

Image Caption Generation

Providing you captions for the image you provide.



Problem Statement

You see an image and your brain gives the description for it. Can a computer do it for you ?Can it suggest you ideas ...



Segments of Problem



Recognising the objects

The first problem in recognizing an image is to recognize what are the objects in it?

Converting to Natural Language

Now we have recognized the objects.
How to give a meaningful description?

Suggesting Captions

Well my caption needs to be impressive.
How can I make it striking?



Our Objective

The objective of this project is to recognize the context of an image and provide a meaningful caption suggestion in natural language.

Easing Human Effort

By Automatically providing description to the image/s.

Providing assistance to visually impaired

The image caption can be converted to audio for helping visually impaired

Reaching to unsafe places

It can be employed at places unsafe for humans.

Significance of Project

Through this project we are trying to discover new boundaries for the image recognition and natural language processing techniques. It carries limitless scope. Some of which are enlisted here.

Section 1: Recognizing Objects

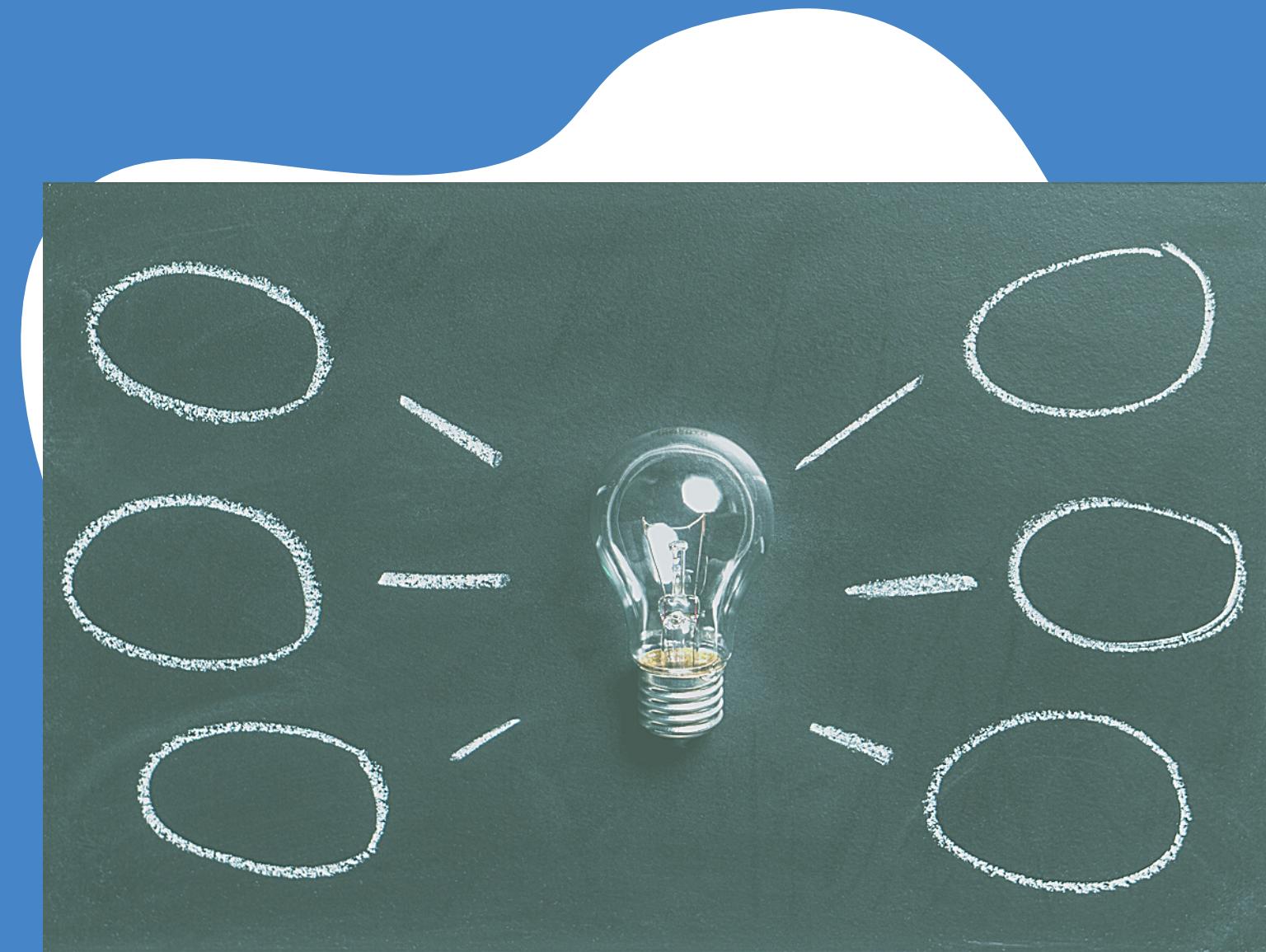
We are using Convolution Neural Network technique.

Section 2 : Natural Language Conversion

We are using Long Short Term Memory technique of RNN.

Section 3 : Providing Captions

For proving meaningful caption suggestion, we will use suggestion API.



Training Model

Every model need to be trained. For that..

Flickr8k Dataset

The primary dataset we will be using.

This contains images and related captions for training our Xception model.

Flickr30k Dataset

We will be using this data set in final stage of our project to improve it accuracy.



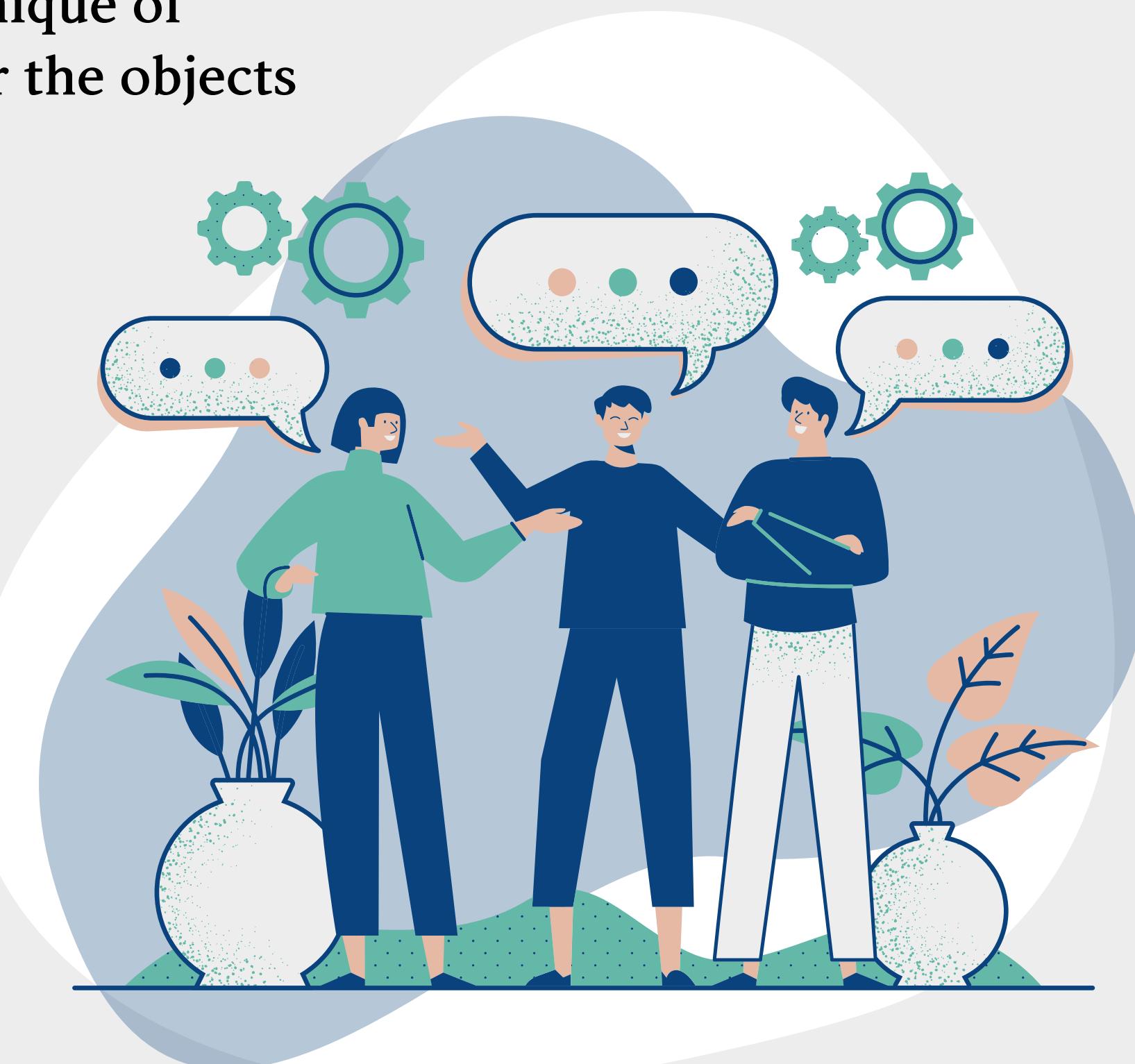
Convolution Neural Network

Convolution Neural Network is a deep learning technique used for recognizing object/s in an image. Different object/s are recognized based on there color and depth in the image.



Long Short Term Memory

Xception model using Long Short Term Memory technique of Recurrent Neural Network is used for generating string for the objects recognized by CNN.

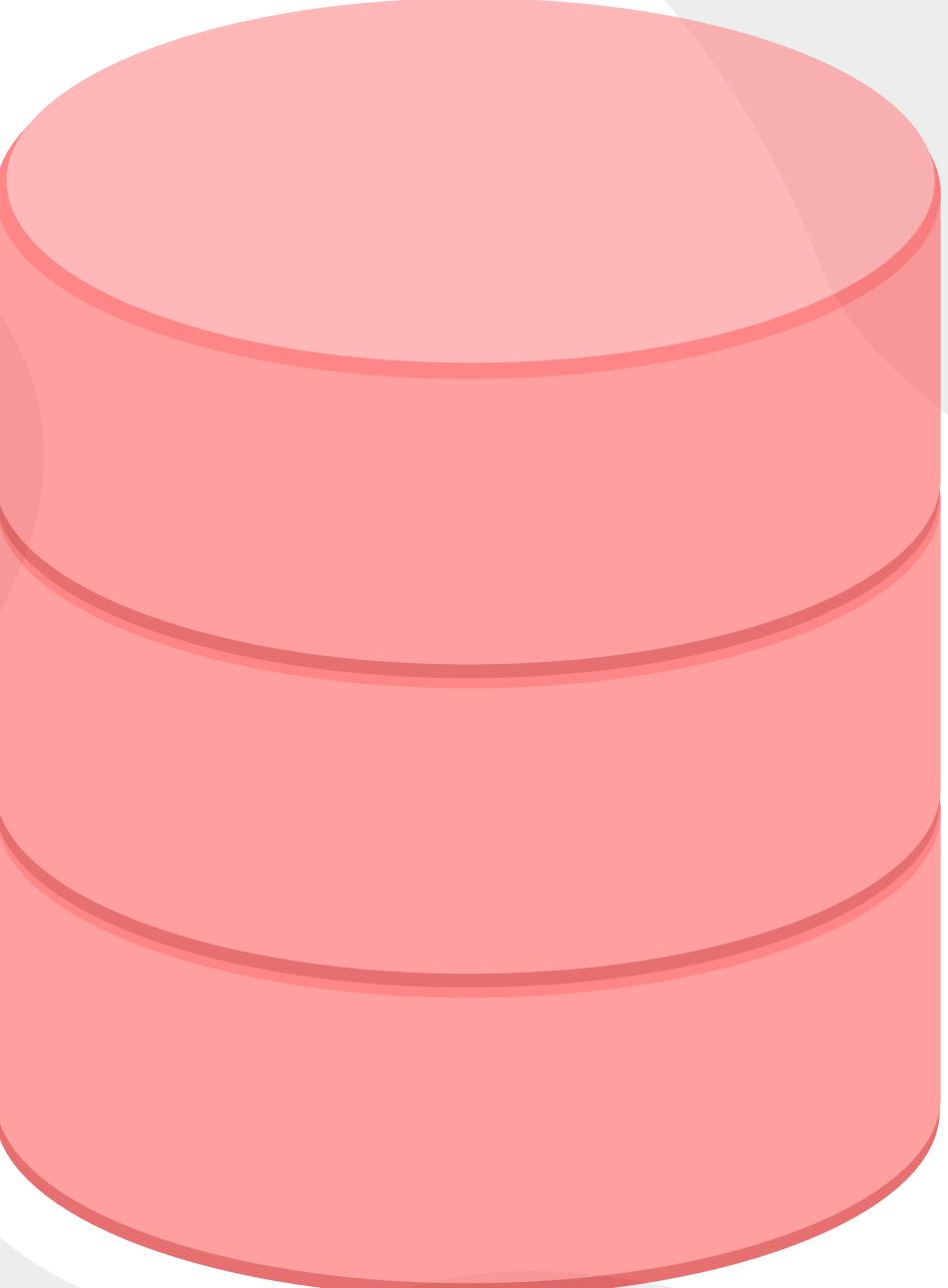


Suggestion API

The string generated by the LSTM is provided to our suggestion API, which then provide few meaningful caption suggestions to the user.



Saving to Database



Although we learned from the dataset. But to grow regularly we need to learn regularly. We will save the images and caption provided by our program in our database to improve our results.



Product or Service

Introduce your company's product or service as the ultimate solution to these problems.

How to use the product?

Just upload the image whose caption you want. We will process the image and provide you with a few suggestions for the caption.

Upload

Upload the image

Processing

By using CNN,
LSTM and
suggestion API, we
process the image.

Results

Finally you will
have suggestions
for captions to
choose.

The Team



Priya Chawla

Title



Mayank Singh

Title



Pritika Rana

Title