Project Report Format

Date	15 February 2025
Team ID	LTVIP2025TMID42449
Project Name	Butterfly Species Classification System
Maximum Marks	10 Marks

1 Introduction

1.1 Project Overview

The Butterfly Species Classification System is a web-based application using transfer learning with MobileNetV2 to classify 6499 images across 75 butterfly species, providing real-time predictions and educational facts via a Streamlit interface to support biodiversity monitoring, ecological research, and citizen science.

1.2 Purpose

The project aims to automate butterfly species identification, reducing manual effort and enabling accessible, accurate tools for researchers, ecologists, and enthusiasts.

2 Ideation Phase

2.1 Problem Statement

Accurate identification of butterfly species is challenging due to species diversity, manual processes, and lack of real-time tools, hindering biodiversity and research efforts.

2.2 Empathy Map Canvas

- **Who**: Researchers, ecologists, citizen scientists, students.
- **Says**: “I need a quick way to identify butterflies.”
- **Thinks**: “Manual guides are slow and error-prone.”
- **Does**: Collects images, uses field guides, searches online.
- **Feels**: Frustrated by delays, excited about conservation.
- **Pains**: Limited access to automated tools.
- **Gains**: Accurate identification, educational insights.

2.3 Brainstorming

Key ideas include CNN-based classification, Streamlit app, and educational content, prioritized for feasibility and impact on biodiversity and education.

3 Requirement Analysis

3.1 Customer Journey Map

[Placeholder: Discovery Image Upload Prediction Fact Review Sharing Results]

3.2 Solution Requirement

- **Functional**: Image classification, preprocessing, educational content delivery.
- **Non-functional**: 95% accuracy, 2-second response time, intuitive UI.

3.3 Data Flow Diagram

DFD Level 0: Image upload preprocessing MobileNetV2 inference SQLite query output display.

3.4 Technology Stack

- **Frontend**: Streamlit, Custom CSS.
- **Backend**: Flask, TensorFlow/Keras.
- **Database**: SQLite.
- **Infrastructure**: Streamlit Cloud, Google Colab.

4 Project Design

4.1 Problem Solution Fit

Addresses manual identification challenges with an automated, user-friendly classification system.

4.2 Proposed Solution

A Streamlit app with MobileNetV2 for classifying 75 butterfly species, integrated with educational facts.

4.3 Solution Architecture

Micro-services architecture with Streamlit frontend, Flask backend, MobileNetV2 model, and SQLite database, deployed on Streamlit Cloud.

5 Project Planning & Scheduling

5.1 Project Planning

Three 5-day sprints:

- **Sprint 1**: Data preprocessing, model training (10 story points).
- **Sprint 2**: Model testing, classification (8 story points).
- **Sprint 3**: Application development (5 story points).

Velocity: 7.67 story points per sprint.

6 Functional and Performance Testing

6.1 Performance Testing

- **Unit Testing**: TensorFlow for model, Pytest for backend.
- **Integration Testing**: Test API with Postman.
- **Performance**: Response time <2 seconds, 95% accuracy.

7 Results

7.1 Output Screenshots

[Placeholder for screenshots of Streamlit UI, predictions, and facts.]

8 Advantages & Disadvantages

- **Advantages**: Automates identification, supports biodiversity, user-friendly.
- **Disadvantages**: Limited dataset diversity, struggles with blurry images.

9 Conclusion

The system successfully automates butterfly species classification, enhancing research and education with a scalable, accurate solution.

10 Future Scope

- Expand dataset for more species.
- Integrate advanced CNNs (e.g., EfficientNet).
- Develop mobile app for field use.

11 Appendix

- **Source Code**: [GitHub repository link]
- **Dataset Link**: [Kaggle dataset link]
- **Project Demo Link**: [Streamlit Cloud link]