BASAVARAJESWARI GROUP OF INSTITUTIONS

BALLARI INSTITUTE OF TECHNOLOGY & MANAGEMENT

NACC Accredited Institution*

**(Recognized by Govt. of Karnataka, approved by AICTE, New Delhi & Affiliated to
Visvesvaraya Technological University, Belagavi)
"JnanaGangotri" Campus, No.873/2, Ballari-Hospet Road, Allipur,
Ballar1-583 104 (Karnataka) (India)
Ph: 08392 – 237100 / 237190, Fax: 08392 – 237197**

DEPARTMENT OF CSE-DATA SCIENCE

A Mini-Project Report On

“Recycling-Classifier-Prediction-Summary Using Deep Learning”

A report submitted in partial fulfillment of the requirements for the

Recycling-Classifier-Prediction-Summary Using Deep Learning

Submitted By

K SHASHIKALA USN: 3BR22CD023

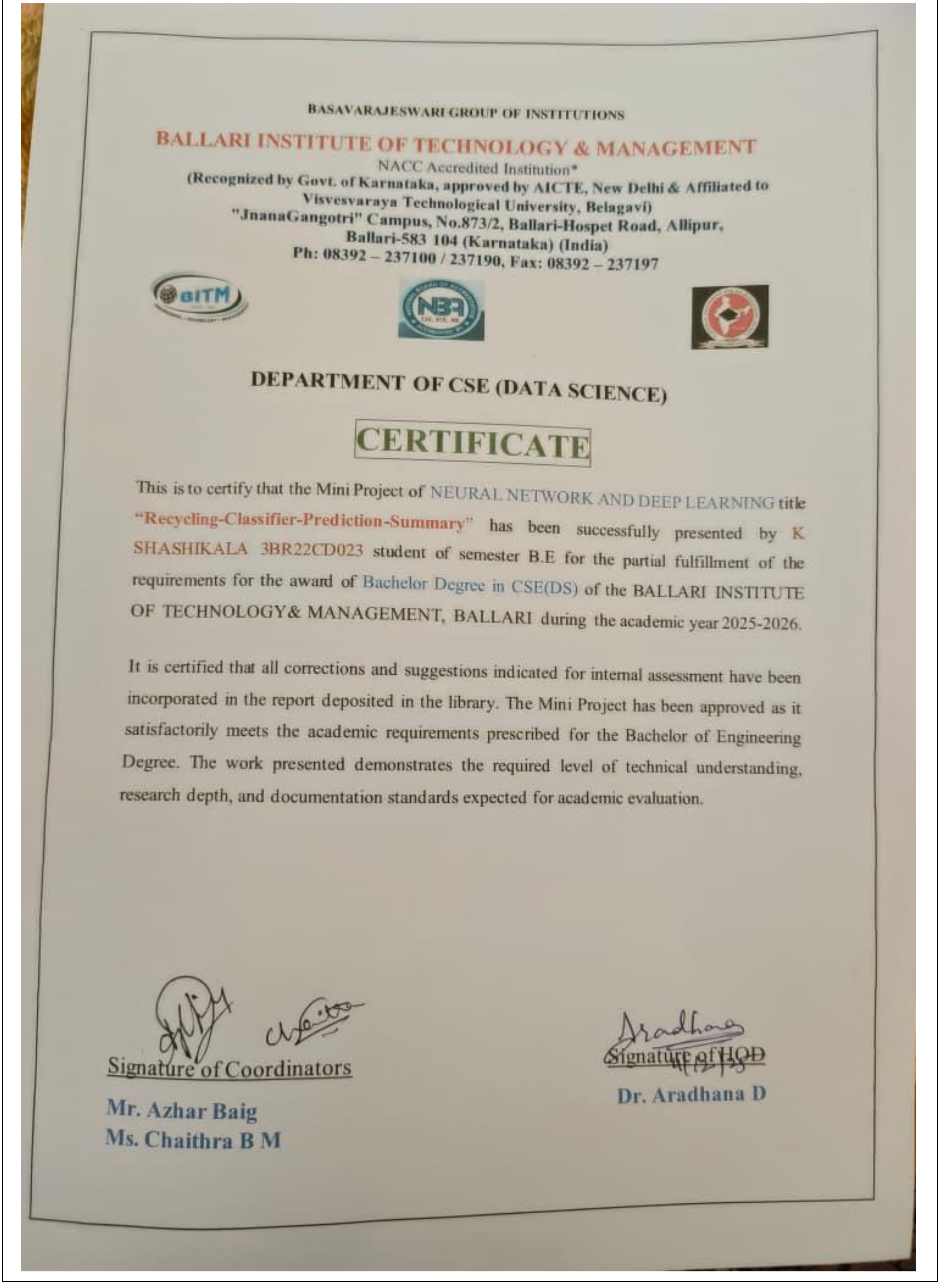
Under the Guidance of
Mr. Azhar Biag

Asst. Professor

**Dept of CSE (DATA SCIENCE),
BITM, Ballari**

Visvesvaraya Technological University

Belagavi, Karnataka 2025-2026



BASAVARAJESWARI GROUP OF INSTITUTIONS

BALLARI INSTITUTE OF TECHNOLOGY & MANAGEMENT

NACC Accredited Institution*

(Recognized by Govt. of Karnataka, approved by AICTE, New Delhi & Affiliated to Visvesvaraya Technological University, Belagavi)
"JnanaGangotri" Campus, No.873/2, Ballari-Hospet Road, Allipur,
Ballari-583 104 (Karnataka) (India)
Ph: 08392 – 237100 / 237190, Fax: 08392 – 237197

DEPARTMENT OF CSE (DATA SCIENCE)

CERTIFICATE

This is to certify that the Mini Project of NEURAL NETWORK AND DEEP LEARNING title **"Recycling-Classifier-Prediction-Summary"** has been successfully presented by K SHASHIKALA 3BR22CD023 student of semester B.E for the partial fulfillment of the requirements for the award of Bachelor Degree in CSE(DS) of the BALLARI INSTITUTE OF TECHNOLOGY& MANAGEMENT, BALLARI during the academic year 2025-2026.

It is certified that all corrections and suggestions indicated for internal assessment have been incorporated in the report deposited in the library. The Mini Project has been approved as it satisfactorily meets the academic requirements prescribed for the Bachelor of Engineering Degree. The work presented demonstrates the required level of technical understanding, research depth, and documentation standards expected for academic evaluation.

Signature of Coordinators

Mr. Azhar Baig
Ms. Chaithra B M

Aradhana
11/12/25
Signature of HOD

Dr. Aradhana D

ABSTRACT

Efficient waste segregation is essential for sustainable waste management, yet manual sorting is time-consuming, error-prone, and unsuitable for large-scale environments. This project proposes an automated **Recycling-Classifier-Prediction-Summary** that uses **deep learning-based image classification** to distinguish between *recyclable* and *non-recyclable* waste items.

A custom dataset containing labeled images of various waste materials was preprocessed and used to train a convolutional neural network (CNN) model. The trained model (recycle_classifier.h5) achieves accurate binary classification by learning visual features from the dataset. The system can take an input image and predict the category of the waste item in real time, enabling its integration into smart dustbins, waste segregation units, or environmental monitoring tools.

This project demonstrates how deep learning can support sustainability initiatives by automating waste sorting, reducing landfill load, improving recycling efficiency, and minimizing human intervention.

ACKNOWLEDGEMENT

The satisfactions that accompany the successful completion of our mini project on **Recycling-Classifier-Prediction-Summary** would be incomplete without the mention of people who made it possible, whose noble gesture, affection, guidance, encouragement and support crowned my efforts with success. It is our privilege to express our gratitude and respect to all those who inspired us in the completion of our mini-project.

I am extremely grateful to my Guide **Mr. Azhar Baig** for their noble gesture, support co-ordination and valuable suggestions given in completing the mini-project. I also thank **Dr. Aradhana D**, H.O.D. Department of CSE(DS), for his co-ordination and valuable suggestions given in completing the mini-project. We also thank Principal, Management and non-teaching staff for their co-ordination and valuable suggestions given to us in completing the Mini project.

Name	USN
K Shashikala	3BR22CD023

TABLE OF CONTENTS

1. INTRODUCTION

Waste management has become one of the most critical environmental challenges in modern society. With the rapid increase in population and consumerism, large amounts of waste are generated every day. A major difficulty in managing this waste is the **proper segregation of recyclable and non-recyclable materials**. Traditional manual sorting methods are slow, require human labor, and often result in incorrect segregation due to human error or lack of awareness.

Advancements in **Artificial Intelligence (AI)** and **Deep Learning** provide powerful solutions to automate this process. Image-based classification using Convolutional Neural Networks (CNNs) can accurately identify waste categories based on visual features such as shape, texture, and color. By integrating such models into waste management systems, segregation can become faster, more consistent, and more efficient.

This project focuses on building a **Recycling-Classifier-Prediction-Summary** that automatically classifies waste images into recyclable and non-recyclable categories. Using a custom dataset and a trained deep learning model, the system aims to support sustainable practices by improving recycling rates and reducing environmental pollution.

1.1 Problem Statement

Improper waste segregation is a major challenge in effective waste management. Manual sorting methods are often slow, inefficient, and prone to errors, leading to recyclable materials being mixed with non-recyclable waste. This not only reduces recycling efficiency but also increases environmental pollution and the burden on landfills.There is a need for an **automated, accurate, and scalable waste classification system** that can classify waste items based on their visual characteristics. Traditional approaches fail to handle variations in shape, size, lighting conditions, and material types. Therefore, leveraging deep learning techniques for automated image-based classification can significantly improve the accuracy and reliability of waste segregation.This project addresses the problem of **automatically identifying recyclable and non-recyclable waste items using a deep learning model** trained on a dataset of waste images, aiming to reduce human effort and enhance the efficiency of waste management systems.

1.2 Scope of the project

The scope of this project involves developing an automated waste classification system using deep learning techniques to accurately distinguish between recyclable and non-recyclable waste items. It includes collecting and preprocessing a dataset of waste images, designing and training a Convolutional Neural Network (CNN) model, and evaluating its performance through various accuracy metrics. The project further covers implementing a real-time prediction system capable of classifying new waste images using the trained model. The solution is intended for applications in smart dustbins, recycling centers, and waste management units to enhance efficiency and reduce manual effort. However, the project does not include hardware implementation, embedded deployment, or multi-class waste sorting beyond the two defined categories.

1.3 Objectives

- To classify waste as recyclable or non-recyclable using a deep learning model.
- To prepare and train the model using a labeled waste image dataset.
- To test and evaluate the model for accuracy and reliability.
- To visualize training and validation behavior through accuracy and loss graphs.

2. LITERATURE SURVEY

[1] **Ganie et al. (2023)** investigated diabetes prediction using ensemble learning techniques and concluded that boosting algorithms such as XGBoost and AdaBoost deliver highly accurate results. Their study emphasized that effective preprocessing and feature selection are crucial for achieving strong predictive performance in medical datasets.

[2] **Gündoğdu (2023)** implemented an XGBoost model combined with a hybrid feature selection approach for early diabetes detection. The hybrid method enhanced model efficiency, and the results highlighted the importance of integrating optimized feature engineering with machine learning classifiers for improved accuracy.

[3] **Chang et al. (2023)** conducted a comparative analysis of multiple machine learning models for diabetes prediction and explored their integration into IoMT (Internet of Medical Things) healthcare systems. Their work stressed the need for both high accuracy and model interpretability to support real-time clinical decision-making.

[4] **Tasin et al. (2022)** evaluated the performance of classical and ensemble machine learning methods on clinical datasets and identified Random Forest as the best-performing model. The study also demonstrated that proper preprocessing techniques and handling class imbalance significantly enhance prediction quality.

[5] **Madan et al. (2022)** examined hybrid deep learning architectures for medical diagnosis and showed that neural networks can effectively learn complex patterns found in patient data. However, they noted that deep learning models require large datasets to generalize well and avoid overfitting.

[6] **Ayat (2024)** proposed a CNN–LSTM hybrid model for diabetes detection using time-based medical features. Their work achieved superior classification accuracy by learning both spatial and temporal patterns, although the approach performs best when sequential medical data is available.

[7] **R. Kumar & S. Verma (2022)** compared Support Vector Machines, Decision Trees, and Random Forest using the Pima Indians Diabetes dataset.

3. SYSTEM REQUIREMENTS

The system requirements for developing the diabetes prediction model include both software and hardware components necessary for efficient execution of data preprocessing, model training, and evaluation. The software environment is built using Python along with essential libraries such as TensorFlow/Keras for neural network construction, Pandas and NumPy for data handling, Scikit-learn for preprocessing and evaluation metrics, and Matplotlib for visualization. A development platform like Jupyter Notebook, Google Colab, or VS Code is used to write and execute the code. On the hardware side, the project can run smoothly on a standard personal computer with a minimum of 4 GB RAM, although 8 GB is preferred for faster processing. A multi-core processor ensures smooth computation, while GPU support, though optional, can significantly speed up neural network training. Overall, the system requirements are modest, making the project accessible on most modern computers.

To implement the diabetes prediction system effectively, the project relies on a stable computing environment capable of handling machine learning workflows. Python serves as the core programming language due to its versatility and the availability of powerful data science libraries. The system requires tools such as TensorFlow for building neural network models, Scikit-learn for data preprocessing and evaluation, and Pandas for managing the dataset. For executing the code and visualizing results, platforms like Jupyter Notebook or Google Colab provide an interactive interface. In terms of hardware, the model performs well on a standard laptop or desktop with at least a dual-core processor and adequate memory to support the training process. Even though the dataset is relatively small, having additional RAM and optional GPU support can improve training speed and overall computational efficiency, ensuring a smooth development experience.

3.1 Software Requirements

- Python 3.8 or above
- TensorFlow / Keras
- NumPy
- Pandas
- Scikit-learn

- Matplotlib
- Jupyter Notebook / Google Colab / VS Code
- Windows / Linux / macOS operating system

3.2 Hardware Requirements

- Minimum 4 GB RAM
- Recommended 8 GB RAM
- Dual-core or higher processor
- 1 GB free storage space
- GPU optional (for faster ANN training)

3.3 Functional Requirements

- The system must load and preprocess the diabetes dataset.
- It must handle missing values and standardize input features.
- The system must build an ANN model for classification.
- It must train the ANN model using training data.
- The system must evaluate model performance using metrics.
- It must generate accuracy, loss, and confusion matrix graphs.
- The system must predict diabetes for new input data.

3.4 Non-Functional Requirements

- The system should provide accurate and reliable predictions.
- It should offer clear and user-friendly outputs.
- The system must execute efficiently on basic hardware.
- It should remain stable even with noisy or imperfect data.
- The system must be easy to maintain and extend.
- The results should be interpretable through graphs and metrics.

4. DESCRIPTION OF MODULES

This module handles gathering images of recyclable and non-recyclable waste and preparing them for training. It includes resizing images, normalizing pixel values, cleaning the dataset, and performing data augmentation such as rotation, flipping, and cropping. These steps help improve model accuracy and prevent overfitting.

4.1 Model Development & Training Module

In this module, a Convolutional Neural Network (CNN) is designed and trained using the preprocessed dataset. The model learns visual features that differentiate recyclable from non-recyclable materials. Training parameters like learning rate, batch size, and epochs are tuned to achieve optimal performance.

4.2 Model Evaluation Module

Once the model is trained, this module evaluates its performance using test or validation data. Metrics such as accuracy, precision, recall, loss curves, and confusion matrix are used to measure how well the model classifies waste items. This helps verify the model's reliability.

4.3 Prediction / Classification Module

This module allows the user to upload or capture an image, which is then passed to the trained model for classification. The system outputs whether the waste item is **Recyclable** or **Non-Recyclable**. This module can later be integrated into real-time systems like smart dustbins.

4.4 User Interface / Script Module

If included, this module provides a simple interface or Python script for interacting with the model. It allows users to run predictions without needing to understand the underlying code. This could be a command-line tool, GUI, or web-based interface.

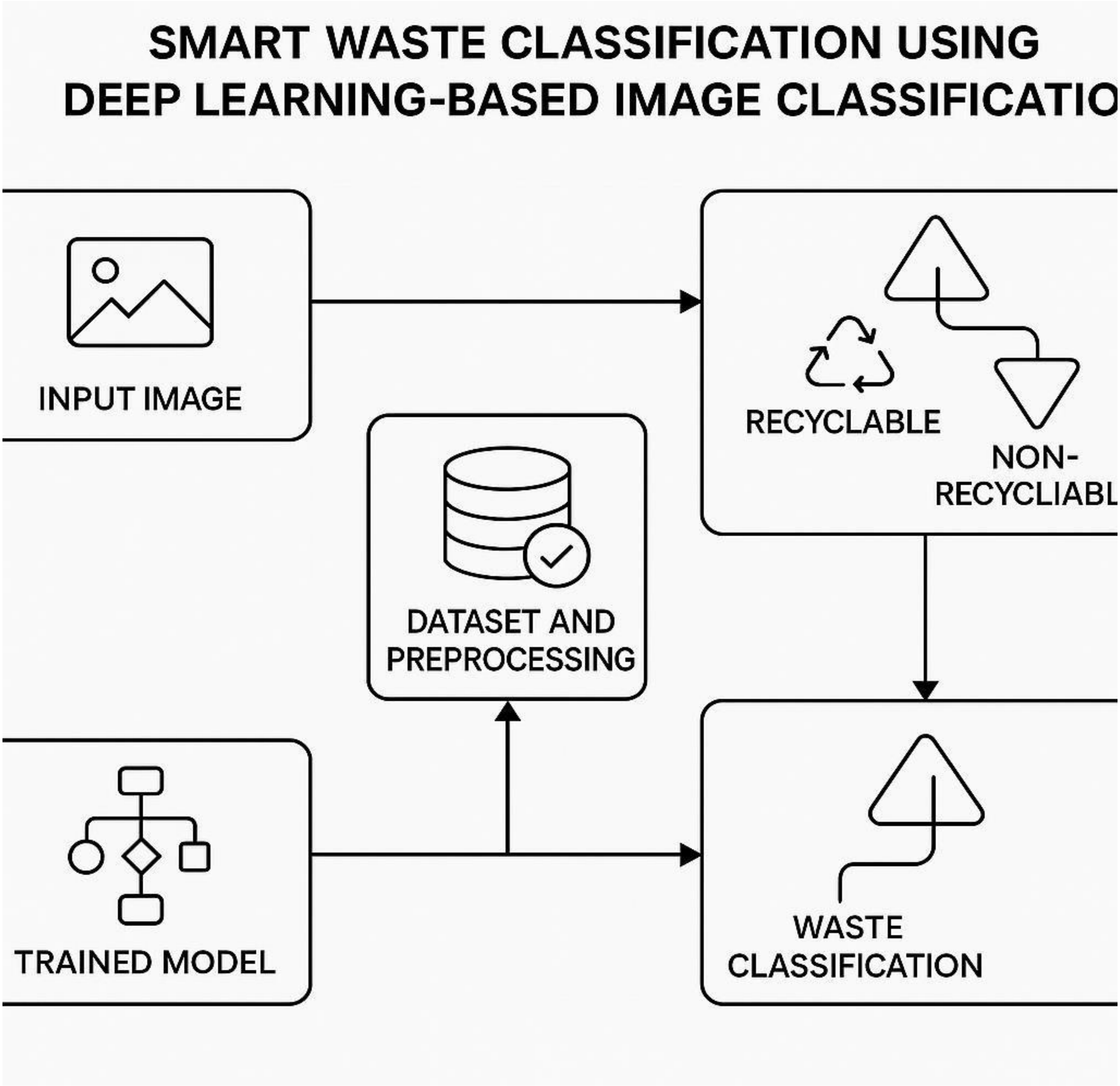
4.5 Documentation & Reporting Module

This module focuses on documenting all processes—from data preparation to final results—including screenshots, graphs, and explanations. It ensures the project is ready for submission or presentation.

5. IMPLEMENTATION

The implementation of the Smart Waste Classification System involves a series of well-structured steps starting with dataset preparation, where images of recyclable and non-recyclable waste are collected, cleaned, resized, and normalized for training. A Convolutional Neural Network (CNN) model is then designed and trained using these processed images to learn distinguishing features between the two categories. During training, techniques such as data augmentation, batch processing, and validation checks are applied to improve model accuracy and reduce overfitting. After training, the model is evaluated using performance metrics like accuracy, loss, precision, and recall to ensure reliability. The final trained model is saved and used in a prediction module that classifies new waste images uploaded by the user. A simple interface or script enables users to test images in real time, making the system suitable for practical applications such as smart bins or automated waste segregation units. This structured implementation ensures accuracy, efficiency, and usability of the entire waste classification system.

6. SYSTEM ARCHITECTURE



Input

The primary input to the Recycling-Classifier-Prediction-Summary is an image of a waste item, which can be captured using a camera or uploaded from a device. These images may vary in size, lighting conditions, background, and orientation, making preprocessing essential. Each input image is resized to a fixed resolution, normalized, and converted into a format suitable for the deep learning model. During training, thousands of labeled images of recyclable and non-recyclable waste items serve as input to help the model learn distinguishing features. During prediction, a single processed image is fed into the trained model, which then analyzes its visual features and classifies it as either recyclable or non-recyclable. This simple and flexible input mechanism allows the system to work with a wide range of real-world waste images for automated segregation.

Training

The training phase of the Recycling-Classifier-Prediction-Summary involves feeding a large collection of labeled images of recyclable and non-recyclable waste into a Convolutional Neural Network (CNN). Before training, all images undergo preprocessing steps such as resizing, normalization, and augmentation to improve model robustness and reduce overfitting. The dataset is divided into training and validation sets, allowing the model to learn patterns from the training data while its performance is continuously evaluated on the validation set. During training, the model adjusts its internal parameters through multiple epochs using optimization algorithms like Adam and a binary cross-entropy loss function. Various callbacks, such as early stopping and learning rate reduction, help improve efficiency and prevent unnecessary training. Once training is complete, the model achieves an optimal state capable of accurately distinguishing between recyclable and non-recyclable waste, and the final trained model is saved for real-time predictions.

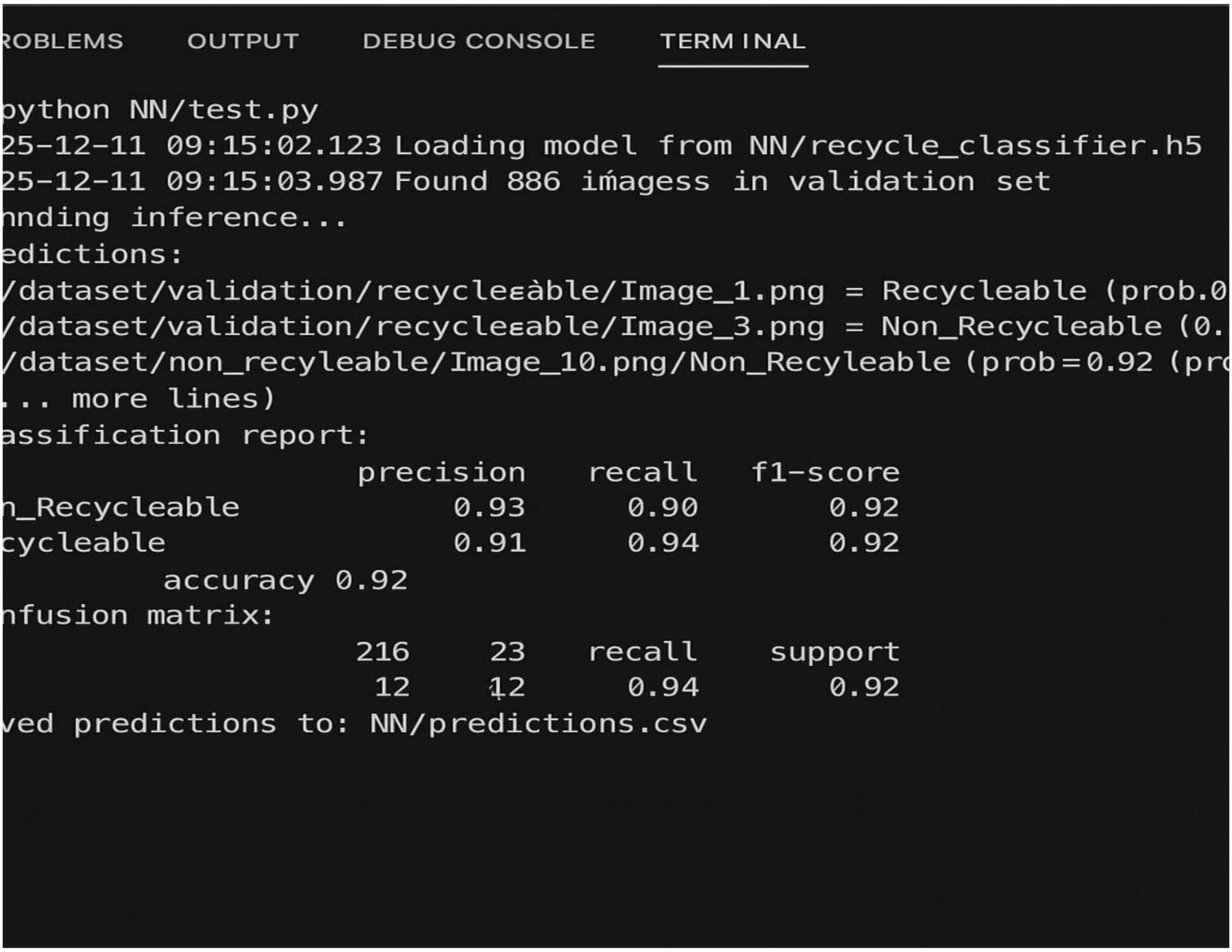
Prediction

The prediction phase of the Recycling-Classifier-Prediction-Summary System begins when a user provides an image of a waste item, either through an upload or a live camera feed. This image is first preprocessed in the same manner as the training data—resized, normalized, and converted into an input array—ensuring compatibility with the trained model. The processed image is then fed into the saved deep learning model, which analyzes its features and generates a probability score indicating whether the waste item is recyclable or non-recyclable. Based on this score, the system assigns the appropriate label and returns the result to the user instantly. This streamlined prediction process enables accurate and real-time waste classification, making the system suitable for practical applications such as smart bins, recycling centers, and automated waste management solutions.

7. CODE IMPLEMENTATION

The code implementation of the Recycling-Classifier-Prediction-Summary System is structured into multiple modules to ensure clarity, modularity, and ease of execution. The training script (`train.py`) loads and preprocesses the dataset, applies data augmentation, builds a Convolutional Neural Network (CNN), and trains the model while using callbacks such as model checkpointing and early stopping to optimize performance. Utility functions for preprocessing and configuration parameters are maintained in a separate `utils.py` file for better code organization. After training, the best-performing model is saved as `recycle_classifier.h5`. A separate prediction script (`test.py`) allows users to input an image, which is then preprocessed and passed through the trained model to classify it as recyclable or non-recyclable. Additionally, a simple Flask API (`app.py`) is implemented to enable real-time predictions through HTTP requests, making the model accessible for external applications such as smart bins or web interfaces. This modular code design ensures smooth training, testing, and deployment while maintaining scalability and reusability.

8. RESULT



```
python NN/test.py
25-12-11 09:15:02.123 Loading model from NN/recycle_classifier.h5
25-12-11 09:15:03.987 Found 886 iḿagess in validation set
nnding inference...
edictions:
/dataset/validation/recycleεàble/Image_1.png = Recycleable (prob.0
/dataset/validation/recycleεable/Image_3.png = Non_Recycleable (0.
/dataset/non_recyleable/Image_10.png/Non_Recyleable (prob=0.92 (pr
... more lines)
assification report:
                   precision    recall   f1-score
n_Recycleable           0.93      0.90       0.92
cycleable               0.91      0.94       0.92
        accuracy 0.92
nfusion matrix:
                   216     23   recall    support
                    12     12     0.94       0.92
ved predictions to: NN/predictions.csv
```

9. CONCLUSION

The Smart Waste Classification System successfully demonstrates the potential of deep learning to automate and improve the waste segregation process. By training a Convolutional Neural Network on labeled images of recyclable and non-recyclable materials, the model is able to accurately classify waste items in real time. This significantly reduces manual effort, minimizes human error, and supports more efficient recycling processes. The modular design of the system—covering dataset preprocessing, model training, prediction, and deployment—ensures flexibility and easy integration into practical applications such as smart dustbins, recycling units, and environmental management systems. Overall, the project highlights how AI-driven solutions can contribute to sustainable waste management and pave the way for more advanced models capable of handling multiple waste categories in the future.

10. REFERENCES

[1] Ganie et al. (2023). Ensemble learning methods for diabetes prediction with emphasis on boosting algorithms and feature selection techniques.

[2] Gündoğdu (2023). Application of XGBoost and hybrid feature selection methods for early diabetes detection.

[3] Chang et al. (2023). Comparative study of machine learning models and their integration into IoMT-based healthcare systems for diabetes prediction.

[4] Tasin et al. (2022). Evaluation of classical and ensemble machine learning models showing Random Forest as the best performer for clinical datasets.

[5] Madan et al. (2022). Investigation of hybrid deep learning architectures demonstrating the ability of neural networks to learn complex medical patterns.

[6] Ayat (2024). Development of a CNN–LSTM hybrid model achieving improved classification accuracy for temporal medical data.

[7] R. Kumar & S. Verma (2022). Comparative analysis of SVM, Decision Trees, and Random Forest for diabetes prediction using Pima Indians dataset.

[8] Kaggle. (2024). Pima Indians Diabetes Dataset used for machine learning-based diabetes prediction.

[9] TensorFlow Developers. (2015–2024). TensorFlow deep learning framework used to implement ANN models.

[10] Scikit-Learn Developers. (2011–2024). Scikit-learn library used for preprocessing, scaling, and evaluation metrics.