Problem Statement 1 - Assembly Ally

Introduction

In our era of self-assembly and DIY projects, the challenge of understanding and following diverse instructional guides is a common hurdle. From assembling a piece of furniture to building a complex model airplane, the range of tasks calling for clear, step-by-step guidance is vast. This creates a unique opportunity for a technological solution that simplifies and enhances this process.

Problem Statement

The task is to create "Assembly Ally" a versatile Conversational AI designed to assist mechanics with a wide array of instructional manuals for the assembly and repair of vehicles. The goal is to provide an interactive, AI-driven assistant that makes the assembly process easier, faster, and more efficient while onboarding a new mechanic into the team.

Data

Some example guides / manuals with a few example questions are mentioned here:

🗀 **Assembly Ally**

Keep in mind that you are free to choose any similar document while sticking true to the problem statement.

Technical Challenges

Robust Assistant Development: Develop a robust AI based assistant capable of interpreting various instructional manuals. This includes sophisticated natural language processing and computer vision algorithms to understand user questions and provide accurate, helpful responses.

Extensive Manual Database Integration: Incorporate a diverse database of instructional manuals. The Conversational AI should be able to navigate these different formats and content types effectively.

Accurate Responses: The assistant should be able to accurately answer queries, summarize sections, and retrieve information from the document that is provided.

Adaptive Learning and Personalization: Ensure the Conversational AI can adapt to different user comprehension levels and learning styles, offering personalized assistance based on user interactions and feedback.

User-Friendly Interface Design: Design an intuitive and engaging interface that facilitates easy interaction with Conversational AI across various devices and platforms that will enable users to find products and repair cars.

Deliverables

- An advanced Conversational AI assisting with a wide range of instructional manuals responsive across mobile, tablet, and desktop screens.
- The Conversational AI should be able to understand a wide range of data such as **text, tables, and images** mentioned in the reference document.
- The AI should be able to extract information from the document and look up key specifications such as part name, part number, etc.
- A full-stack web or mobile application interfacing with the AI model.
- Detailed documentation on the development, AI implementation strategies, and user instructions.

Brownie Point

Generate reports from the documents that will summarize the extracted information in simple and understandable chunks.

Judging Criteria

- Versatility in handling queries across various types of instructional manuals.
- Precision and clarity in the Conversational AI's responses and guidance.
- Innovation in AI technology, particularly in image recognition and adaptive learning.
- The time taken for the application to come up with responses for any user query.
- Quality of user experience, including interface design and interaction fluidity.

Problem Statement 2 - Smart Inventory Tracking System

Introduction

The world is moving in a direction where life-changing decisions are made in seconds. This has never been more important than the healthcare sector where thousands of patients are being treated every day. Each patient requires several dozens if not hundreds of items to be available in the health care provider's inventory. The inventory managers of such institutions are tasked with managing thousands of tools, medicines, and equipment that need to be updated in real-time and procured if required.

Problem Statement

Participants are required to develop a system that can analyze images of shelving systems containing various medical objects and generate a digital map for every type of object in the image. The unique aspect of this challenge is achieving this with minimal or no training data, in a single-shot manner.

Data

You are provided videos of the shelf system out of which the required images of the shelves and the objects can be extracted.

You can find the data here: 🗀 **Smart Inventory Tracking System**

Technical Challenges

Image Recognition and Analysis: Accurately identify and categorize objects in diverse shelving environments.

Data Sparsity: The solution must work with limited or no prior training data about the shelves or the objects that can be kept on the shelves. Some objects might look similar to each other but are different objects kept in similar packages.

Algorithm Efficiency: It is crucial to process information in real time so that decisions about inventory are made as soon as possible. The algorithms you are developing are expected to run on the edge or the cloud in real time.

Environmental Awareness: The algorithm must understand and interpret the spatial arrangement of objects on shelves and the shape and orientation of the objects themselves. The algorithm mustn't misinterpret the location of one object over another.

Deliverables

- An intuitive AI application where users can upload images of the shelving units and generate a map of different items present in the image.
- The AI should be able to differentiate between similar looking objects that might have similar shapes and sizes but are kept different because of the content in them.
- A full-stack web or mobile application to onboard and visualize the map of every object on the image.
- Detailed documentation on the development, AI implementation strategies, and user instructions.

Brownie Point

Given an image of an item on the shelves, provide the location of the object based on the map that was generated and provide the count of the same type of object visible on the image.

Judging Criteria

- Versatility in handling different types of orientations and positions of the objects.
- Ability to handle different types of objects kept on the shelves.
- Easy to understand the mapping of the different objects.
- Quality of user experience, including interface design and interaction fluidity.

Problem Statement 3 - Trigger Word Detection

Introduction

The importance of touch free interaction has never been more important in the post covid world. Touch free interfaces are not only safer from contact born diseases they are also much more convenient to use. You use touch free and voice based tools every day from your vacuum machine to your mobile phone and your car has touch free components to make the interaction seamless. The issue is that even the labs at Apple, Google, and Amazon have yet to crack the holy grail of creating a robust wakeword detection system that is resilient to mispredictions. That's where you come in.

Problem Statement

You are required to develop a web based or mobile application that can interphase with the microphone on the device and provide visual feedback to the user based on the word that was spoken.

The model at a later stage is expected to run on an edge device with limited resources, you are expected to design the model in such a way that it can be deployed on a Raspberry Pi 3 or a STM32H747II.

Data

You will be provided samples of the wake words **door open, door stop,** and **door close.**
You can find the dataset here: 🗀 **Trigger Word Detection**

Keep in mind that you are free to choose your own data as well as long you believe that it is relevant to the given problem statement.

Technical Challenges

Extremely diverse data conditions: The model is expected to perform with a high accuracy on a diverse set of environmental conditions such as areas with background noise and the audio patterns of a diverse gender and pronunciations of the same word.

Compute Limitation: Being that the model must be deployed on a low compute device the model is expected to utilize under 8MB of RAM and 16MB of ROM at any given time. You are expected to produce some kind of proof to back up the claims that the model is only utilizing the above mentioned resources.

Algorithm Efficiency: The time taken to perform a prediction must be less than the length of the audio used for the inference to be considered realtime.

Deliverables

- An intuitive full stack application where users can record audio using the microphone of the device it is running on.
- A model capable of notifying the user whenever the words **door open, door close, or door stop** are spoken by the user.
- A description to support the claim that the model can be deployed on a Raspberry Pi or an STM32H747II
- Detailed documentation on the development, AI implementation strategies, and user instructions.

Judging Criteria

- Performance of the model on an unknown test dataset.
- The accuracy of the model in recognizing wake words for a diverse set of people and environmental conditions.
- Quality of user experience, including interface design and interaction fluidity.

Problem Statement 4 - Handwriting Recognition

Introduction

In a world where handwritten data remains prevalent, especially in sectors like healthcare, the challenge of accurately interpreting and digitizing this information is critical. Handwritten medical prescriptions, often difficult to decipher due to varied handwriting styles and medical shorthand, pose a significant challenge for data entry and analysis. This presents an opportunity for a technological breakthrough that can efficiently and accurately convert handwritten prescriptions into digital format.

Problem Statement

The objective is to develop a sophisticated Optical Character Recognition (OCR) system specifically tailored for interpreting and digitizing handwritten medical prescriptions. The aim is to create an AI-driven tool to accurately read, interpret, and convert varied handwriting styles into digital text. This tool should facilitate easier data management, improve prescription accuracy, and enhance patient safety.

Data

You are provided images of handwritten here: 🗀 **Handwriting Recognition**

You are free to choose your own dataset as long as you make sure to stick true to the problem statement.

Technical Challenges

Advanced OCR Development: Design a highly efficient OCR system capable of recognizing diverse handwriting styles and medical abbreviations. This requires integrating advanced image processing and machine learning techniques.

Comprehensive Training Data: Utilize a vast and varied dataset of handwritten prescriptions to train the OCR model, ensuring high accuracy and reliability in real-world applications.

User-Friendly Interface: Develop an intuitive interface that allows healthcare professionals to easily scan, upload, and verify digitized prescriptions.

Deliverables

- A state-of-the-art OCR system specifically designed for medical prescriptions.
  The OCR tool should reliably interpret text, symbols, and abbreviations found in handwritten prescriptions.
- A user-friendly full stack application, accessible on various devices, for scanning and verifying digitized prescriptions.
- Comprehensive documentation detailing the OCR system's development, machine learning models used, and user guidelines.

Brownie Point

Develop a feature that automatically flags potential prescription errors or anomalies for further review by healthcare professionals.

Judging Criteria

- Performance of your algorithm on an unknown test dataset
- Effectiveness in handling various handwriting styles and medical shorthand.
- Accuracy of the digitized text, especially in correctly identifying medication names and dosages.
- Innovation in OCR technology and machine learning algorithms.
- Speed and efficiency of the OCR system in processing and digitizing prescriptions.
- Quality of user experience, including ease of use, interface design, and system integration capabilities.