Ideation Phase

Empathize & Discover

Date: 29 October 2023

Project Name Project - Image Caption Generation

Project Description:

In this project, we aim to create captions for images using deep learning techniques. This involves teaching a computer to understand the content of an image and describe it in human-readable language. To achieve this, we'll leverage Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. While humans find image captioning easy, it's a challenging task for computers. The key is to combine the power of CNNs, which excel at image processing, with LSTMs, which are effective in understanding sequences and generating natural language. The project will require substantial datasets and computing power, considering the growing capabilities of deep learning techniques. Through Python, we'll implement and train models that can analyze images and generate descriptive captions. This fusion of CNN and RNN (LSTM) will enable the computer to comprehend visual content and express it in a language that humans can understand.

Empathy Map

Image Caption Generation

WHO are we empathizing with?
Who is the person we want to understand?
What is the situation they are in?
What is their role in the situation?

people who use social media platforms, sharing image,

GOAL

By training the model with diverse datasetsand tools

What do they need to DO?
What do they need to do differently?
What job(s) do they want or need to get done?
What decision(s) do they need to make?
How will we know they were successful?

Professionals who analyze and interpret images, depends on accurate and correct captions to understand the content.

By using LSTML model and CNN to get high accuracy of image caption

What do they HEAR?
What are they hearing others say?
What are they hearing from friends?
What are they hearing from colleagues?
What are they hearing second-hand?

What do they THINK and FEEL?

PAINS
What are their fears, frustrations, and anxieties?

GAINS
What are their wants, needs, hopes, and dreams?

users see the diversity of images ,varying contexts ,design features

Not correct and nonsensical captions

Time takeing is more

Enhancing the user experience

User hear about the developments in image captioning

What do they SEE?
What do they see in the marketplace?
What do they see in their immediate environment?
What do they see others saying and doing?
What are they watching and reading?

Developers may find it difficult to train models that handle diverse languages and contexts

Leading to improved application functiones

User hear about the actions and strategies of image captioning field through reports and discussions

What other thoughts and feelings might influence their behavior?

What do they SAY?
What have we heard them say?
What can we imagine them saying?

User may feel frustrated with the limitations of current image captioning systems

Emerging technologies and alternative approcahes of image captioning

The image captions are not just perfect but also provide meaningful context

Actively monitor trends in image captioning

What do they DO?
What do they do today?
What behavior have we observed?
What can we imagine them doing?

They collab with developers to the needs of the use