# IMAGE CAPTION GENERATOR USING DEEP LEARNING

*A project report submitted in partial fulfillment of the requirements for*

*the award of the degree of*

**Bachelor of Technology**

*by*


| | | |
|---|---|---|
| **Sufiyan Ali Khan** | **:** | **2111CS020573** |
| **G Swetha** | **:** | **2111CS020584** |
| **G Thejashwini** | **:** | **2111CS020589** |
| **H Varshith** | **:** | **2111CS020621** |
| **M Varun Teja** | **:** | **2111CS020626** |



Under the guidance of

**Prof. A. Siva Kumar**

Professor


**DEPARTMENT OF AI & ML**

**MALLA REDDY UNIVERSITY**

**(As per Telangana State Private Universities Act No.13 of 2020 and G.O.Ms.No.14, Higher Education (UE) Department)**

**HYDERABAD – 500043**

**TELANAGANA**

**INDIA**


**2023-24**

**MALLA REDDY UNIVERSITY**

**(As per Telangana State Private Universities Act No.13 of 2020 and G.O.Ms.No.14, Higher Education (UE) Department)**

**HYDERABAD – 500043**

**TELANAGANA**

# *Certificate*

This is to certify that this is the bonafide record of the application development entitled,"IMAGE CAPTION GENERATOR USING DEEP LEARNING" submitted by **Sufiyan Ali Khan(2111CS020573), G Swetha(2111CS020584), G Thejashwini (2111CS020589), H Varshith(2111CS020621), M VarunTeja(2111CS020626)** of B.Tech lII year II$^{nd}$ semester, Department of CSE (AI&ML) during the year 2023- 24.The results embodied in thereport have not been submitted to any other university or institute for the award of any degree or diploma.

**INTERNAL GUIDE**
**A Siva Kumar**
**Professor**

**HEAD OF THE DEPARTMENT**
**Dr. Thayaba Khatoon**
**Professor & HoD**

# ACKNOWLEDGEMENT

We would like to express our gratitude to all those who extended their support and suggestions to come up with this application. Special Thanks to our Guide A Siva Kumar whose help and stimulating suggestions and encouragement helped us all time in the due course of project development.

We sincerely thank our HOD Dr. Thayyaba Khatoon for her constant support and motivation all the time. A special acknowledgement goes to a friend who enthused us from the back stage. Last but not the least our sincere appreciation goes to our family who has been tolerant understanding our moods, and extending timely support.

# Abstract

Image Captioning is a task where each image must be understood properly and are able generate suitable caption with proper grammatical structure .Here it is a hybrid system which uses multilayer CNN (Convolutional Neural Network) for generating keywords which narrates given input images and Long Short Term Memory(LSTM) for precisely constructing the significant captions utilizing the obtained words .Convolution Neural Network (CNN) proven to be so effective that there is a way to get to any kind of estimating problem that includes image data as input. LSTM was developed to avoid the poor predictive problem which occurred while using traditional approaches. We used an encoder-decoder based model that is capable of generating grammatically correct captions for images. This model makes use of VGG16(Visual Geometry Group) as an encoder and LSTM as a decoder. The model will be trained like when an image is given model produces captions that almost describe the image. The efficiency is demonstrated for the given model using Flickr8K data sets which contains 8000 images and captions for each image but we use CNN and LSTM to capture dependencies and tell both the spatial relationships of images and contextual information of captions and generate contextually relevant captions.

# CONTENTS