
  - Scans and translates Malayalam text from images, enhancing user engagement.
- **Optimized Model Training**
  - Trained on multiple datasets for improved recognition efficiency.

**D. Software Requirements**

- Pytorch - For the handwritten Character recognition model and translator model.
- SuryaOCR - For Optical Character Recognition of Malayalam Words.
- OpenCV - For Preprocessing images.
- Django - For the backend of the platform.
- ReactJS - For an interactive and user friendly Interface
- SQLite - For storing User data, malayalam words and sentences

# Results and Discussion

## Home Page

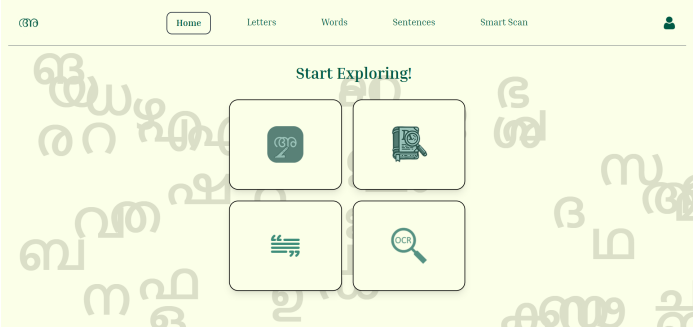



Fig. 5: Home Page

## Letters Page

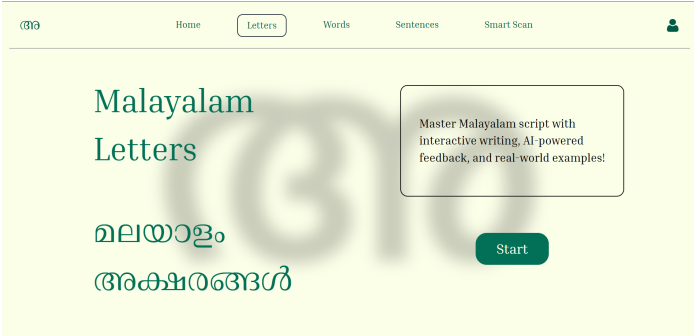



Fig. 6: Letters Page

## Letters Page

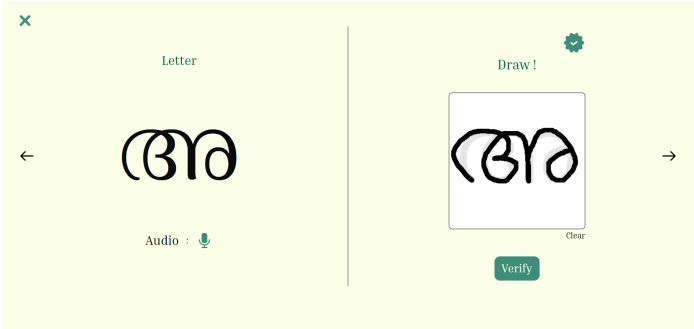



Fig. 7: Teaching Letters

## Smart Scan


Fig. 8: Smart Scan Feature

## Smart Scan


Fig. 9: Smart Scan Translation

# Results and Discussion

## Confusion Matrix

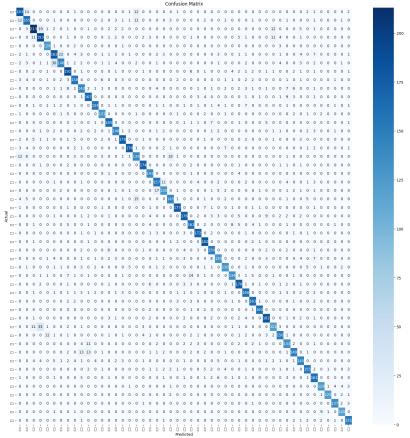



Fig. 10: Letters Confusion Matrix

# Results and Discussion

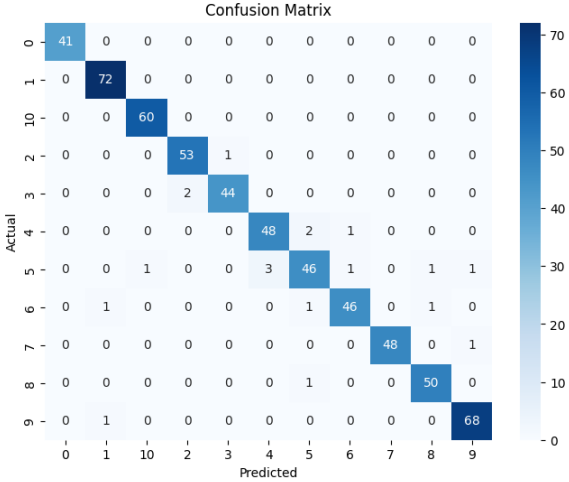
## Accuracy Scores

| Accuracy Scores | | | | | |
|---|---|---|---|---|---|
| അ | 77.83 | ജ | 85.79 | ഫ | 82.72 |
| ആ | 83.9 | ഝ | 86.5 | ബ | 79.8 |
| ഇ | 80.25 | ഞ | 80.46 | ഭ | 81.71 |
| ഉ | 82.85 | ട | 96.45 | മ | 89 |
| എ | 80.93 | ഠ | 94.64 | യ | 80.95 |
| ഏ | 77.03 | ഡ | 85.86 | ര | 91.62 |
| ഒ | 79.82 | ഢ | 74.1 | റ | 97.2 |
| ക | 79.8 | ണ | 87.25 | വ | 80.79 |
| ഖ | 74.23 | ത | 87.89 | ശ | 79.43 |
| ഗ | 86.41 | ഥ | 83.68 | ഷ | 74.42 |
| ഘ | 85 | ദ | 78.95 | സ | 84.11 |
| ങ | 89.02 | ധ | 93.62 | ഹ | 86.59 |
| ച | 90.15 | ന | 93.94 | ള | 72.78 |
| ഛ | 87.63 | പ | 89.94 | ഴ | 84.75 |

Fig. 10: Letter Accuracy Scores

# Results and Discussion

## Confusion Matrix



# Results and Discussion

## Accuracy Scores

| Accuracy Scores | |
|---|---|
| ◌ | 100 |
| ◌ാ | 98.15 |
| ◌ി | 95.65 |
| ◌ീ | 94.12 |
| ◌ു | 86.79 |
| ◌ൂ | 93.88 |
| ◌ൃ | 97.96 |
| ◌െ | 98.04 |
| ◌േ | 98.55 |
| ◌ൗ | 100 |
| ◌ൊ | 100 |

Fig. 10: Symbol Accuracy Scores

# Comparison Table

Table 1: Comparison Study

| Features | Aksharam | Malayalam 101 | Enjuvadi | Malayalam Alphabets 360 | Ling Learn |
|---|---|---|---|---|---|
| Handwritten Character Recognition | ✅ Available | ❌ Not Available | ❌ Not Available | ❌ Not Available | ❌ Not Available |
| Real-time Writing Feedback | ✅ Available | ❌ Not Available | ❌ Not Available | ❌ Not Available | ❌ Not Available |
| Image-based Text Translation | ✅ Available | ❌ Not Available | ❌ Not Available | ❌ Not Available | ❌ Not Available |
| Contextual Learning Examples | ✅ Available | ❌ Not Available | ✅ Available | ❌ Not Available | ❌ Not Available |
| Interactive Learning Interface | ✅ Available | ✅ Available | ✅ Available | ❌ Not Available | ✅ Available |
| AI-based Classification Model | ✅ Available | ❌ Not Available | ❌ Not Available | ✅ Available | ❌ Not Available |
| Self-paced Learning Structure | ✅ Available | ✅ Available | ✅ Available | ❌ Not Available | ✅ Available |
| User Progress Analytics | ✅ Available | ❌ Not Available | ✅ Available | ✅ Available | ❌ Not Available |

# Conclusion and Future Scope

- The AI-powered Malayalam learning platform enhances language acquisition through handwritten character recognition, OCR-based text extraction, and AI-driven translation for an interactive learning experience.
- By providing contextual learning with real-world examples and a self-paced structure, the platform promotes accessibility, cultural awareness, and modernizes language education.
- Improve recognition of handwritten Malayalam words using advanced AI models.
- Enhance the OCR translator for accurate extraction and translation of Malayalam scripts.
- Integrate a chatbot for real-time conversations, enabling interactive language learning.
- Expand user engagement with quizzes and questionnaires to reinforce learning.

# References


- Vaishakh.V.K ,Lyla.B.Das, “Handwritten Malayalam Character Recognition System using Artificial Neural Networks”, 2020.
- Baiju.K.B ,Sabna.T.S, Lajish.V.L, “Segmentation of Malayalam Handwritten Characters into Pattern Primitives and Recognition using SVM”, 2020.
- Manjusha K, Anand Kumar Madasamy, Soman Kp, ”On developing handwritten character image database for Malayalam language script”, 2019
- Anish S, Preeja V, “A Novel Method for Malayalam Handwritten Character Recognition”, 2015
- M. Abdul Rahiman and Rajasree M. S., “An Efficient Character Recognition System for Handwritten Malayalam Characters Based on Intensity Variations”, 2011

**Thank You!**