

Title of Research

**Speech Recognition System Using Deep Learning
Techniques for Assamese Sign Language.**

Prepared by

Name	Eng. Sultan Ali Senan
Academic ID No	253002563
Program	Master of Science in Artificial Intelligence
Course Name	Final Project (1) " AI900a"
Supervised	Dr. Amjad Qatish

2026

Task 5

Thesis Completion Project Plan (12–14 Weeks)

Phase 1: Methodology & System Architecture (Weeks 1–3)

Goal: Complete Chapter 3 (Proposed Approach).

- **Week 1: Data Finalization.** Finalize the collection of the Assamese Sign Language dataset (the 2,094 data points mentioned in your review). Perform data augmentation (rotation, flipping, brightness) to increase model robustness.
- **Week 2: Preprocessing Pipeline.** Implement the **MediaPipe** landmark extraction scripts. Convert raw video/images into coordinate-based sequences (21 hand landmarks).
- **Week 3: Model Design.** Define the hybrid **CNN-LSTM** architecture.
 - Draft Chapter 3: Explain the mathematical flow, the layers of the CNN (for spatial features), and the LSTM (for temporal movement).

Phase 4: Implementation & Training (Weeks 4–7)

Goal: Build the engine for Chapter 4.

- **Week 4: Environment Setup.** Set up Python, TensorFlow/PyTorch, and GPU acceleration.
- **Week 5: Training.** Train the model on the Assamese dataset. Perform hyperparameter tuning (learning rate, dropout layers, batch size).
- **Week 6: Speech Integration.** Integrate the **Text-to-Speech (TTS)** module (e.g., using gTTS or pyttsx3) so that recognized signs trigger audio output.
- **Week 7: Real-Time Prototype.** Develop a simple GUI or OpenCV window that captures webcam feed and displays the recognized Assamese sign and plays the corresponding audio.

Phase 3: Testing & Results Analysis (Weeks 8–10)

Goal: Complete Chapter 4 (Experimental Results).

- **Week 8: Performance Metrics.** Calculate Accuracy, Precision, Recall, and F1-Score. Generate **Confusion Matrices** to see which Assamese signs are being confused by the AI.
- **Week 9: Comparative Analysis.** Compare your CNN-LSTM results against baseline models (like a simple CNN or Random Forest) to justify your methodology.
- **Week 10: Writing Chapter 4.** Document the results with charts, graphs (Accuracy/Loss curves), and screenshots of the real-time system in action.

Phase 4: Finalizing & Submission (Weeks 11–12)

Goal: Complete Chapter 5, References, and Formatting.

- **Week 11: Conclusion & Future Work.** Draft Chapter 5. Discuss the limitations (e.g., lighting, background noise) and future goals (e.g., mobile app deployment).
- **Week 12: Final Review.**
 - Ensure all **References** are in the correct format (as seen in your list).
 - Check **List of Figures** and **List of Tables** for correct page numbering.
 - Submit the final draft to your supervisor for feedback.

Key Deliverables for the Term:

1. **A Functional Model:** A trained .h5 or .pth file capable of recognizing Assamese signs.
2. **The Speech Module:** A working script that converts detected signs into audio.
3. **Complete Chapters 3, 4, and 5:** Moving your thesis from an "overview" to a "completed work."
4. **Final Thesis Document:** Ready for submission to the College of Informatics, Midocean University.