PicLingo: AI-Powered Image Caption Generator

Institute of Engineering & Technology

Department of Computer Science and Engineering

B.Tech Final Year Project Report

Project Title: PicLingo – AI-Powered Image Caption Generator

Submitted by: Tabish Javed

Roll No: <Your Roll Number>

Under the guidance of: Mr. Avadesh Yadav

Submission Date: July 2025

# Certificate

This is to certify that the project titled 'PicLingo: AI-Powered Image Caption Generator' has been completed by Tabish Javed under my guidance in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering.

Signature of Guide  
  
Mr. Avadesh Yadav  
Project Supervisor

# Acknowledgment

I would like to express my sincere gratitude to my guide Mr. Avadesh Yadav for his continuous support and guidance throughout the development of this project. I would also like to thank my family, friends, and faculty members for their encouragement and help.

# Abstract

PicLingo is an AI-powered image caption generator designed to bridge the gap between visual data and natural language. The system uses OpenAI’s CLIP model to extract features from uploaded images and retrieve the most relevant caption from a predefined set. Built with React.js, Node.js, and Python, the project demonstrates the effective integration of full-stack development with AI capabilities.

Key features include a responsive frontend, efficient backend processing, and integration of a state-of-the-art vision-language model. PicLingo has potential applications in accessibility, content automation, and e-commerce product tagging.

# Introduction

## Introduction to Image Captioning

Image captioning is a computer vision task that involves generating a textual description of an image. It requires understanding of visual elements and translating them into natural language. Advances in AI have enabled models to comprehend image semantics and provide relevant captions.

## Motivation Behind PicLingo

The primary motivation was to create an assistive tool for content accessibility and automation. Many applications rely on captions for SEO, accessibility, and user interaction, but generating these manually is time-consuming. PicLingo simplifies this using AI.

## Objectives

- Build a user-friendly web interface for uploading images  
- Integrate OpenAI’s CLIP model for zero-shot image captioning  
- Ensure backend handles model inference efficiently  
- Deploy a working prototype for public demonstration

## Problem Statement

Manual captioning of images is not scalable in today’s fast-paced content-driven world. There’s a need for a solution that can automatically generate accurate, relevant, and context-aware captions for diverse visual content.

## Scope

PicLingo serves as a foundational tool for image captioning with potential for scalability across industries such as e-commerce, education, and accessibility platforms.