

**TRIBHUWAN UNIVERSITY**

**INSTITUTE OF SCIENCE AND TECHNOLOGY**

**TEXAS INTERNATIONAL COLLEGE, KATHMANDU**

**A Project Proposal on**

**“Shishu-Care: Infant Cry Classifier”**

**Submitted To:**

**DEPARTMENT OF CSIT**

**TEXAS INTERNATIONAL COLLGE**

**A Project Proposal**

**Submitted By:**

**Nishan Kharel(5-2-926-25-2021)**

**Rup Jung Malla(5-2-926-32-2021)**

**Nabin Katwal(5-2-926-40-2021)**

Contents

[Introduction 2](#_Toc200830687)

[Problem Statement 2](#_Toc200830688)

[Objectives 3](#_Toc200830689)

[Methodology 4](#_Toc200830690)

[a.Requirement Identification 4](#_Toc200830691)

[I.Literature Review 4](#_Toc200830692)

[II. Requirement Analysis 5](#_Toc200830693)

[i. Functional Requirement 5](#_Toc200830694)

[ii. Non-Functional Requirements 5](#_Toc200830695)

[b. Requirement Feasibility 5](#_Toc200830696)

[I. Technical Feasibility 5](#_Toc200830697)

[II. Operational Feasibility 6](#_Toc200830698)

[III. Economic Feasibility 6](#_Toc200830699)

[IV. Schedule Feasibility 7](#_Toc200830700)

[c. High-Level Design of System 7](#_Toc200830701)

[I. Methodology of the Proposed System 8](#_Toc200830702)

[II. Flow Charts/Working Mechanism of Proposed System 9](#_Toc200830703)

[III. Description of Algorithms 10](#_Toc200830704)

[Expected Outcome 10](#_Toc200830705)

[References 11](#_Toc200830706)

# Introduction

Newborn babies' major form of communication is crying. It is through crying that they get to communicate their needs, whether it is hunger, pain, or discomfort. New parents, however, cannot interpret the cries. New parents believe all cries sound alike, and this results in late or inappropriate responses. This works against the well-being of the baby and reduces parents' confidence. Shishu**-**Care, a web application can automatically identify and categorize baby cries in real-time. It will capture sound through a website, using trained deep learning models to detect whether a baby is crying and categorize the cry by specific needs. The goal is to give quick and reliable feedback to parents and caregivers. It will allow them to better react to infant behavior. When the project building is finished, we expect to have a working model that can effectively determine and classify infant cries. This will be valuable aid to parents, especially new parents, and caregivers.

# Problem Statement

It is a great challenge for caregivers and parents in Nepal and the world to understand the needs of the newborn since the babies are unable to speak. This leads to several problems which are as follows:

• Most of the children's cry cannot be differentiated by parents, and so it is hard to understand and react accordingly.

• There is no easy-to-use digital gadget in Nepal currently that can hear baby cries as they happen and help identify why they are crying. Lack of technology adds stress to parents and caregivers and causes delays in the proper reactions.

• Research shows that human accuracy for crying cause detection is low, leading to misinterpretation and unmet infant needs. This is especially alarming in families with low parenting experience.

• AI and deep learning have demonstrated high prospects in audio classification applications, their use in infant care particularly in a localized, affordable manner continues to be underdeveloped in the context of Nepal.

# Objectives

The general aim of the Shishu-Care project is to create an intelligent, simple, and reliable deep learning-based online system that will analyze and classify baby cries in real time. The objectives of this project are:

* To help caregivers understand what the baby needs by listening to and learning baby cries. They can categorize these cries into hunger, pain, and so on, so parents can react quickly and positively.
* To bridge the gap in a parent-baby communication by using deep learning models to decode baby cries effectively and emotional tension are reduced.
* To create a basic website layout that has a special interface and works using audio input.
* To improve infant care by providing feedback immediately by delivering short information about the cry type so that caregivers can understand earlier and make the right decisions.
* To promote caring parental support through technology by making the platform accessible to everyone, including new parents and caregivers.
* To assist future research and integration with healthcare by offering structured, data sets and cry categorizations that would assist children's research or be utilized in intelligent childcare systems.

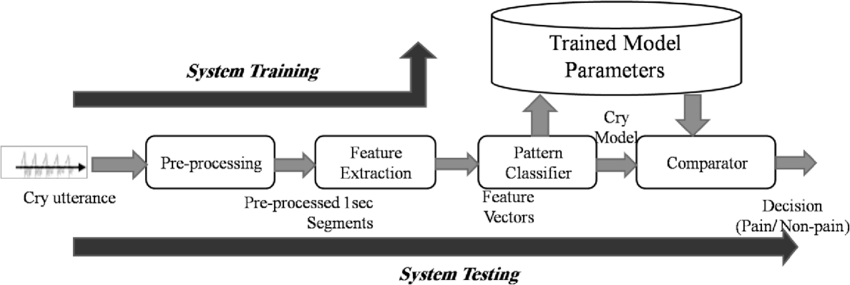
# Methodology

## a.Requirement Identification

The requirement identification phase is crucial in understanding what the Shishu care system must achieve. This section outlines all the essential functional and non-functional requirements of the system to ensure its successful development and implementation.

### I.Literature Review

Infant cry recognition usually comprises audio signal processing and machine learning model classification. Numerous works, including those of [Petroni et al., 1995] and [Aureli et al., 2014], concentrated on the acoustic description of various kinds of baby cries in relation to acoustic features comprising pitch, intensity, and frequency patterns. The research came to the conclusion that there are specific types of cries, e.g., hunger, pain, and discomfort, which have recognizable audio patterns that can be represented computationally.



It is the past project which has been done already decades ago. Here, pre-processing has been done and feature extraction with pattern classifier with trained Model Parameter. Chorna et al. (2014) used cry acoustics for assessing neurological development, focusing more on clinical diagnosis than daily care. Ravichander et al. (2017) used MFCCs and KNN/SVM to differentiate types of cries in a controlled environment.

### II. Requirement Analysis

### i. Functional Requirement

The functional Requirement define the basic functionality of Shishu Care system which perform operations like:

* The system must capture real time data audio whenever baby cries using microphone or audio recognition devices.
* The system must classify the cry “hunger” or “Not Hungry” using trained model.

### ii. Non-Functional Requirements

Non-functional requirements define the overall attributes and constraints of the system performance:

* Usability: The system must be easy to use, with minimal user interaction.
* Accuracy: The system must be easy to use with accurate more than 80%.
* Performance: The detection of crying and classification must be done automatically.

## b. Requirement Feasibility

### I. Technical Feasibility

The proposed project, Shishu-Care, is to recognize a baby's cry and determine if the baby is hungry or not using sound classification and machine learning algorithms. The project is feasible from a technical perspective because it takes audio as an input from a microphone, preprocessing techniques, and deep learning models like CNNs.

* Technology Stack: The platforms and tools that are to be used are Python, TensorFlow, Convolutional Neural Network and Classification and VS-code as IDE platforms, which are all accessible to the development team.
* Algorithm:

Shishu Care, the project begins by extracting audio features from recorded baby cries using the librosa library. This typically involves the use of Mel-Frequency Cepstral Coefficients (MFCCs), which effectively capture the important acoustic properties of sound. KNN is a simple yet powerful classification algorithm that works by comparing a given data point with its closest neighbors in the feature space. Additionally, Support Vector Machine and Random forest classifer algorithm are also used together contributing toward the development of an intelligent assistant for infant care.

### II. Operational Feasibility

The system will be simple and intuitive for users such as parents or caregivers. It will provide clear indications (e.g., visual or audio signals) if the baby is crying due to hunger. As the system automates an important part of infant care, it is likely to be accepted and prove beneficial to the users, especially for new parents.

### III. Economic Feasibility

Shishu-care project is economically feasible as it mainly relies on free and open-source software tools. The only potential costs may involve acquiring a high-quality microphone or sound sensor, but these costs are minimal. We team members will contribute our time and effort without any monetary compensation, making it a cost-effective academic project.

### IV. Schedule Feasibility

The development timeline is planned to accommodate the academic semester. Initial weeks are assigned for planning and requirement analysis, followed by design and, training, testing, and integration of models i.e. development. Weeks in subsequent periods are assigned for testing, and documentation. Although the phases seems sequential, but we will be following agile approach.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Week | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Planning |  |  |  |  |  |  |  |  |  |  |  |  |
| Requirement Analysis |  |  |  |  |  |  |  |  |  |  |  |  |
| System Design |  |  |  | |  |  |  |  |  |  |  |  |
| Implementation |  |  |  |  | | | | |  |  |  |  |
| Testing |  |  |  |  |  |  |  |  | |  |  |  |
| Documentation |  |  |  |  |  |  |  |  |  | |  |  |
| Deployment |  |  |  |  |  |  |  |  |  |  |  | |

## c. High-Level Design of System

System Architecture

The ShishuCare system is to be developed as a dynamic web application that is built on top of artificial intelligence to provide real-time analysis of infant cries. It will be designed based on a Three-Tier Architecture, a standard client-server model, separating the application into following layers:

• Presentation Layer, UI & front-end

• Application Layer, Back-end logic

• Data Layer, Database & trained model

### I. Methodology of the Proposed System

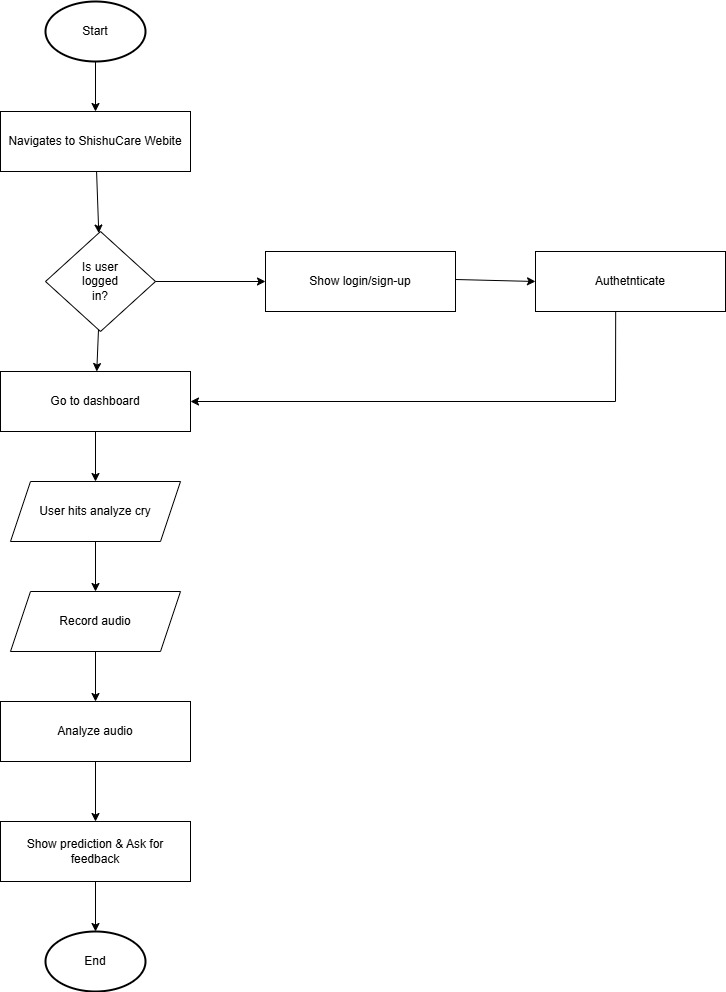
All the project phases will be executed using an Agile development Methodology in sprints. There is high technical uncertainty, especially in the nature of the data, which affects accuracy of the measurement. Any AI-based system is probabilistic in nature. The model requires continuous fine-tuning and adjustments which contradicts the principles of sequential workflow. If the model fails deliberately in agile approach, we have plenty of time to search for different dataset, try a different model and even adjust the project’s scope.

The system will be designed under the object-oriented paradigm. It will consist of collection of interacting objects, each representing a distinct entity with its own attribute and behavior. We have following classes:

1. User: Parent with the attribute email, username and password. He/She will record the audio of baby crying.

2. CryClassifier: This class will encapsulate the trained model object and contains simple methods like .predict(), hiding the complex TensorFlow/Keras logic from the rest of the application.

### II. Flow Charts/Working Mechanism of Proposed System



### III. Description of Algorithms

1. Cry Classification Pipeline

This will be the central algorithm of the project. It automates the tasks from pre-processing to final prediction or classification of the audio. The pipeline will consist of audio preprocessing, feature extraction, Convolutional Neural Network and, softmax activation.

1. Password hashing algorithm

It securely stores user passwords in the database in a way, preventing them from being read, in case of a data breach. For this purpose, we will be using adaptive hashing algorithm like bcrypt.

1. RESTful API Request Routing

It maps the incoming HTTP requests from the client to the appropriate function on the backend which will handle the request. Therefore, it is important for proper flow of data between front-end and back-end of the system.

# Expected Outcome

The goal of the Shishu-Care project is to be achieved by creating a web application that would hear and classify the baby cries in real-time when and if they happen. The expected outcomes are as below:

* Cry Detection Module: To detect whether the baby is crying or not whenever the baby cries, deep learning algorithm will be used.
* Cry Classification Engine: A model trained on a number of different classified cries and classify the type of crying.
* Web-Based Interface: A simple web application that will provide parents with the ease of uploading audio, seeing the classification outcome, and getting recommendations.
* Improved Caregiver Response: The system allows caregivers and parents to respond in the correct time and in the appropriate manner to infant crying by detecting cries in real time with precision. The infant's health is improved and the stress level of the caregivers is lessened.
* Documents and Model Reports: Complete information on the data set, data preparation, model construction, training, and how success will be evaluated. This will allow for understanding and future research or development.

# References

***Liang, Y.-C.****,* ***Wijaya, I.****,* ***Yang, M.-T.****,* ***Cuevas Juarez, J. R.****, &* ***Chang, H.-T.*** *(2022).* Deep Learning for Infant Cry Recognition*.* Frontiers in Public Health

***Sun, B.****,* ***Liu, C.****,* ***Yang, S.****,* ***Wang, W.****, &* ***Mei, Y.*** *(2025).* ResCapsnet: a capsule network with CRAM and BiGRU for sound event detection*.* EURASIP Journal on Audio, Speech, and Music Processing

***Abbaskhah, A.****,* ***Sedighi, H.****, &* ***Marvi, H.*** *(2023).* Infant cry classification by MFCC feature extraction with MLP and CNN structures*.* Biomedical Signal Processing and Control

***Hammoud, M.****,* ***Getahun, M. N.****,* ***Baldycheva, A.****, &* ***Somov, A.*** *(2024).* Machine learning‑based infant crying interpretation*.* Frontiers in Artificial Intelligence

***Maghfira, T. N.****,* ***Basaruddin, T.****, &* ***Krisnadhi, A.*** *(2020).* Infant cry classification using CNN–RNN*.* Journal of Physics: Conference Series