

# High Level Design (HLD)

---

## Image Captioning

Revision No: 1.0

Last Date of Revision: 23/11/2023.

Document Version Control-

<u>Date Issued</u>	<u>Version</u>	<u>Description</u>	<u>Author</u>
23/11/2023	1.0	HLD – V- 1.0	Mahesh. A

## Contents -

Description	Page No.
Document Version Control	2
1. Introduction	4
1.1 Necessity of HLD	4
2. General Description	4
2.1 Product Perspective	4
2.2 Problem Statement	4
2.3 Data Requirements	4
2.4 The format would as the below mentioned sample	6
2.5 Tools Required	7
3. Design Details	8
3.1 Process Flow-	8
3.2 Model Training and Evaluation-	8
3.3 Deployment Process	9
3.4 Event Log	9
3.5 Error Handling	9
4 Performance-	
4.1 Reusability	10
4.2 Application Compatibility	10
4.3 Resource Utilization	10
4.4 Deployment	10
5 Conclusion	

## **1 Introduction**

### **1.1 Necessity of HLD-**

The purpose of this document is to add the necessary information to the current project for Modelling and its coding. This can be treated as a user manual.

## **2. General Description**

### **2.1 Product Perspective-**

This Image Captioning is the project is based on the Deep learning sequence model which helps us to caption an image.

### **2.2 Problem S-**

To create the AI solution for the Image Captioning

. To detect to caption an image

### **2.3 Data Requirements-**

The data requirements completely depend on the problem statement.

We need the data in the .text and images.

This text file contains line that have the information of the names of .jpg images and their predefined captions separated by tab.

## 2.4 The format would as the below mentioned sample

File Path	Caption	Text
1305564994_00513f9a5b.jpg#0		A man in street racer armor be examine the tire of another racer 's motorbike .
1305564994_00513f9a5b.jpg#1		Two racer drive a white bike down a road .
1305564994_00513f9a5b.jpg#2		Two motorist be ride along on their vehicle that be oddly design and color .
1305564994_00513f9a5b.jpg#3		Two person be in a small race car drive by a green hill .
1305564994_00513f9a5b.jpg#4		Two person in race uniform in a street car .
1351764581_4d4fb1b40f.jpg#0		A firefighter extinguish a fire under the hood of a car .
1351764581_4d4fb1b40f.jpg#1		a fireman spray water into the hood of small white car on a jack
1351764581_4d4fb1b40f.jpg#2		A fireman spray inside the open hood of small white car , on a jack .
1351764581_4d4fb1b40f.jpg#3		A fireman use a firehose on a car engine that be up on a carjack .
1351764581_4d4fb1b40f.jpg#4		Firefighter use water to extinguish a car that be on fire .
1358089136_976e3d2e30.jpg#0		A boy sand surf down a hill
1358089136_976e3d2e30.jpg#1		A man be attempt to surf down a hill make of sand on a sunny day .
1358089136_976e3d2e30.jpg#2		A man be slide down a huge sand dune on a sunny day .

## 2.5 Tools Required-

Python Programming language and the frameworks below mentioned are used to build the model.

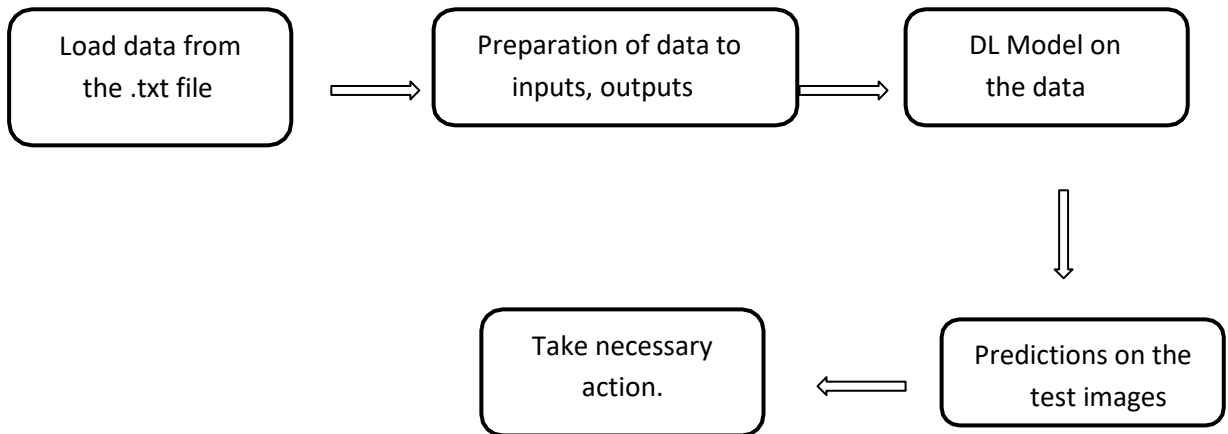
- Numpy
- Matplotlib
- Logging
- Flask
- NLTK
- Tensorflow
- Keras
- Pickle
- Joblib
- Unicorn
- OpenCV
- Pillow
- Regex

- a. VS code is used as the IDE
- b. For visualization we use Mat plot library for images
- c. We use local host for deploying the model
- d. HTML/CSS for the front end development
- e. Python flask used of the backend development
- f. Git is used for version control
- g. Docker is used of creation of model that works on any architecture.



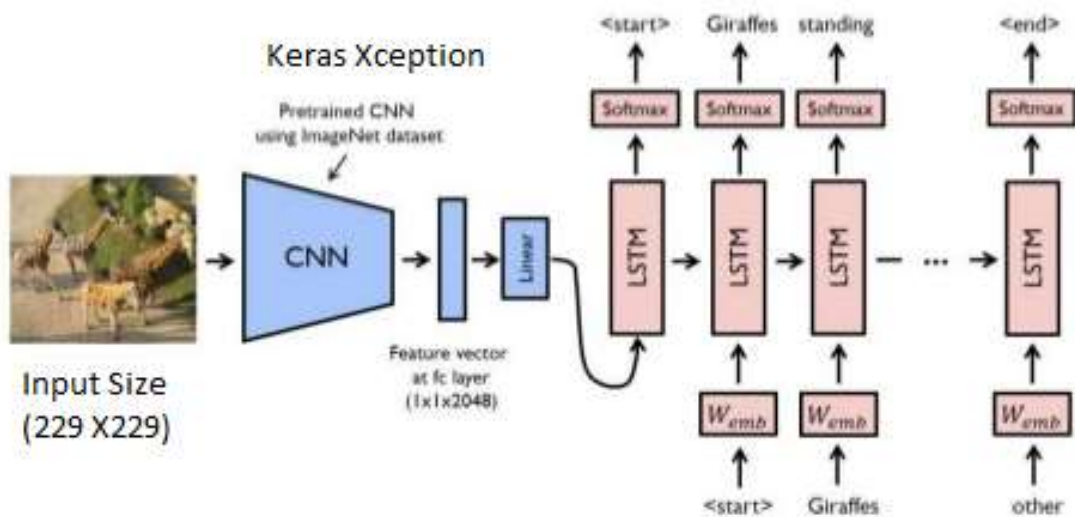
### 3 Design Details

#### 3.1 Process Flow-

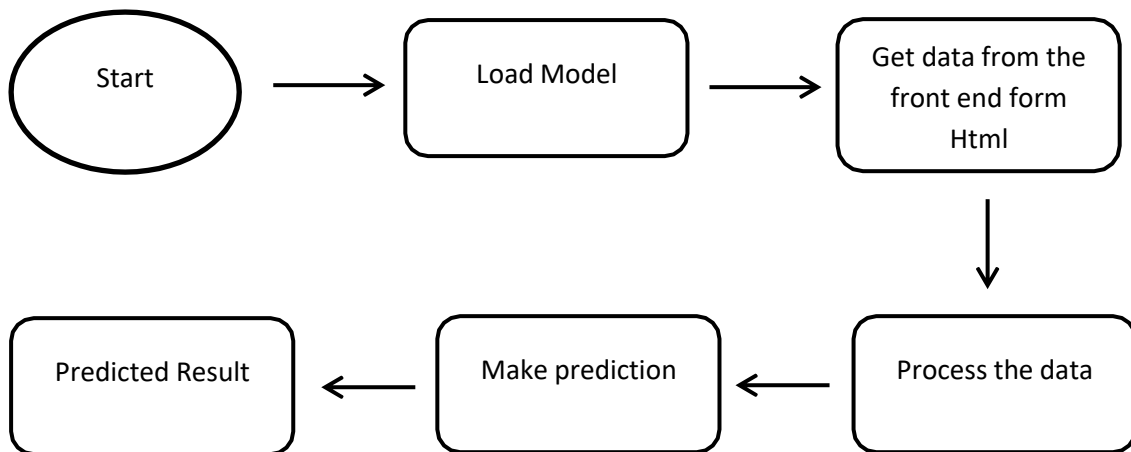


#### 3.2 Model Training and Evaluation-

### Model



### 3.3 Deployment Process:



### 3.4 Event Log-

The system should log every event so the user will know the process that happens on the time.

The system should identify the different logs

The system should note down the logs for further usage

Developer uses this data

### 3.5 Error Handling-

By using the logs file all errors can be noted and find the solution for them by the developer.



## 4 Performance-

This Image captioning DL model (CNN and RNN) is used to generate the captions of the given images.

### 4.1 Reusability-

The code written should be reused without any problem.

### 4.2 Application Compatibility-

We are using python as an interface.

### 4.3 Resource Utilization-

When any task is performed, it will use all the processing power to do the task assigned.

### 4.4 Deployment-



## 5 Conclusion-

This Image captioning DL model (CNN + RNN) which helps user to caption the images....