Assignment.

Name : Varshini S
Rollno : 2407401B

ASR (Automatic Speech Recognition) :-

ASR is the process of converting spoken langu (audio/speech) into written text using computati model It's a key technology behind voice assiste like Siri, Alexa, google Assistant and tools like youtube captions.

ASR works :-

Step 1 : Speech input (Audio)

Step 2 : preprocessing the Audio

Step 3 : feature Extraction

Step 4 : Acoustic model

Step 5 : language model.

Step 6 : Decoding.

Example

> "find action movies from the 90's"

* your voice is captured through the tv remote's microphone

* The audio is split into short frames (20 ms)

* The system calculates MFCCs and other features that response the sound patterns

* The system converts the sound into likely phonemes:

/f/, /aɪ/, /n/, /d/, /æ/, /k/, /s/, /ə/ etc.

* The system understand that find action movies from the 90s" a likely phrase

* The phonemes and context are combined to output the best matching sentence

final output :-

> "find action movies from the 90s".

Application :-

* Voice typing in smartphones
* Virtual assistance (Siri, Alexa)
* Transcription services
* Subtitles on videos
* Hands-free control in cars or devices

Information retrival :-

* IR (Information retrival) finding relevant document or data form a large collection based on a query

Task :- you search "symptoms of lung cancer" on google.

IR system :- google finds and returns web pages with that information

Knowledge-based question answering :-

* The system answers question using structure knowledge from database (like wikidata)

* Question :- "who is the president of france ?"

KAQA :- It checks a knowledge base and replies "

classic QA Models :-

* These models extract answers from a passage of text (often using) Machine learning or deep learning)

Text : water boils at 100 degrees celsius"

Question : "At what temperature does water boil"?

Answer : 100 degrees celsius (extracted from the text)

Simple frame-based dialogue systems :-

* systems that follow fixed templates (frames) for specific tasks like booking or customer support

use case : Booking a hotel room

System Ask : "what city?", check-in-date?' How many guests ?"

Evaluating Dialogue systems :-

* checking how well a chatbot or voice assistant works based on metrics like correctness, fluency and user satisfaction

Example :-

* you use a customer service chatbot. After the session, you rate it, was your problem solved ? (yes/no) Rate from 1-5 stars

The Automation speech Recognition (ASR) :-

* converting spoken language into text

Example :-

"Set a timer for 10 minutes"

* ASR system transcribes it as text and passes it to the assistant to act