Project Report

# 1. INTRODUCTION

## 1.1 Project Overview

This project focuses on classifying butterfly species using deep learning, specifically through transfer learning with the VGG16 model. It leverages a dataset consisting of 75 classes with 6499 butterfly images. The goal is to build a robust classification system capable of real-time predictions, aiding biodiversity research, conservation efforts, and public education.

## 1.2 Purpose

The primary objective is to implement a machine learning model that can automatically classify butterfly species from images, thereby simplifying ecological surveys, enabling citizen science contributions, and facilitating rapid biodiversity assessments.

# 2. IDEATION PHASE

## 2.1 Problem Statement

Manual classification of butterfly species is tedious and error-prone. Researchers and conservationists need an automated, accurate solution for identifying species efficiently in the field.

## 2.2 Empathy Map Canvas

The target users include ecologists, students, and nature enthusiasts. They seek ease of use, reliability, and visual support for accurate species identification.

## 2.3 Brainstorming

- Automate classification using transfer learning  
- Integrate model with web interface  
- Enable use by students and researchers  
- Keep system scalable for more classes

# 3. REQUIREMENT ANALYSIS

## 3.1 Customer Journey Map

User uploads image → Model processes image → Species name predicted → Result shown in UI

## 3.2 Solution Requirement

- Annotated image dataset  
- Pretrained VGG16 model  
- Python with TensorFlow and Flask  
- GitHub for version control

## 3.3 Data Flow Diagram

A basic image → preprocessing → VGG16 model → prediction → displayed on Flask interface

## 3.4 Technology Stack

Python, TensorFlow, Keras, Pandas, Seaborn, Flask, HTML/CSS, GitHub

# 4. PROJECT DESIGN

## 4.1 Problem Solution Fit

The solution automates the complex and time-consuming task of species identification with a pre-trained deep learning model.

## 4.2 Proposed Solution

Use VGG16 transfer learning to classify butterfly images and deploy it using Flask.

## 4.3 Solution Architecture

Input Image → VGG16 Model → Dense Output Layer → Predicted Class → Flask UI Output

# 5. PROJECT PLANNING & SCHEDULING

## 5.1 Project Planning

- Week 1: Dataset setup and cleaning  
- Week 2: Model training with VGG16  
- Week 3: Flask integration  
- Week 4: Testing and documentation

# 6. FUNCTIONAL AND PERFORMANCE TESTING

## 6.1 Performance Testing

Accuracy evaluated using training/validation split. Confusion matrix, precision, recall, and F1-score used for evaluation.

# 7. RESULTS

## 7.1 Output Screenshots

Flask UI screenshots, prediction outputs, and terminal logs are attached. The model predicts species based on uploaded butterfly images with high accuracy.

# 8. ADVANTAGES & DISADVANTAGES

Advantages:  
- Fast and reliable  
- Lightweight and scalable  
- Easy web-based UI  
  
Disadvantages:  
- Requires clean input images  
- Relies on pretrained class accuracy

# 9. CONCLUSION

This project achieved its goal of creating an automated butterfly classifier using deep learning. The VGG16 model provided high accuracy, and the Flask app ensured user-friendly deployment.

# 10. FUTURE SCOPE

- Add real-time camera integration  
- Expand dataset to include more species  
- Convert to Android/iOS app  
- Improve model with custom CNNs

# 11. APPENDIX

Source Code: https://github.com/sravan0937/butterfly-classifier-sravs  
Dataset: https://www.kaggle.com/datasets/gpiosenka/butterfly-images40-species  
Model File: butterfly\_model\_sravs.h5  
Demo Video: (Optional)