Real-Time Classroom Monitoring System For Student Performance Using Machine Learning

Project Submitted By

Asmita Ghorai

Pallabi Biswas

Ratri Dey

Riddhi Basu

BACHELOR OF TECHNOLOGY (CSE)

DEPARTMENT OF COMPUTER SCIENCE

Sister Nivedita University, Newtown

Kolkata, West Bengal

2025

Real-Time Classroom Monitoring System For Student Performance Using Machine Learning

Project Submitted in Partial Fulfilment of the Requirements for Award of the Degree of

Bachelor of Technology (CSE)

Student Name	**Reg. Number**	**Email ID**
Asmita Ghorai	**220010317574**	**ghoraiasmita@gmail.com**
Pallabi Biswas	**220010788506**	**pallabibiswas4002@gmail.com**
Ratri Dey	**220010988119**	**ratridey0@gmail.com**
Riddhi Basu	**220011009400**	**riddhib4321@gmail.com**

Submission Date: November 25, 2025

Under the supervision of

Dr. Bidyut Saha
Assistant Professor
Department of Computer Science
Sister Nivedita University, Newtown
Kolkata, West Bengal

A Satyam Roychowdhury initiative

SNU
SISTER NIVEDITA UNIVERSITY

DEPARTMENT OF COMPUTER SCIENCE

Sister Nivedita University, Newtown

Kolkata, West Bengal

November, 2025

Declaration

We hereby declare that this dissertation is the product of our own work, and we attest that it contains no material that resulted from collaboration, except where explicitly acknowledged in the text. Furthermore, we confirm that this project has not been previously submitted, either in part or in its entirety, to any other University or Institution for the purpose of obtaining any degree, diploma, or other qualification. All sources used and referenced in this dissertation are duly credited, and any borrowed ideas or information are appropriately cited in accordance with academic standards and guidelines.

. . . .

Date: 25-11-2025
Place: SNU, WB

Asmita Ghorai
Registration Number:220010317574

. . . .

Pallabi Biswas
Registration Number: 220010788506

. . . .

Ratri Dey
Registration Number: 220010988119

. . . .

Riddhi Basu
Registration Number: 220011009400

Certificate

This is to certify that the project entitled "Real-Time Classroom Monitoring System For Student Performance Using Machine Learning", submitted by Asmita Ghorai , Pallabi Biswas, Ratri Dey, Riddhi Basu to Sister Nivedita University, West Bengal for the award of the degree of Bachelor of Technology (CSE) is a bonafide record of the project work carried out by them under my supervision and guidance. The content of the project, in full or parts have not been submitted to any other institute or university for the award of any degree or diploma.

Date: 25-11-2025
Place: SNU, WB

....

(Dr. Bidyut Saha)
Assistant Professor
Dept. of Computer Science
Sister Nivedita University

Contents

Sister Nivedita University, WB

List of Tables

List of Figures

 Sister Nivedita University, WB

Acknowledgement

We would like to express our heartfelt gratitude to our esteemed guide, **Dr. Bidyut Saha**, Assistant Professor at Sister Nivedita University, WB, for his invaluable guidance, unwavering support, and encouragement throughout our project journey. His profound knowledge, insightful suggestions, and mentorship have been instrumental in shaping our project and academic pursuits.

We extend our sincere thanks to the members of the Project Committee at the Department of Computer Science , Sister Nivedita University, for their valuable feedback and constructive criticism during the evaluation of our work. Their expertise and insights have been crucial in refining the quality of our project.

We would like to express our heartfelt gratitude to our HOD, the esteemed Vice-Chancellor, the Class teachers and all the faculty members of our department for their unwavering support and guidance throughout our project journey. Your invaluable contributions have been instrumental in shaping this project endeavor.

We are deeply indebted to our family members whose constant love, encouragement, and support have been the driving force behind our academic pursuits. Their unwavering belief in us has given us the strength to overcome challenges and achieve milestones.

We also want to express our heartfelt gratitude to our friends, colleagues, juniors, and seniors for their camaraderie, encouragement, and support throughout this academic journey. Their presence and interactions have enriched us with diverse perspectives, making this journey all the more rewarding.

Lastly, we extend our sincere appreciation to all those who have been our pillars of support during this project endeavor. Your encouragement, motivation, and belief in our abilities have been invaluable, and we are truly grateful for your presence in our life.

Asmita Ghorai (Registration No.: 220010317574)	....
Pallabi Biswas (Registration No.: 220010788506)	....
Ratri Dey (Registration No.: 220010988119)	....
Riddhi Basu (Registration No.: 220011009400)	....

Abstract

Traditional classroom monitoring approaches fail to capture the real-time emotional states and engagement levels of students, especially in large or digitally enhanced learning environments. With recent advancements in machine learning and computer vision, automated analysis of facial expressions has emerged as a reliable method for estimating attentiveness and behavioural patterns. This project presents a **Real-Time Classroom Monitoring System for Student Engagement Using Machine Learning**, which classifies facial expressions into three sentiment categories—Positive, Neutral, and Negative—using a CNN-based deep learning model.

The FER dataset from Kaggle is used as the primary source, where the original seven emotion classes are remapped into sentiment-level categories suitable for classroom analysis. The dataset undergoes several preprocessing steps including sentiment mapping, balancing, resizing, normalization, and augmentation to improve robustness. A customized CNN architecture is trained and evaluated using accuracy/loss curves and a confusion matrix. With a validation accuracy of approximately 73%, the model demonstrates effective sentiment classification performance suitable for preliminary classroom analytics.

The system can assist teachers in identifying disengaged or struggling students early, thereby supporting academic intervention and personalized learning. This work also lays the foundation for future smart classroom technologies integrating real-time behavioural analytics and multi-modal engagement monitoring.

Keywords: Machine Learning, CNN, Facial Expression Recognition, Student Engagement, Classroom Analytics

Chapter 1

Introduction

1.1 Background / Context of the Study

In recent years, digital learning environments and large classroom settings have created a growing need for objective and continuous monitoring of student engagement. Student attentiveness is a key indicator of academic performance, yet traditional monitoring methods—manual observation, attendance records, and periodic assessments—provide limited insight into real-time behavioural patterns. Advances in artificial intelligence, machine learning, and computer vision have enabled automated analysis of facial expressions, making sentiment-based monitoring a promising tool for modern educational systems.

Facial expressions provide meaningful cues about a student's emotional state, interest level, and level of understanding. By leveraging Convolutional Neural Networks (CNNs), it is now feasible to develop systems that classify sentiment from facial images with high accuracy. This project contributes to this growing field by proposing a sentiment-based classroom monitoring system that classifies facial expressions into Positive, Neutral, and Negative categories.

1.2 Problem Context / Motivation

In real classroom environments, teachers face challenges in continuously observing every student. Factors such as class size, multitasking demands, and subjective interpretation further limit the accuracy of manual engagement monitoring. Students who lose focus, experience confusion, or struggle academically may not always express these challenges verbally. As a result, teachers often miss early indicators of disengagement.

With increasing adoption of smart classrooms, there is a strong motivation to use machine learning techniques to automate behavioural and emotional analysis. A sentiment classification system can help teachers gain deeper insights into student engagement patterns, improve teaching strategies, and ensure timely academic intervention.

1.3 Problem Statement

The problem addressed in this project is the absence of a scalable, automated, and accurate system for monitoring classroom engagement through facial sentiment. Traditional monitoring methods are insufficient for capturing real-time emotional states. The proposed solution aims to classify student facial expressions into three sentiment categories—Positive, Neutral, and

 Sister Nivedita University, WB

Negative—using a CNN-based model trained on the FER dataset.

1.4 Objectives of the Project

The primary objectives of this project are:

- To develop a CNN-based model for classifying facial expressions into sentiment categories.
- To preprocess the FER dataset through sentiment mapping, balancing, resizing, and normalization.
- To apply data augmentation techniques to improve model generalization.
- To evaluate model performance using accuracy/loss graphs and a confusion matrix.
- To design system-level diagrams including the Flowchart, DFD Level 0, and System Architecture.

1.5 Scope of the Project

This project focuses on sentiment classification using facial expressions. The scope includes dataset preprocessing, CNN training, evaluation, and system design documentation. The project does not include real-time deployment, video stream processing, or analysis of multi-modal behavioural cues such as voice, body posture, or gaze direction. The implementation is restricted to static images from the FER dataset.

1.6 Significance / Contribution

The proposed system enhances classroom analytics by providing automated, objective insights into student engagement. It enables early identification of disengaged or confused learners, supports personalized teaching strategies, and reduces the cognitive load on educators. The project also contributes a simplified sentiment-based model, which is more suitable for classroom environments compared to traditional seven-emotion classification systems.

1.7 Methodology Overview

The system uses a structured pipeline involving dataset preprocessing, CNN model development, training, and evaluation. The FER dataset is preprocessed using sentiment mapping, normalization, balancing, and augmentation techniques. A custom CNN architecture is trained to classify facial expressions into three sentiment classes. Performance is evaluated using validation accuracy, loss curves, and a confusion matrix.

1.8 Organization of the Report

This dissertation is organized as follows:

Chapter 2 presents a detailed literature survey on facial expression recognition, sentiment analysis, and machine learning methods relevant to this work. Chapter 3 defines the problem, describes the current challenges, and outlines the objectives of the project. Chapter 4 discusses the proposed methodology, including system architecture, preprocessing techniques, and model design. Chapter 5 includes experimental results, analysis, and performance evaluation. Chapter 6 concludes the dissertation with key findings, limitations, and future research directions.

Chapter 2

Literature Survey

2.1 Introduction

A literature survey provides an understanding of the existing research, methodologies, and technological advancements related to facial expression recognition and engagement analysis.
It helps identify research gaps, limitations of existing techniques, and opportunities for improvement.
This chapter reviews existing facial expression recognition systems, sentiment classification methods, and deep learning approaches, with a focus on their applicability to classroom monitoring.

2.2 Existing Systems / Related Work

Facial expression recognition (FER) has been an active research area within computer vision, human–computer interaction, and behavioural computing. Early FER systems relied on handcrafted features such as Local Binary Patterns (LBP), Gabor filters, and Histogram of Oriented Gradients (HOG). Although computationally lightweight, these traditional methods struggled with variations in lighting, pose, and occlusion.

With advancements in deep learning, CNN-based models have become the dominant approach for FER. The FER dataset, CK+, JAFFE, and AffectNet are widely used benchmarks for emotion classification tasks. Deep architectures such as VGGNet, ResNet, and custom CNN models have achieved significant performance improvements by learning hierarchical spatial features directly from images.

Several works in the domain of education analytics have applied affective computing to understand student behaviour. Researchers have used CNNs for emotion detection, eye-gaze tracking for attention analysis, and multimodal systems to estimate cognitive load. These studies demonstrate the feasibility of using facial-based affective cues to understand student engagement in real-time learning environments.

2.3 Limitations of Existing Systems

Despite substantial progress, existing approaches have several limitations:

- Most FER systems classify seven or more emotion categories, which may be too fine-grained for educational applications.

- Many models require high-quality, frontal-face images, making them less robust to realistic classroom conditions.
- Traditional FER datasets are often imbalanced, affecting the reliability of classification results.
- Real-time FER applications may require substantial computational resources not feasible for all classrooms.
- Existing educational systems lack simplified sentiment-level classification (Positive, Neutral, Negative) that aligns better with engagement analysis.

2.4 Comparative Gap Analysis

A comparison of previous research highlights the following gaps:

- **Lack of sentiment-focused models:** Most studies emphasize seven-emotion classification without grouping emotions into high-level sentiment categories.
- **Limited educational context:** Many FER models are evaluated in controlled environments but not in classroom-like scenarios.
- **Insufficient preprocessing pipelines:** Variations in class distribution, lighting, and pose are not consistently addressed, limiting model robustness.
- **Need for lightweight architectures:** Classroom systems require efficient models suitable for real-time execution.

These gaps justify the development of a simplified, sentiment-based FER model tailored for classroom engagement monitoring.

Table 2.1: Comparison of Selected Facial Expression Recognition (FER) Methods

Reference	**Method**	**Dataset**	**Key Outcome / Limitation**
Barsoum et al. (2016)	CNN with label distribution learning	FER2013	Improved robustness to noisy labels; limited by annotation quality.
Whitehill et al. (2014)	Facial features + classifier for engagement	Classroom video data	Demonstrated engagement prediction; dataset too small for generalization.
Simonyan & Zisserman (2014)	VGGNet (deep convolutional model)	ImageNet (transfer learning)	High accuracy but very computationally heavy for real-time classroom use.
Shorten & Khoshgoftaar (2019)	Image augmentation strategies	—	Showed augmentation improves generalization; widely adopted in FER.
Proposed Work (2025)	Custom lightweight CNN + augmentation	FER (mapped to sentiments)	73% validation accuracy; optimized for classroom sentiment monitoring.

2.5 Summary of Literature Review

This chapter reviewed the major FER techniques, deep learning models, and educational engagement systems. CNN-based models outperform traditional handcrafted approaches due to their superior feature extraction capabilities. However, existing methods often lack sentiment-level classification, robustness to real-world variation, and applicability in educational settings. The insights derived from this survey helped shape the design of a sentiment-based CNN model optimized for classroom analytics, as described in the subsequent chapters.

Chapter 3

Problem Identification

3.1 Context / Background of the Problem

The increasing integration of digital learning platforms and smart classrooms has highlighted the need for objective methods to monitor student engagement. Student inattentiveness, confusion, or emotional disengagement often goes unnoticed in large classroom environments. While facial expressions provide rich indicators of emotional and cognitive states, traditional observation methods are subjective, inconsistent, and impractical for continuous monitoring. Machine learning and computer vision offer an automated means to bridge this gap through real-time sentiment analysis based on facial cues.

3.2 Problem Definition

The core problem addressed in this project is the absence of an automated, accurate, and scalable system for identifying student engagement level through facial sentiment. Existing classroom monitoring methods fail to detect real-time emotional responses, making it difficult for educators to understand when students are struggling or disengaged. The project aims to develop a CNN-based facial expression recognition model capable of classifying images into three sentiment categories—Positive, Neutral, and Negative—to support engagement analysis.

3.3 Current Scenario / Limitations of Existing Approaches

Current classroom engagement monitoring relies heavily on:

- Manual observation by teachers, which is subjective and limited.
- Periodic assessments that do not reflect real-time emotional states.
- Traditional FER systems that classify seven detailed emotions, often too granular and unstable for classroom conditions.
- Systems that require high-quality images and struggle under variations in lighting, pose, and facial occlusion.

These limitations demonstrate the need for a robust, sentiment-level classification system tailored specifically for educational settings.

 Sister Nivedita University, WB

3.4 Need / Importance of Solving the Problem

Addressing this problem is crucial because:

- Student engagement directly impacts learning outcomes and academic performance.
- Early detection of negative or neutral sentiments helps teachers intervene in a timely manner.
- Automated sentiment recognition reduces teacher workload and improves classroom analytics.
- Educational institutions increasingly require data-driven methods to enhance teaching strategies.

A sentiment-based monitoring system supports both teachers and students by enabling a more responsive and personalized learning experience.

3.5 Scope and Boundaries of the Problem

The scope of this project is limited to:

- Sentiment classification using facial expressions (Positive, Neutral, Negative).
- Use of static images from the FER dataset.
- Implementation of a CNN model for sentiment classification.
- Performance evaluation using accuracy, loss curves, and confusion matrix.

The project does **not** include:

- Real-time deployment in classroom environments.
- Video processing, body posture analysis, or gaze tracking.
- Multi-modal behavioural analytics.

3.6 Objectives (High-Level)

The major objectives of the project are:

- To preprocess and map the FER dataset into three sentiment categories.
- To design and develop a CNN-based sentiment classification model.

- To evaluate the model using appropriate performance metrics.
- To create system-level diagrams including Flowchart, DFD Level 0, and System Architecture.
- To document the methodology clearly for academic and implementation reference.

3.7 Expected Challenges

Some anticipated challenges include:

- Handling class imbalance in the FER dataset.
- Ensuring the model remains robust to lighting variations and facial angles.
- Reducing overfitting through appropriate augmentation techniques.
- Achieving high accuracy with a lightweight model suitable for classroom applications.

Chapter 4

Proposed Methodology

4.1 Introduction to Methodology

A systematic methodology is essential to develop a robust sentiment classification system capable of supporting classroom engagement analysis. This chapter outlines the overall workflow of the proposed system, the architecture used, techniques applied for preprocessing and modelling, and the expected outcomes of the designed approach. The system follows a machine learning–based framework that integrates data preprocessing, CNN model development, evaluation, and system-level design components.

4.2 System Architecture / Conceptual Framework

The proposed system is designed as a pipeline consisting of four major stages: data preprocessing, model training, sentiment classification, and performance evaluation. The system architecture diagram provides a clear overview of how raw facial images are processed and transformed into sentiment predictions.

 Sister Nivedita University, WB

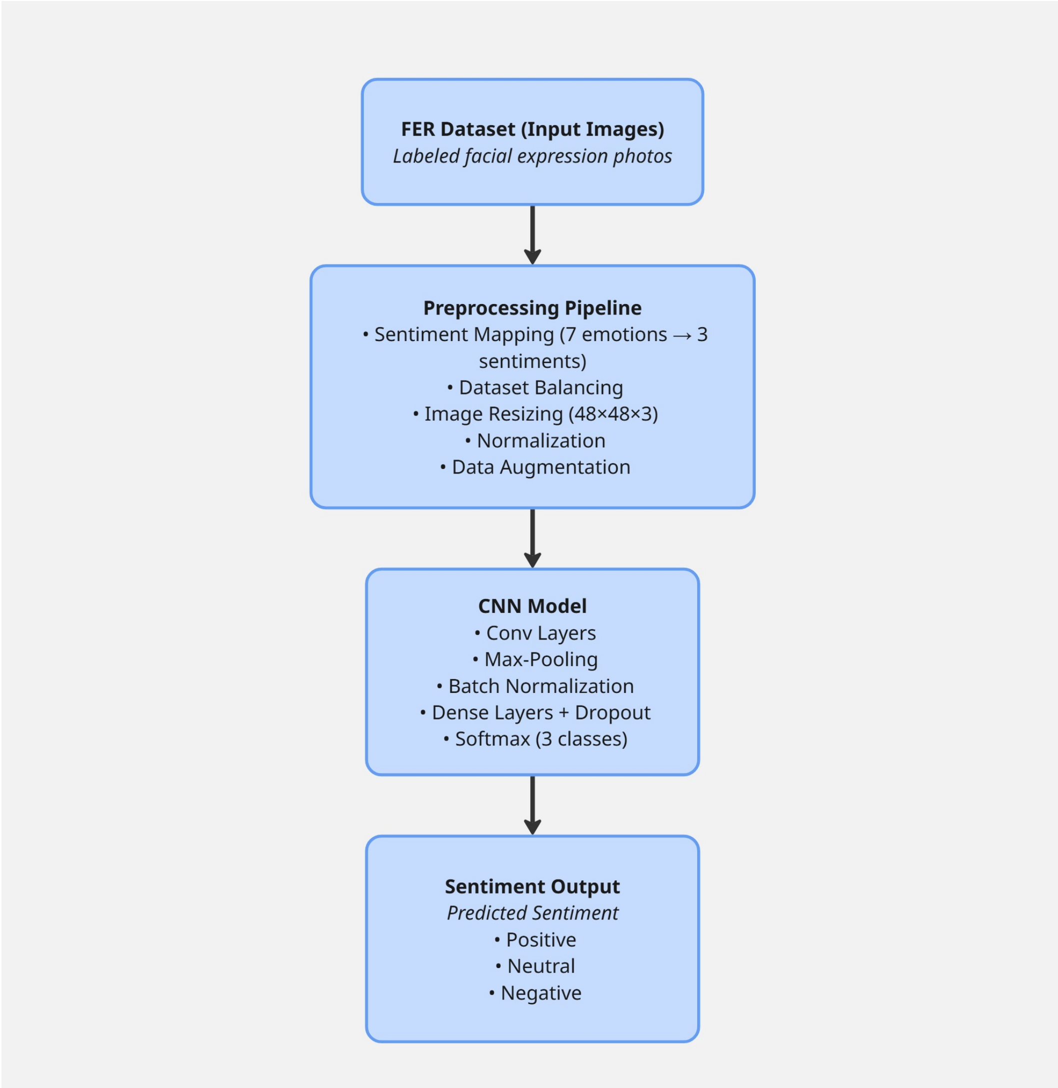

Figure 4.1: System Architecture of the Proposed Model

4.3 Algorithms / Techniques Used

The core algorithm used in this project is a Convolutional Neural Network (CNN), which is widely recognized for its high performance in image classification tasks. CNNs automatically learn spatial features through convolutional operations, making them suitable for facial expression analysis.

Justification for Algorithm Choice

CNNs were selected due to:

- Their superior capability to learn facial features from images.
- High robustness to spatial variations such as expression intensity or partial occlusion.

- Their proven success in multiple FER (Facial Expression Recognition) benchmarks.

The CNN architecture used in this project consists of multiple convolutional, pooling, dropout, and dense layers, designed to balance performance and generalization.

4.4 Data Handling and Processing

The model uses the FER dataset obtained from Kaggle. The following data preprocessing steps were performed:

- **Sentiment Mapping:** The seven original emotion classes (Angry, Disgust, Fear, Sad, Happy, Surprise, Neutral) are mapped into three sentiment labels:
  - Positive: Happy
  - Neutral: Neutral
  - Negative: Angry, Disgust, Fear, Sad, Surprise
- **Image Resizing:** All images are resized to $48 \times 48 \times 3$.
- **Normalization:** Pixel values are scaled to the range [0,1] to improve learning stability.
- **Dataset Balancing:** Undersampling is used to equalize the number of images per sentiment category.
- **Data Augmentation:** Techniques include random flipping, rotation, zoom, and contrast adjustments to increase dataset diversity and reduce overfitting.

4.5 Workflow of the Proposed System

The overall workflow of the proposed sentiment classification system is shown in the flowchart below:

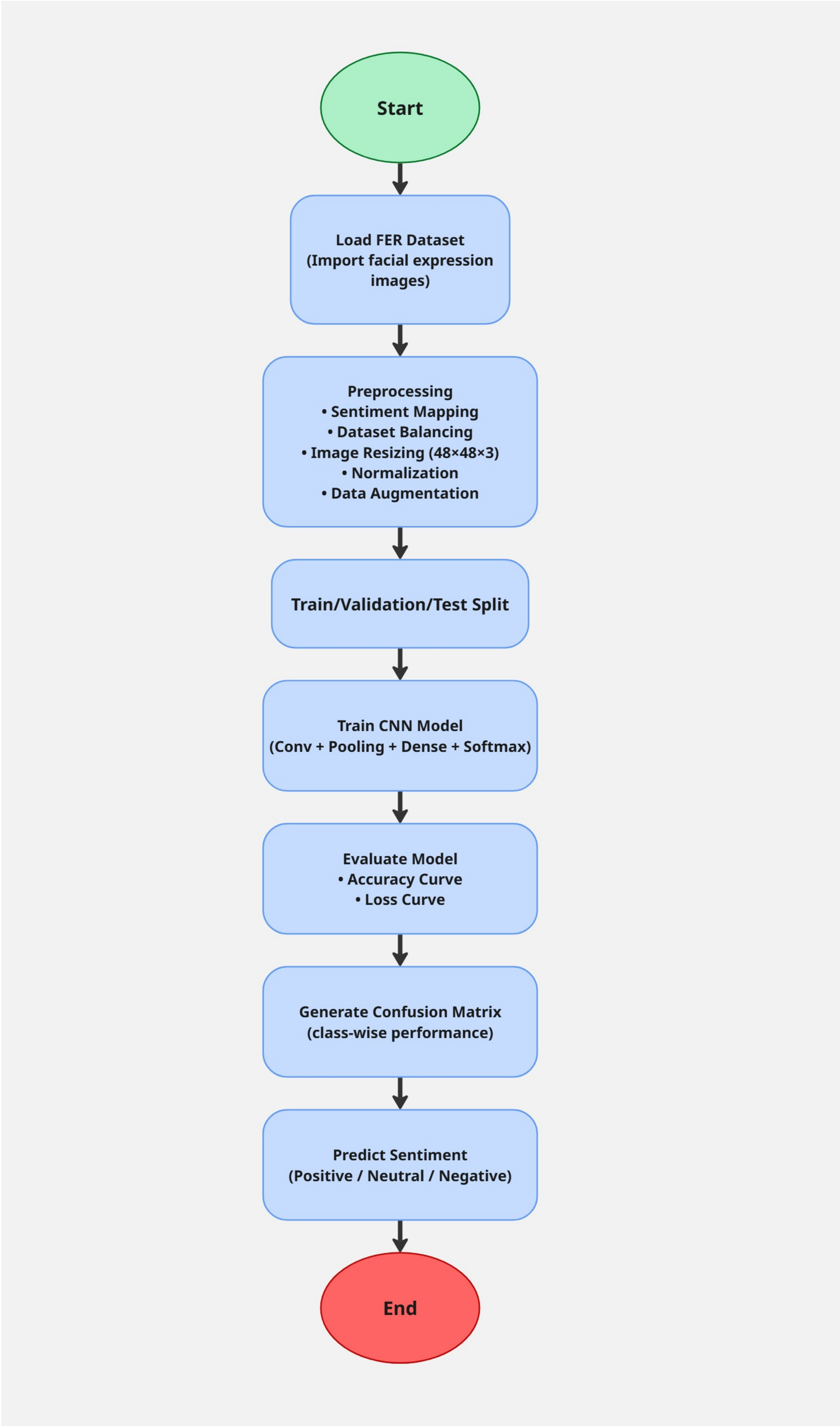

Figure 4.2: Flowchart of the Proposed System

The step-by-step workflow can be summarized as:

Input Image → Preprocessing → CNN Classification → Sentiment Output

4.6 Data Flow Diagram (DFD Level 0)

The DFD Level 0 diagram illustrates the high-level flow of data through the system components.

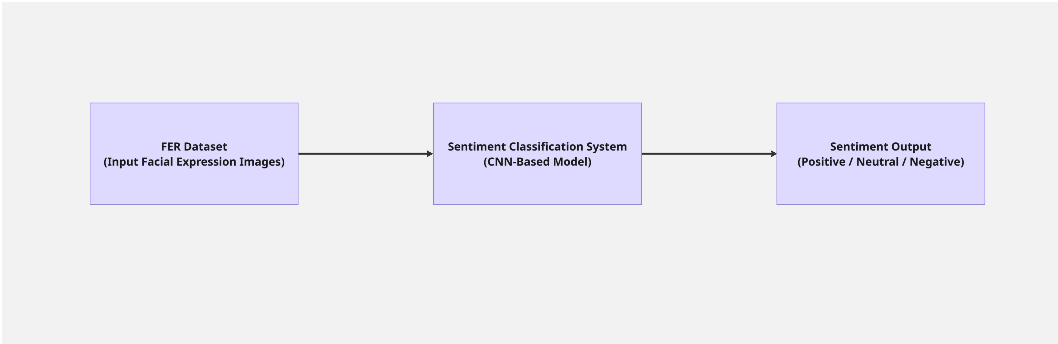

Figure 4.3: DFD Level 0 of the System

4.7 Experimental Evaluation

The model's performance is evaluated using accuracy, loss curves, and a confusion matrix. These metrics help determine how well the model classifies the three sentiment categories.

Confusion Matrix

The confusion matrix provides insight into class-wise prediction accuracy.

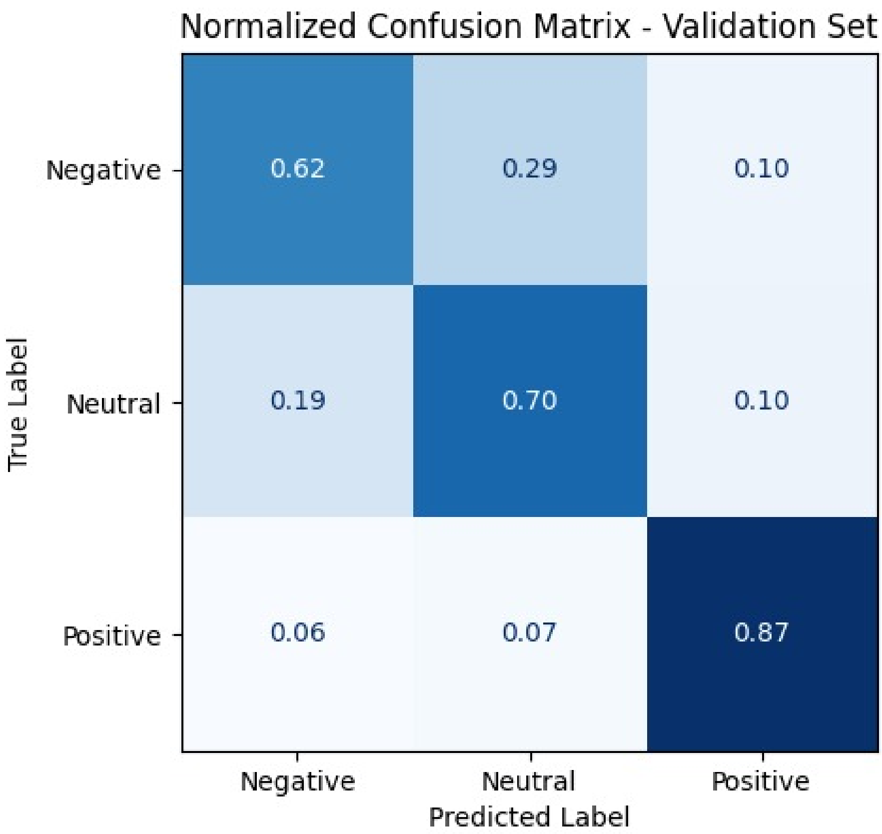

Figure 4.4: Confusion Matrix of the CNN Model

4.8 Expected Outcomes

The proposed methodology is expected to achieve the following:

- Accurate classification of facial images into Positive, Neutral, and Negative sentiments.
- Improved generalization through augmentation and proper preprocessing.
- Generation of interpretable metrics such as confusion matrix and accuracy curves.
- A modular pipeline that can later be extended to real-time classroom environments.

4.9 Novelty / Uniqueness of the Methodology

The uniqueness of the proposed methodology lies in its sentiment-level classification specifically tailored for classroom engagement analysis. Unlike traditional seven-emotion FER models, this system simplifies classification into three pedagogically relevant sentiment categories. The preprocessing pipeline, dataset balancing strategy, and optimized CNN architecture contribute to performance stability and suitability for educational applications.

Bibliography

[1] Goodfellow, I., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., & Bengio, Y. "Challenges in Representation Learning: A Report on Facial Expression Recognition." *arXiv:1307.0414*, 2013.

[2] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE*, 86(11), 2278–2324, 1998.

[3] Simonyan, K., & Zisserman, A. "Very Deep Convolutional Networks for Large-Scale Image Recognition." *arXiv:1409.1556*, 2014.

[4] He, K., Zhang, X., Ren, S., & Sun, J. "Deep Residual Learning for Image Recognition." In: *Proceedings of CVPR*, 2016.

[5] Barsoum, E., Zhang, C., Ferrer, C. C., & Zhang, Z. "Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution." In: *Proceedings of ICMI*, 2016.

[6] Whitehill, J., Serpell, Z., Lin, Y. C., Foster, A., & Movellan, J. R. "Toward Automatic Prediction of Student Engagement from Facial Expressions." *IEEE Transactions on Affective Computing*, 5(2), 201–216, 2014.

[7] Bosch, N., D'Mello, S., & Mills, C. "Automatic Detection of Learner Engagement Using Video-Based Facial Expression Recognition." *International Conference on Educational Data Mining*, 2015.

[8] Chollet, F. *Deep Learning with Python*. Manning Publications, 2018.

[9] Shorten, C., & Khoshgoftaar, T. M. "A Survey on Image Data Augmentation for Deep Learning." *Journal of Big Data*, 6(60), 2019.