



# TRANSCRIBE GLASSES USING SUPERCAPACITOR TECHNOLOGY



1. Mahmoud Wahba abdelhady..... 20196779



2. Ahmed Emad asklany..... 20196174



3. Mariem Essam Saber Ghazy..... 20196073



4. Mohamed Hamdy Abdelrahiem..... 20196125



5. Fahd Hany Mohamed ..... 20197193



6. Mustafa Ashraf Mustafa..... 20196200

Under Supervision of

**Dr. Nariman Abdelsalam**

**Communications and Electronics Engineering Department  
Canadian International College**

# (HOH) people suffering from hearing loss or Hearing problems

Audio or video documentaries or educational videos available in a foreign language



5

Interact with people who speak another languages



4



1



2

International filmmaking event



3

Global, international, political, economic conferences

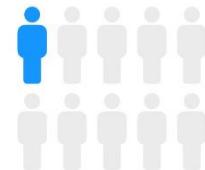


## Statistics and evidence of people's need for this project.

### World health organization



- Over **5%** of the world's population (**430 M**) – suffering from hearing loss (34 million children).
- by **2050** over **700 million** people – or **1 in every 10 people** – will have disabling hearing loss
- The prevalence of hearing loss increases in **60 years, over 25%** are affected by disabling hearing loss.



### European statistical system



- More than **1.5 billion** people worldwide are currently affected by hearing loss in at **least one ear**.
- More than **14%** of adults ages and 65 over use a hearing aid
- Approximately **13%** of adults ages 18 and older experience some difficulty hearing even when using a **hearing aid**.



In the UK, around **40%** of people aged **50 years** old and **71%** of people aged **70** years and older have hearing loss

- An estimated **37 %** of students in low- and middle-income countries are required to learn in a different language, putting them at a significant disadvantage throughout their school life and limiting their learning potential

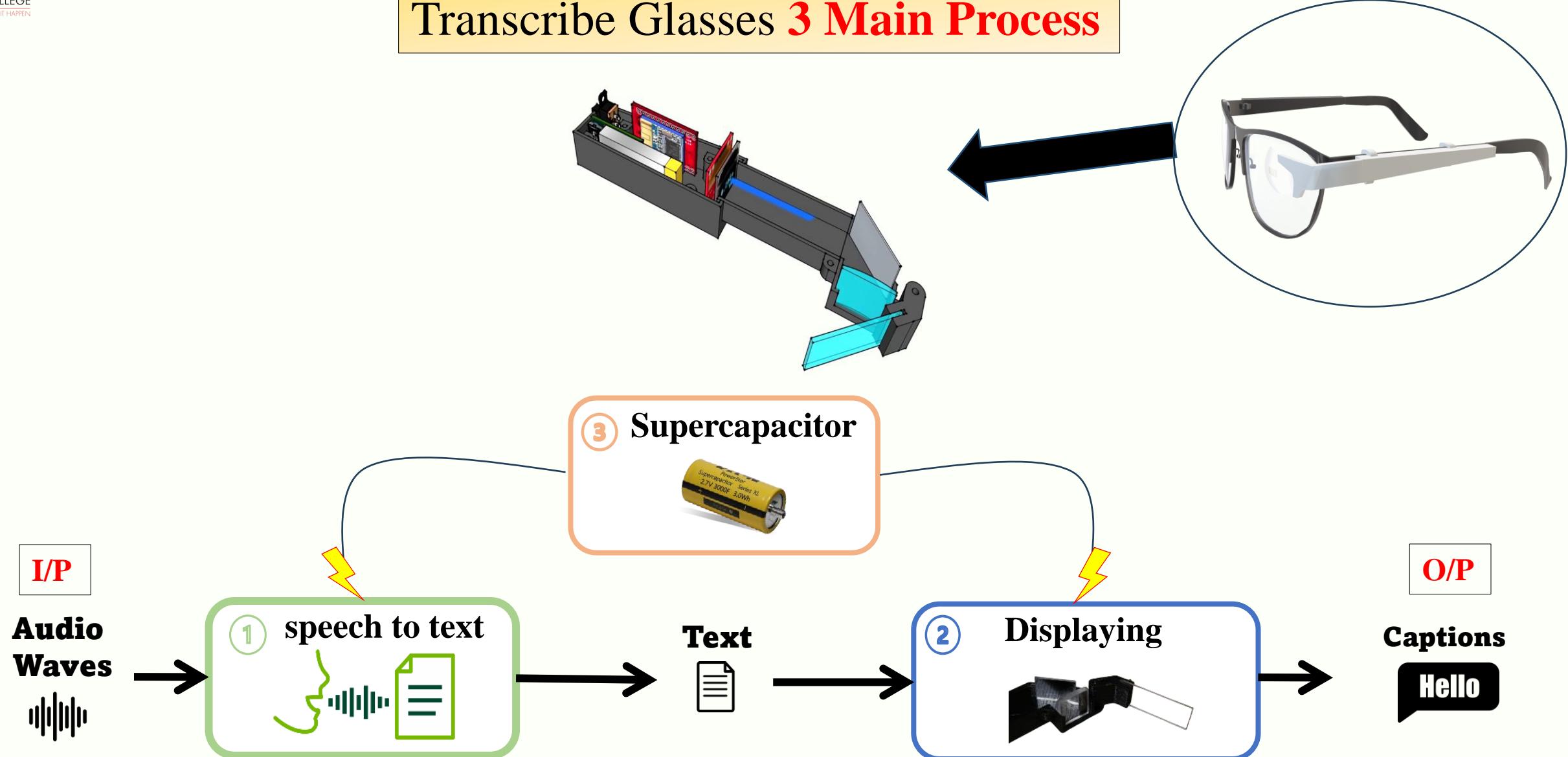
## PROPOSED SOLUTION

### Transcribe Glasses

**Captions**  
**in real life**

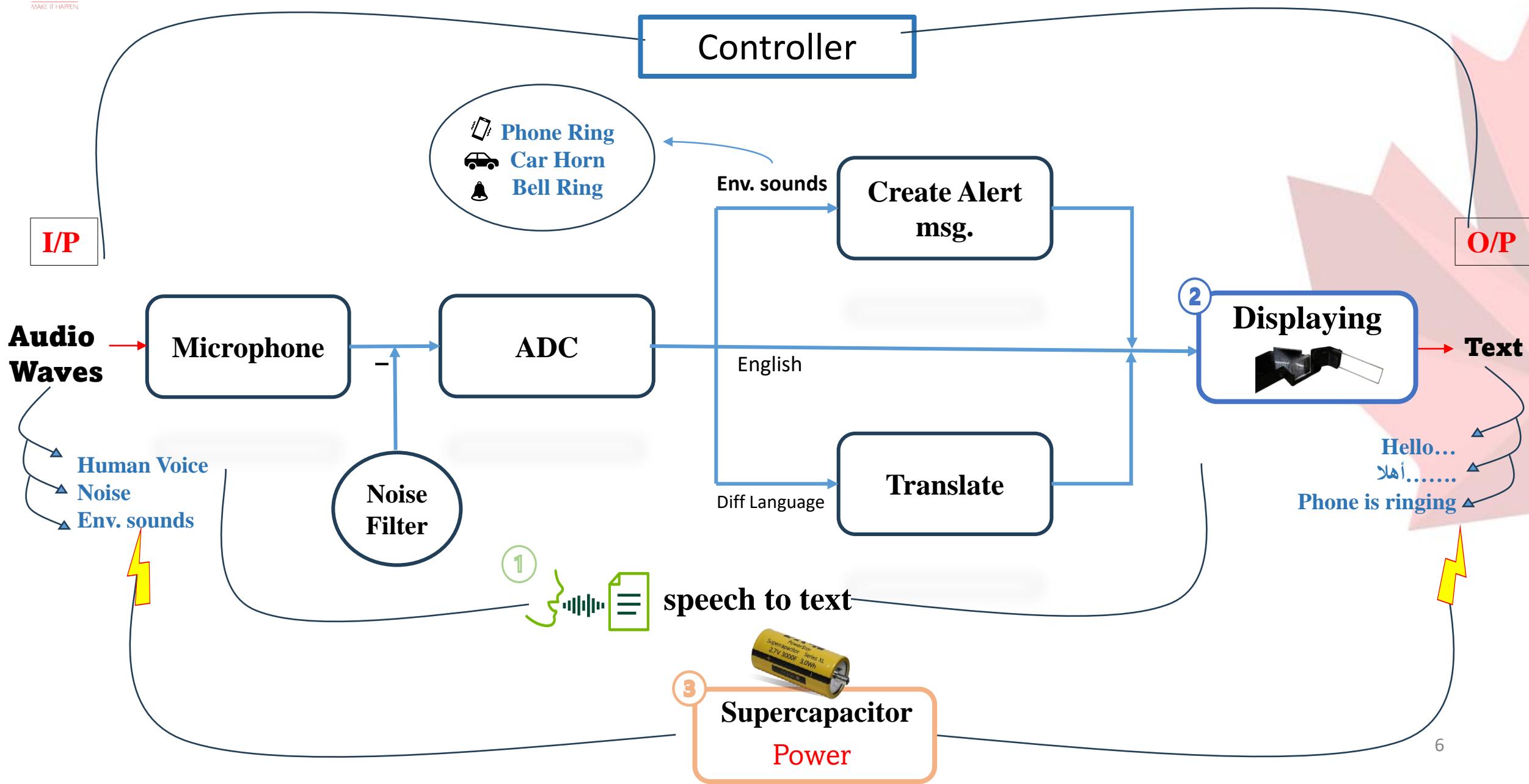


## Transcribe Glasses **3 Main Process**



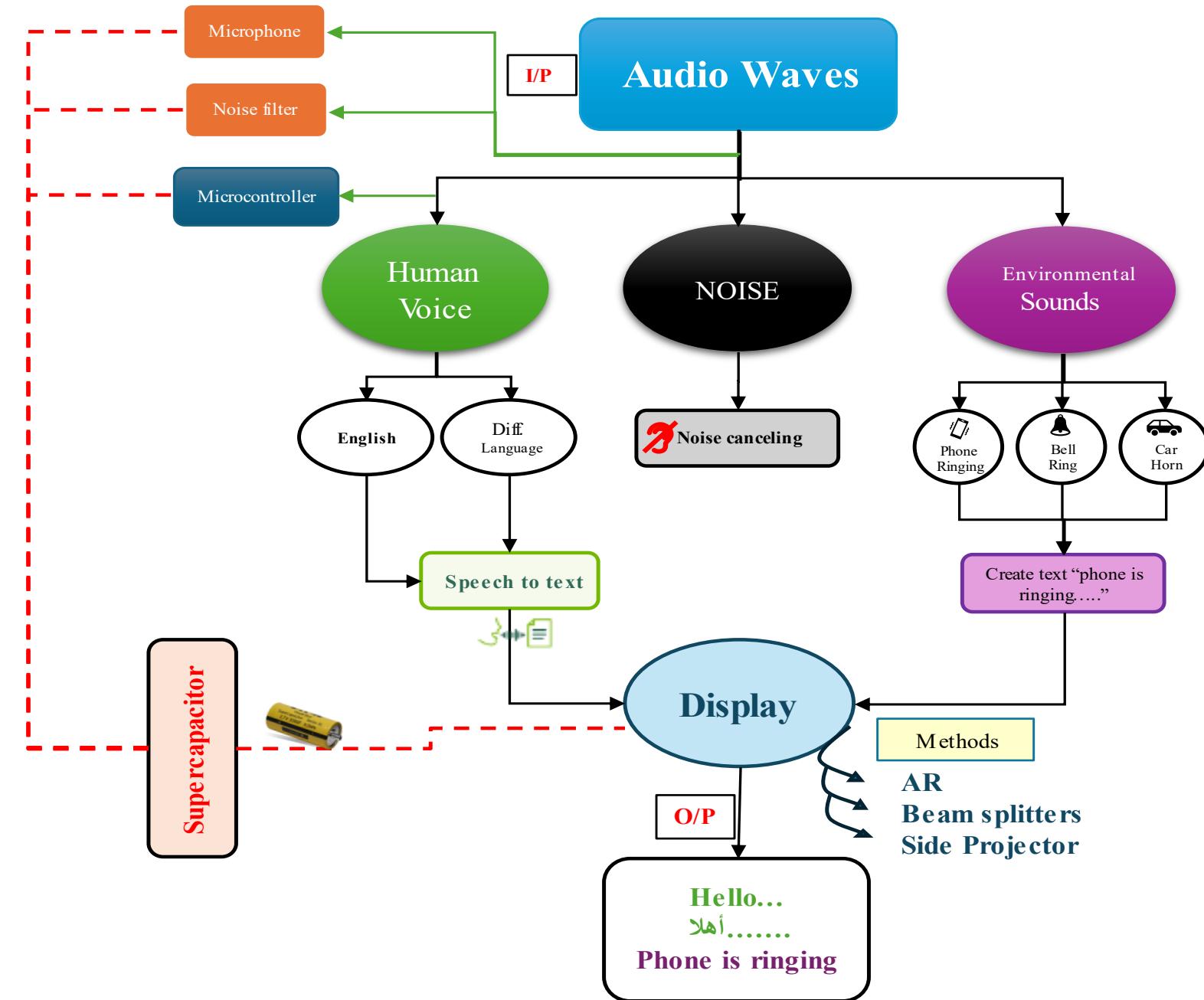


## Transcribe Glasses **BLOCK DIAGRAM**





# Transcribe Glasses **FLOW CHART**



**An explanatory video of what the user sees when trying on the glasses**



## Other similar projects



The Vuzix Blade Smartglasses



Xander Captioning Glasses



TranscribeGlass™



Focals by North smart glasses



XRAI Glass



## Problems

- Limited lithium-ion batteries
- Need to be wired
- Need to use your phone
- Expensive prices



benefit from the problems of previous projects



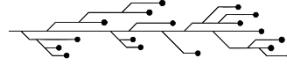
Wearable devices are the future, and we want to keep up with the times



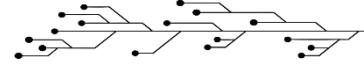
Using our knowledge in the field of com.& electronics to create project that benefits society

## OUR TARGET





# PRESNTATION CONTENT



Part 1

## Transcribe Glasses Overview

- Block diagram of the project
- People benefit from the project
- Similar projects to our project
- Our scope(Goal) of the project

Part 2

## Speech To text

- Data Set
- 1<sup>st</sup> Trained model
- 2<sup>nd</sup> Trained model
- 3<sup>rd</sup> Trained model
- 4<sup>th</sup> Trained model(Whisper)

Part 3

## Controller & Display

- Process of controller in the system
- Esp32
- Problem and solution
- Displaying process

Part 4

## Supercapacitors Overview

- Supercapacitors (definition, uses)
- Super capacitors vs Batteries
- Supercapacitors Types

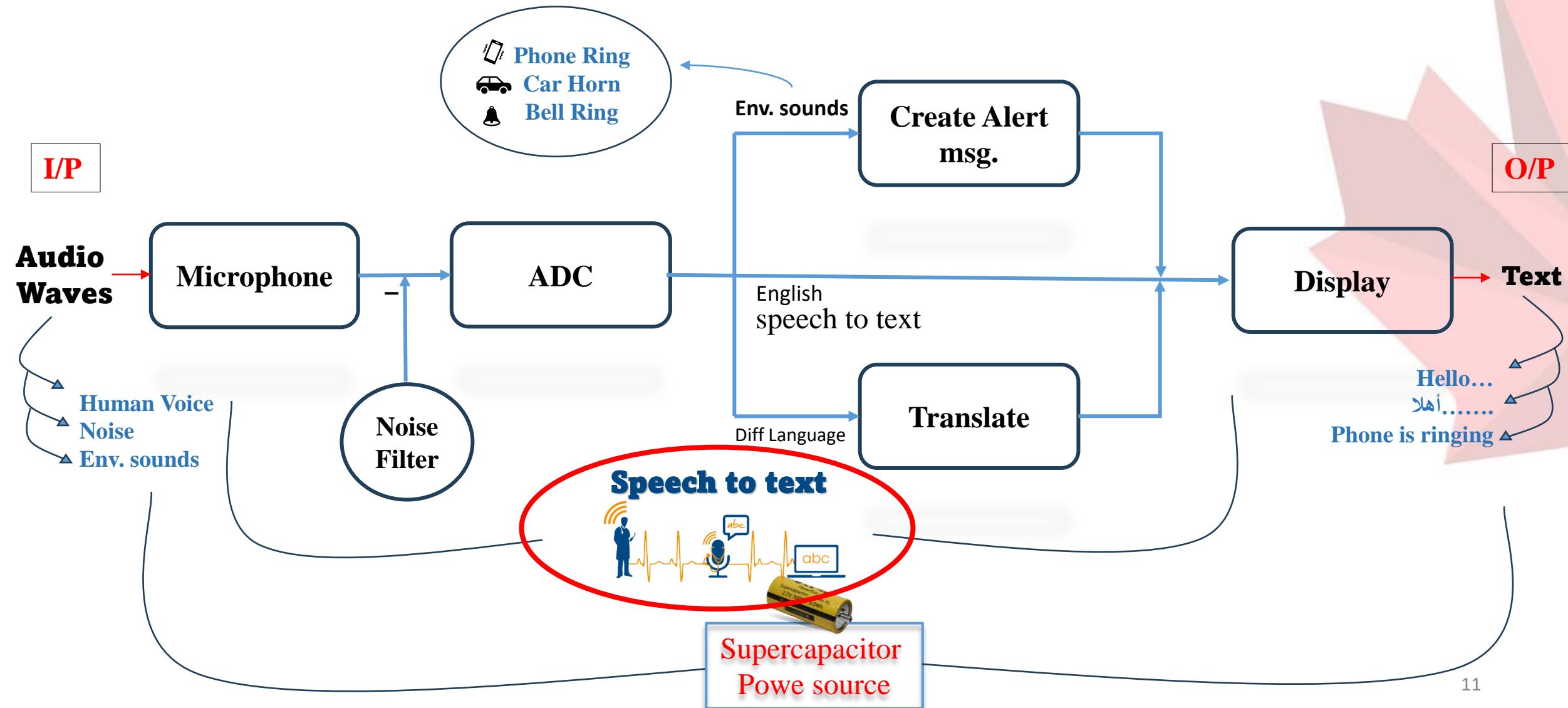
- EDLS
- pseudocapacitor
- Hybrid Supercapacitors

Part 5

## interfacing

- Site Principles
- Front-End , Back-End & API
- Website Architecture

## Transcribe Glasses **BLOCK DIAGRAM**



Part 2

# Speech to text



## ❖ Objective

- Helping HOH people to interact with technology more easily through voice commands and interact easily with people.
- Integration with other system blocks.

## ❖ KPIs

- High accuracy
- Low loss
- High performance

## ❖ Models

- 1st Trained model
- 2nd Trained model
- 3rd Trained model
- 4th Trained model(Whisper)

# Data set

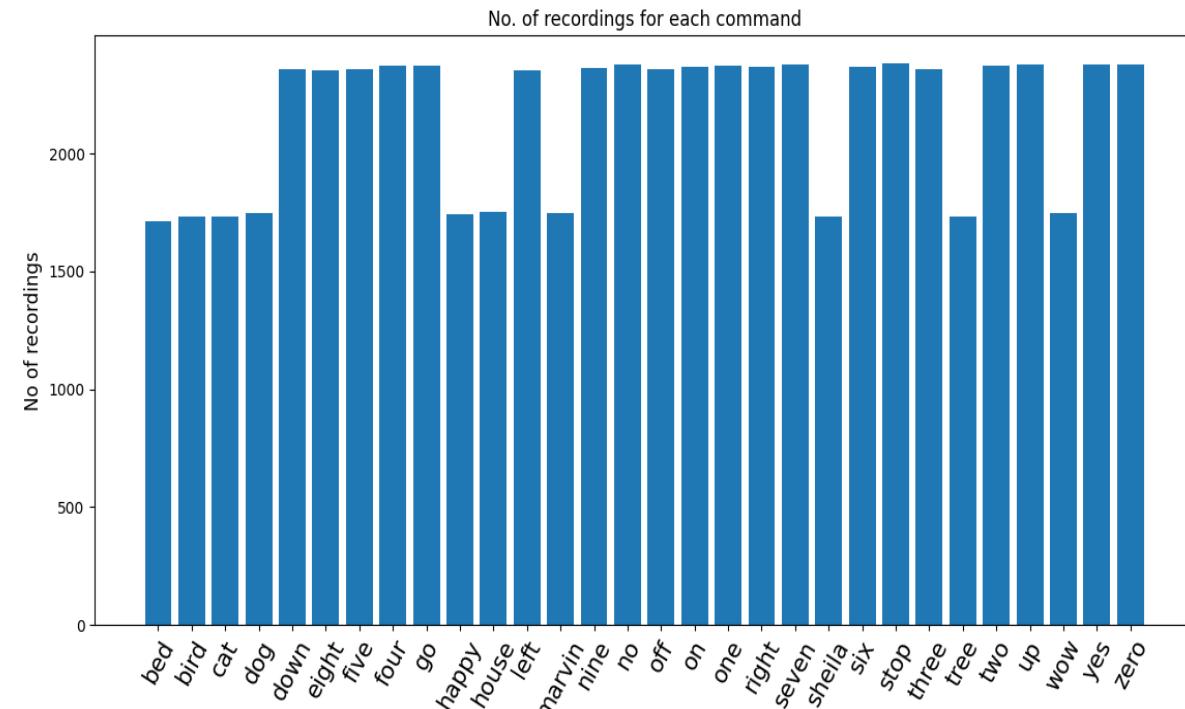
## Data set collection:

- About 64,727 audio files consisted of 30 commands were recorded under different conditions.

## Data set organization:

- The audio files are organized into folders based on the word they contain.

**Problem** Imbalance in data



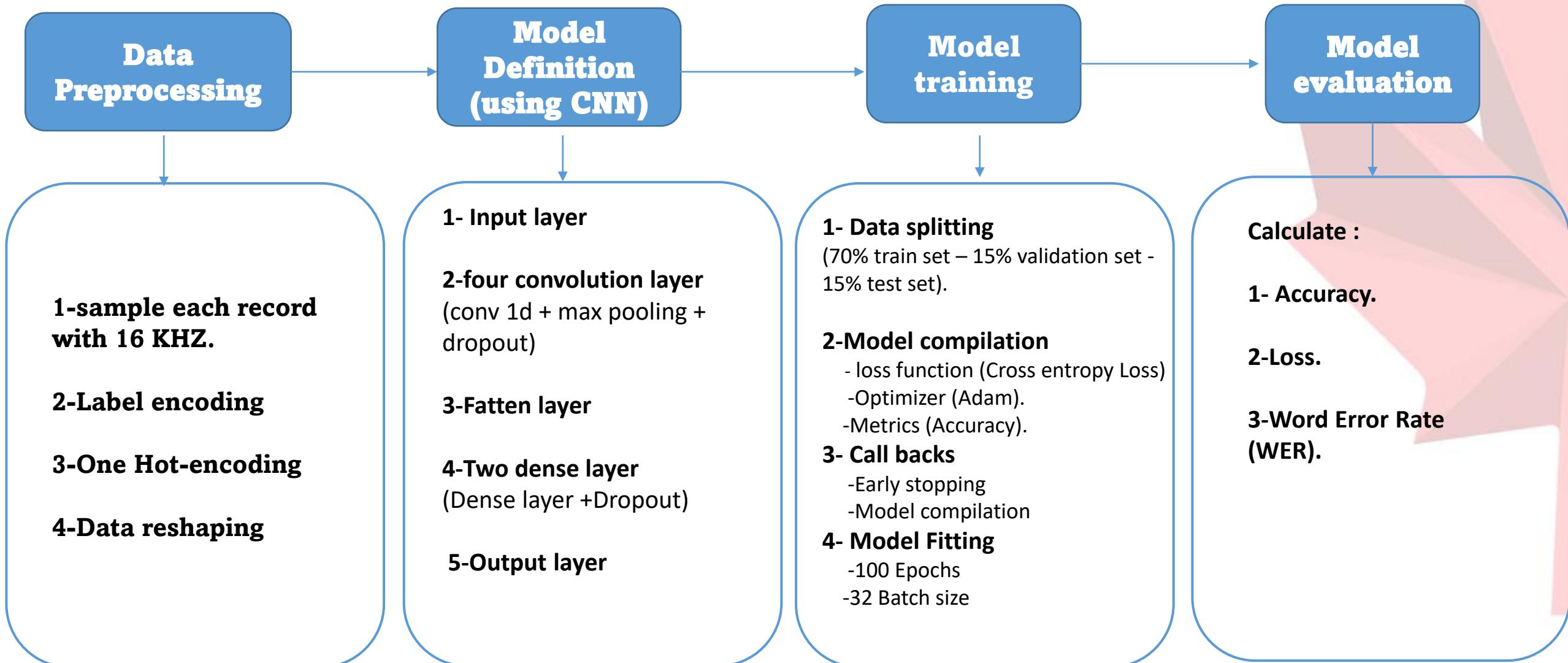
## Solution:

-balance in data by Make each word has 1700 record

## New Size of data set

data set containing 51000 audio files

## First trained model



## First trained model

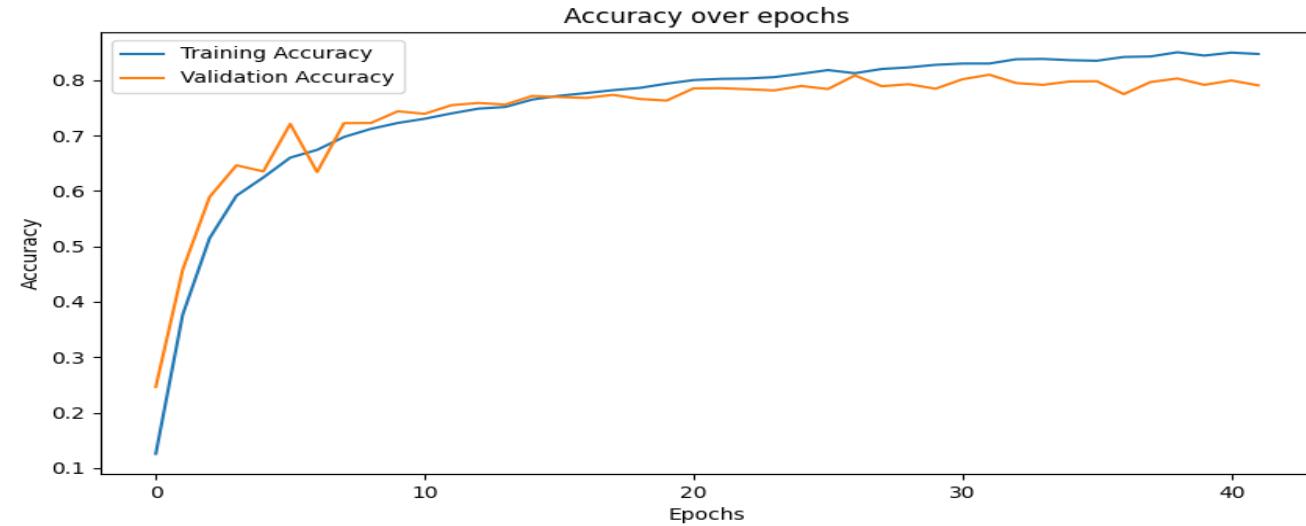
	Accuracy	Loss
Training	83.36%	53.20%
Validation	81%	63.99%
Test	79.33%	75.32%

**WER= 20.54%**

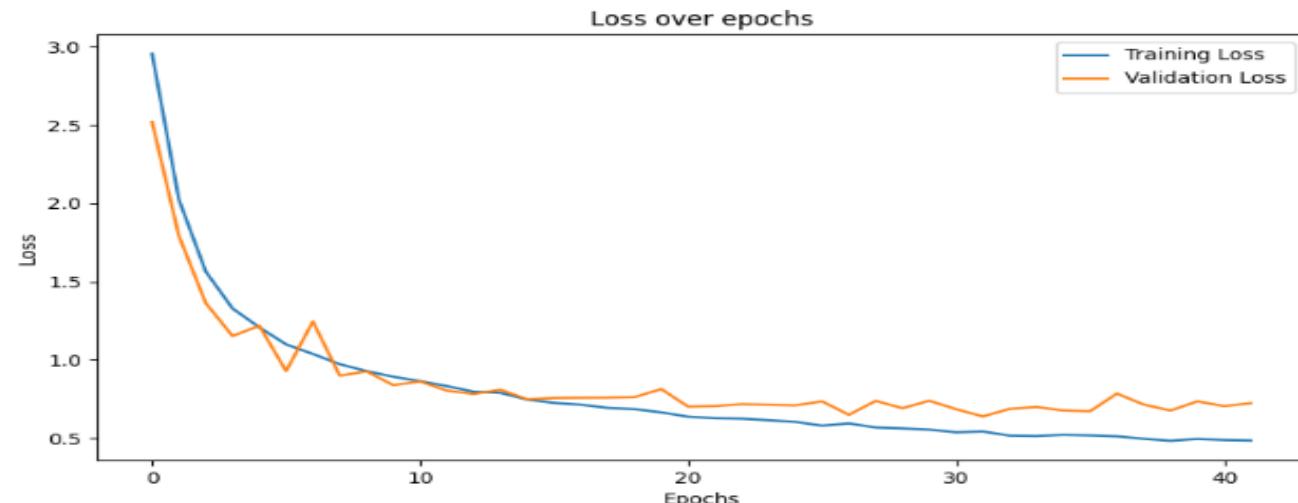
**Problem**

Over fitting

Accuracy between validation and training data:

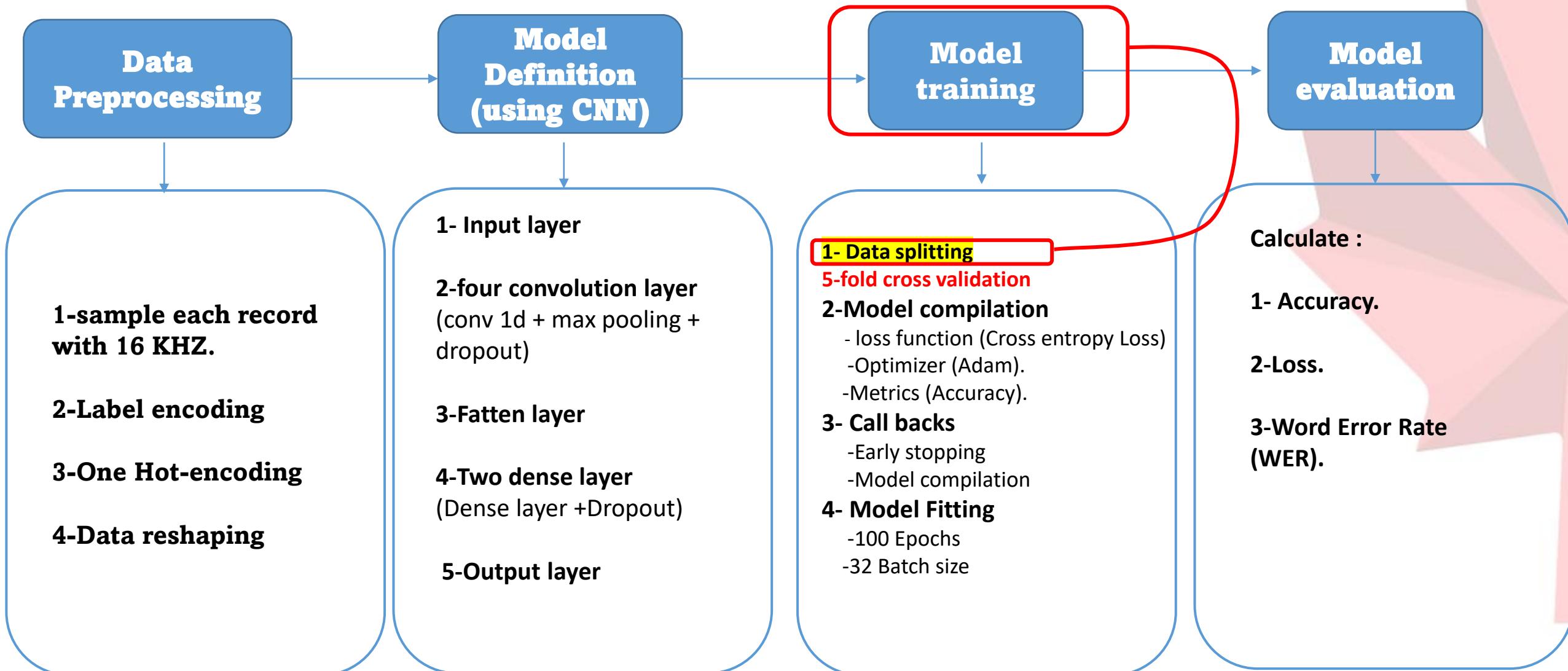


Loss between Validation and training data:



## Second trained model

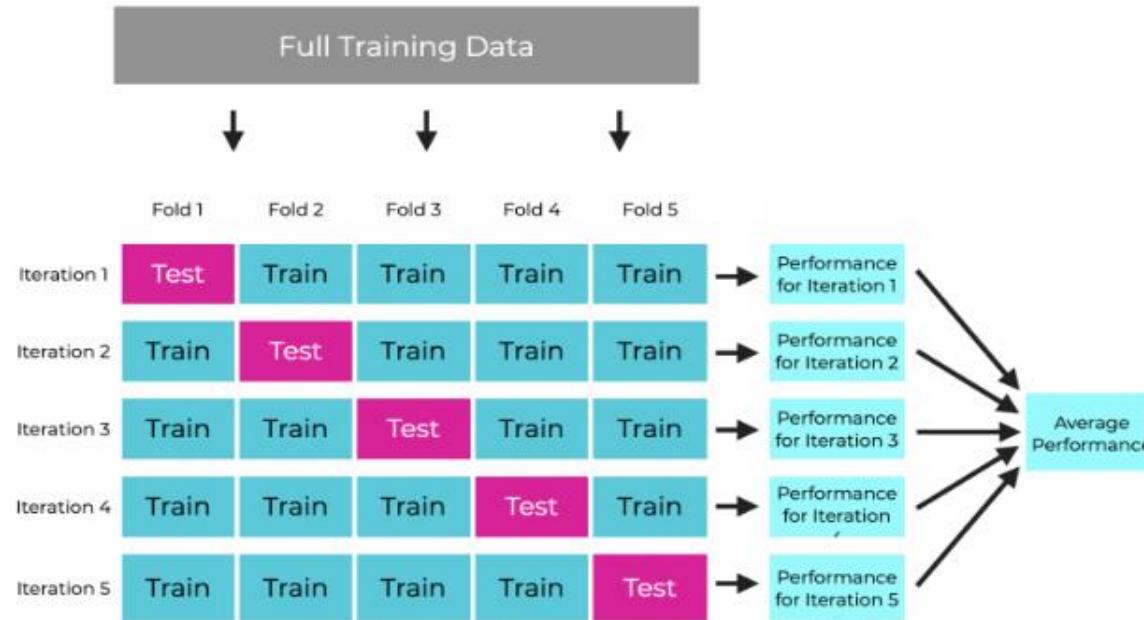
Using (5-fold cross validation)



## Second trained model

### Second trained model

#### 5-Fold cross validation



- Training accuracy and Training loss is the same.
- Testing accuracy **increase** by 2.61% 
- Testing loss **decreased** by 11.71% 
- WER is **reduced** by 2.61% 

#### Results of all folds:

Number of folds	Training accuracy	Training loss	Testing accuracy	Testing loss
Fold1	83.76	51.74	82.30	61.55
Fold2	83.58	53.50	82.92	58.77
Fold3	83.65	52.99	81.79	63.69
Fold4	82.36	65.27	81.46	67.29
Fold5	83.45	54.88	81.24	66.73

#### Average performance:

	Accuracy	Loss
Training	83.36%	53.20%
Test	81.94%	63.61%

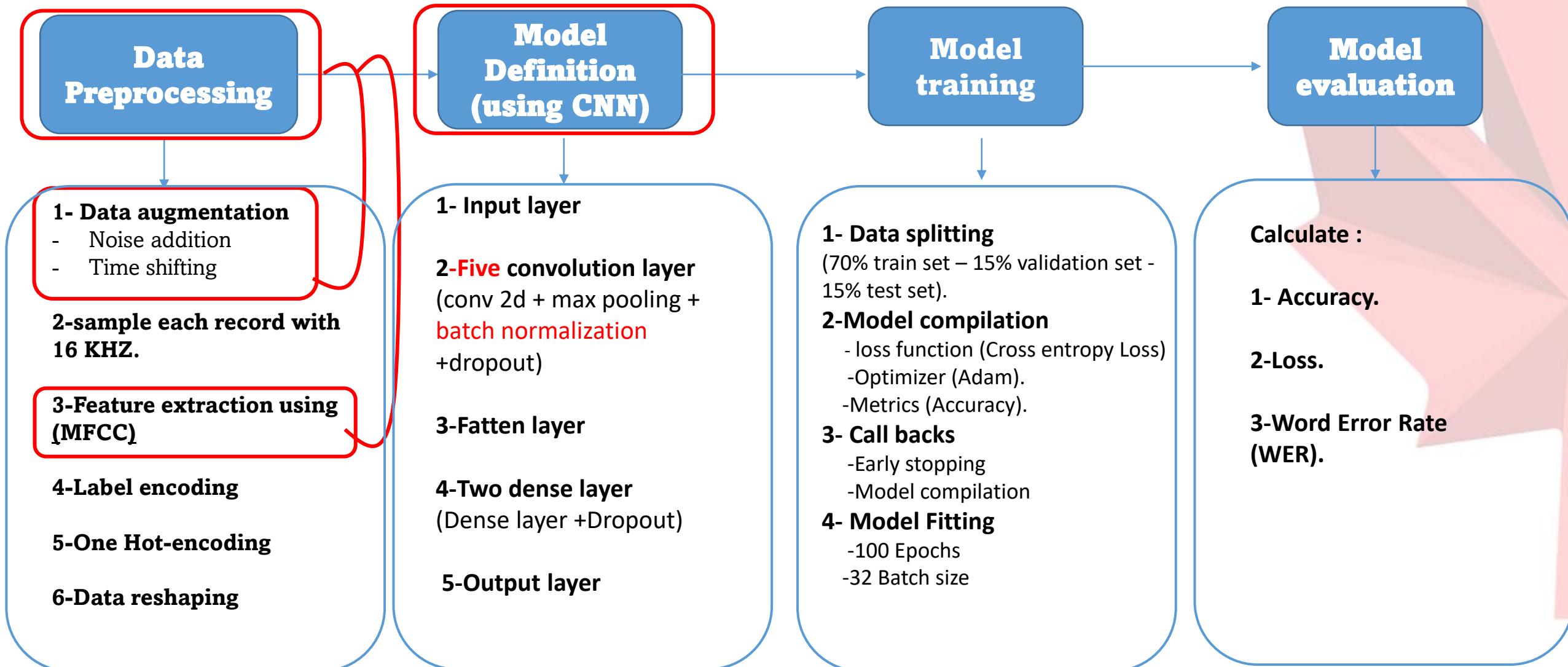
- WER = 18.06%

Difference between loss of train and test is (-10.41%) this is still **over fitting**, but overfitting reduce from initial step.

**Problem**

## Third trained model

### Using (Data augmentation &MFCC& batch normalization )



## Third trained model

	Accuracy	Loss
Training	92.11%	26.52%
Validation	95.12%	16.37%
Test	95.08%	16.28%

**WER= 4.92%**

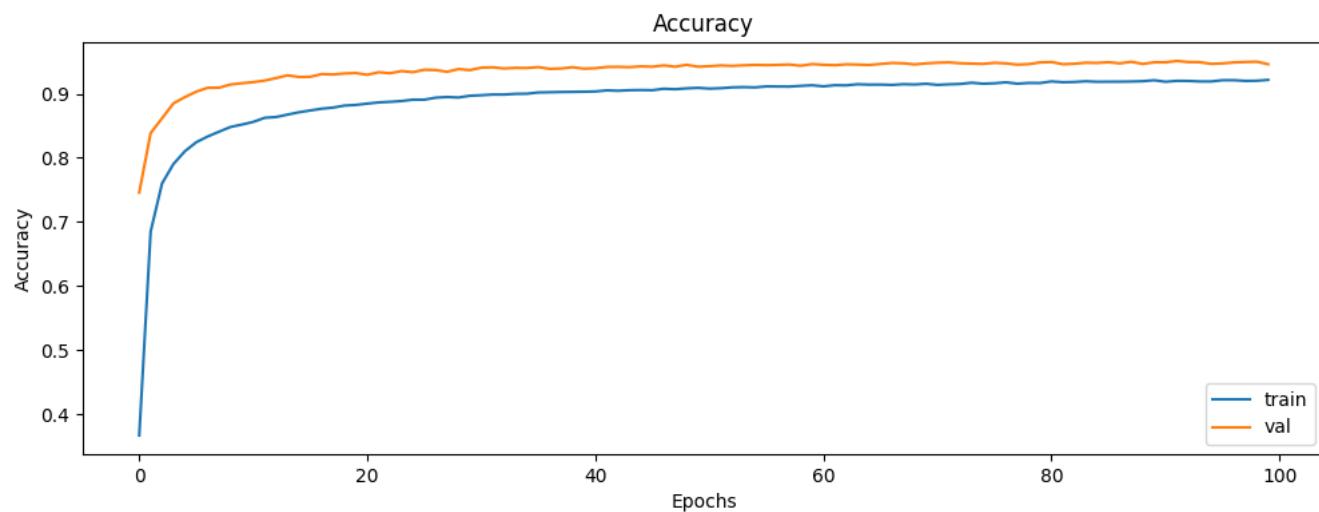
- Training accuracy **increased** by 8.75% 
- Testing accuracy **increased** by 13.14% 
- Training loss **decreased** by 26.68% 
- Testing loss **decreased** by 47.33% 
- WER **decreased** by 13.14% 

**Over fitting is solved**

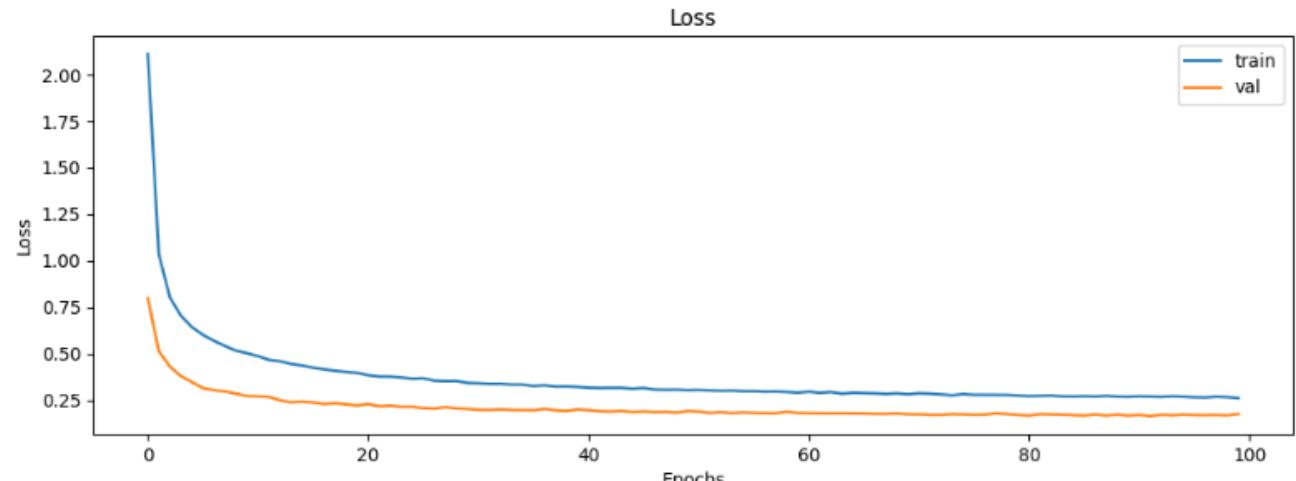
**It can easily Predict audio files.**

**Problem** is bad in real-Time

**Accuracy between validation and training data:**



**Loss between Validation and training data:**



## Fourth Trained model

### Whisper Model

#### whisper Description :

Whisper is a powerful speech-to-text model developed by OpenAI. It excels at transcribing speech and it challenges and stands with the following:

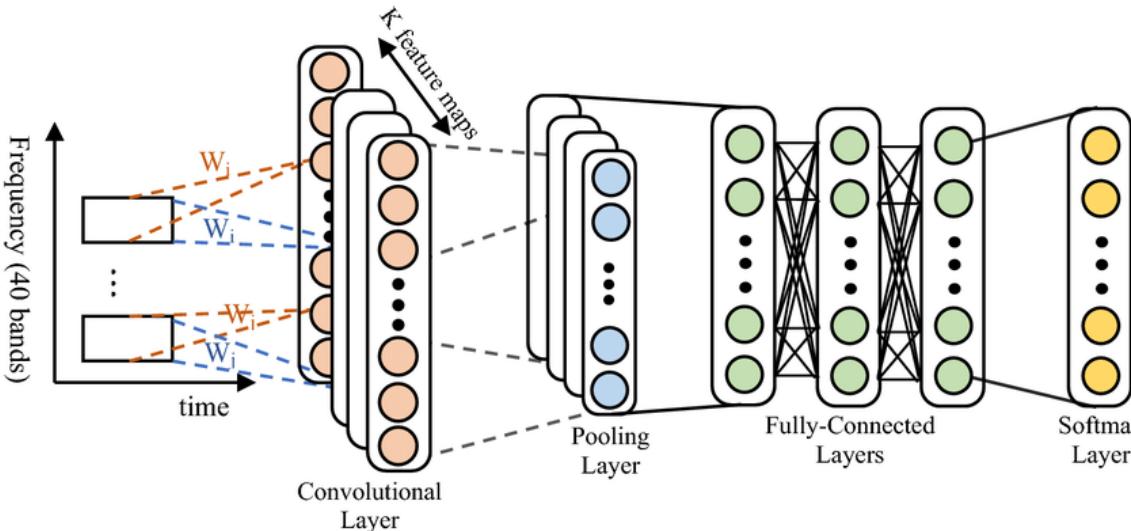
1. Low Speech Volume.
2. Background Noise.
3. Advanced Architecture.
4. Large-Scale Training.

#### Why whisper ?

- As in the previous trained model we didn't have a real-time prediction ,but based on the advanced architecture of whisper we do have real-time prediction and higher performance.

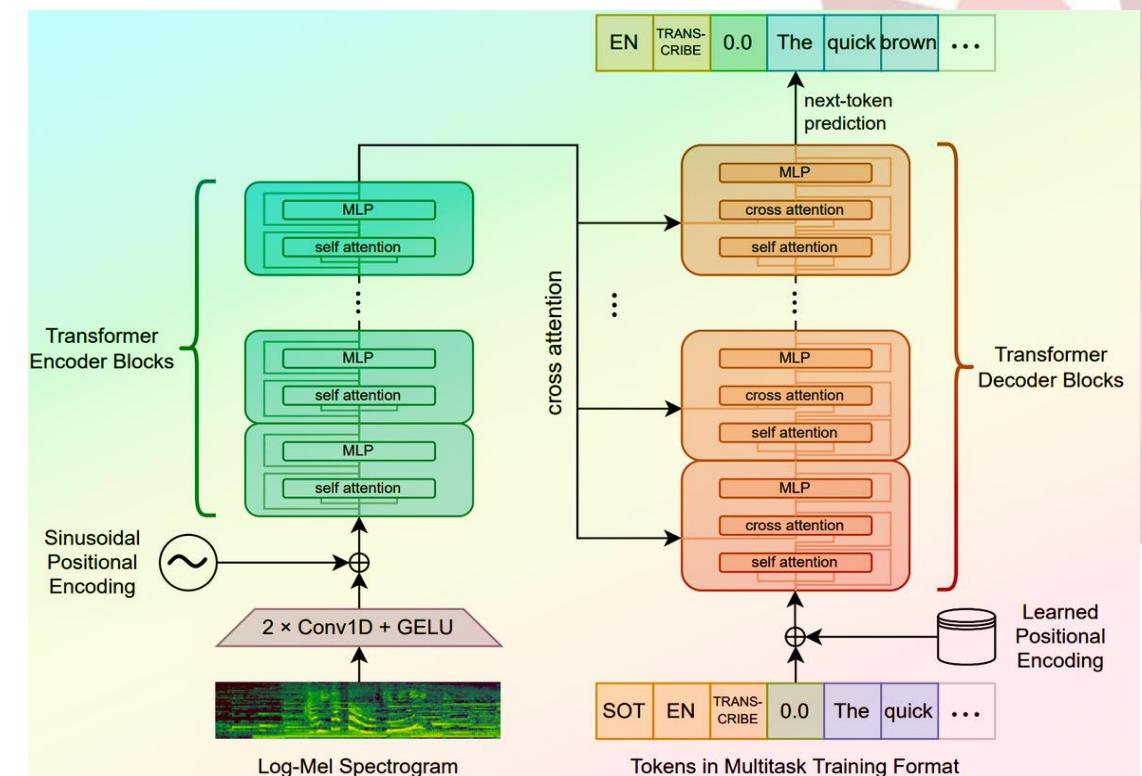
# Whisper Model

## CNN Architecture Vs Transformers Architecture



**CNN**

**Vs.**



**Transformers**

# Whisper Model

## Strengths & Weaknesses of CNN and Transformers

Feature	CNN	Transformer
<b>Strengths</b>	<ul style="list-style-type: none"> <li>- Efficient extraction of local features from spectrograms.</li> <li>- Faster training times.</li> <li>- Generally requires less data</li> </ul>	<ul style="list-style-type: none"> <li>- Captures long-range dependencies between words</li> <li>- Handles complex sentence structures</li> <li>- Often achieves state-of-the-art accuracy.</li> </ul>
<b>Weaknesses</b>	<ul style="list-style-type: none"> <li>- May struggle with long-range dependencies</li> <li>- Less effective with complex sentences</li> </ul>	<ul style="list-style-type: none"> <li>- Computationally expensive to train and run</li> <li>- Requires larger amounts of training data</li> </ul>
<b>Ideal for</b>	<ul style="list-style-type: none"> <li>- Limited datasets</li> <li>- Faster training</li> </ul>	<ul style="list-style-type: none"> <li>- Complex speech with long-range dependencies</li> <li>- High accuracy requirements</li> </ul>

# Whisper Model

## Result of training and testing data using whisper:

Training Loss	Epoch	Step	Testing Loss	Testing Accuracy
0.33	1.0	780	0.0272	0.9938
0.0002	2.0	1560	0.0420	0.9876
0.0001	3.0	2340	0.0487	0.9913
0.0011	4.0	3120	0.0789	0.9802
0.0001	5.0	3900	0.0915	0.9851
0.0014	6.0	4680	0.1017	0.9839
0.0	7.0	5460	0.0993	0.9888
0.0	8.0	6240	0.0694	0.9913
0.0	9.0	7020	0.0760	0.9926
0.0	10.0	7800	0.0842	0.9901

## Performance Results:

- Training Accuracy: 100%
- Test Accuracy: 99.01%
- Training Loss: 0%
- Testing Loss: 8.42%
- WER: 0.99%

- Training accuracy increased by 7.89%
- Testing accuracy increased by 3.93%
- Training loss decreased by 26.52%
- Testing loss decreased by 7.86%
- WER decreased by 3.93%

# Whisper Model

## Comparison between all models

Parameter	First Model	Second Model	Third model	Whisper Model
Training accuracy	83.36%	83.36%	92.11%	100%
Test accuracy	84.55%	81.94%	95.08%	99.01
Training loss	53.20%	53.20%	26.52%	0%
Testing loss	51.9%	63.61%	16.28%	8.42%
WER	20.45%	18.06%	4.92%	0.99%

# Whisper Model

## **Choosing the Right Model for our project:**

From the previous comparisons between architectures of models & models parameters we find:

- ❖ **Transformer architecture is more suitable for our project for the following reasons:**

1. Complex speech with long-range dependencies
2. High accuracy requirements

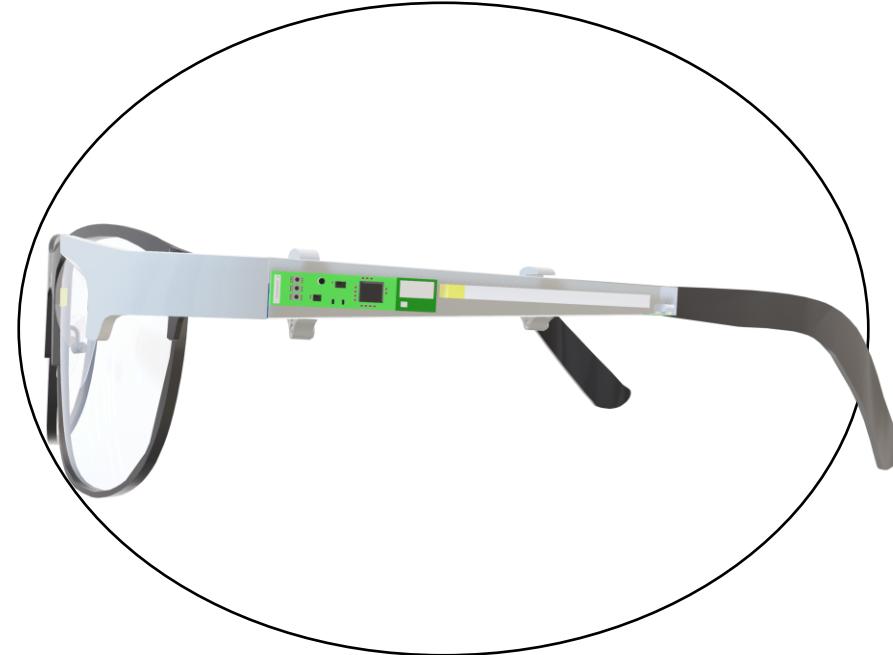
- ❖ **Whisper as a model based on Transformer Architecture it has:**

1. Higher accuracy.
2. Smaller loss.
3. Real-Time Prediction

Part 3



# Controller



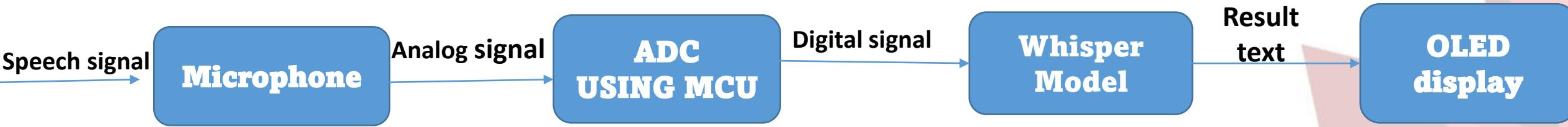
## ❖ Objective

- Develop Transcribe Glass capable of capturing audio, converting it to text, and displaying it in real-time

## ❖ KPIs

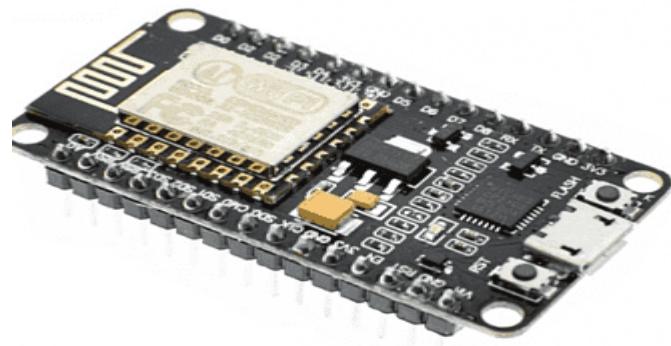
- lightweight
- Efficient audio capture
- Seamless integration with Wi-Fi for data transmission.
- Low Transmission Latency

## Process of controller of the system



### ESP32:

- A powerful microcontroller with a dual-core processor.
- Offers Wi-Fi and Bluetooth communication interfaces.
- Low power consumption.



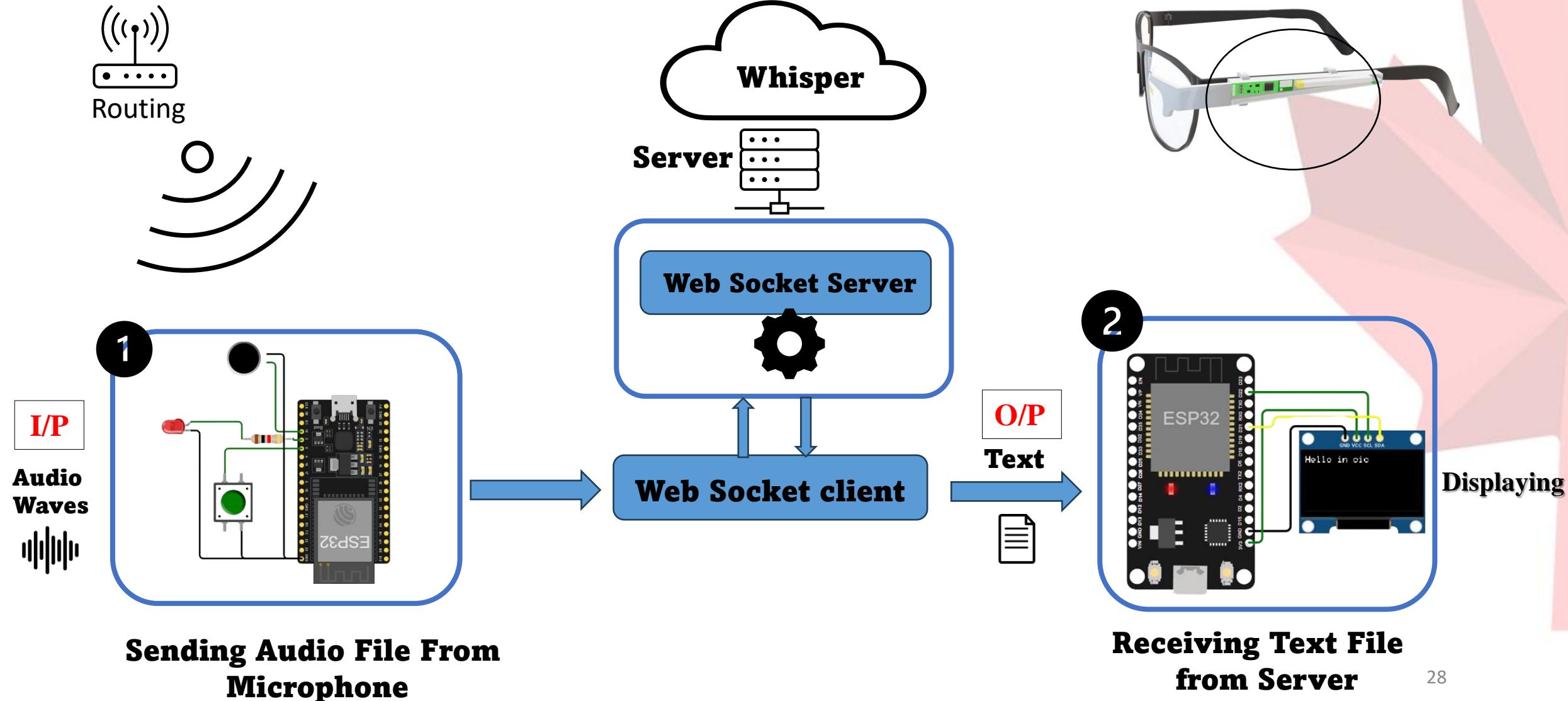
## Problem

- Limiting processing power
- Limited memory

# Solution



## Esp 32 controlling The processes



Part 3

# Displaying



## ❖ Objective

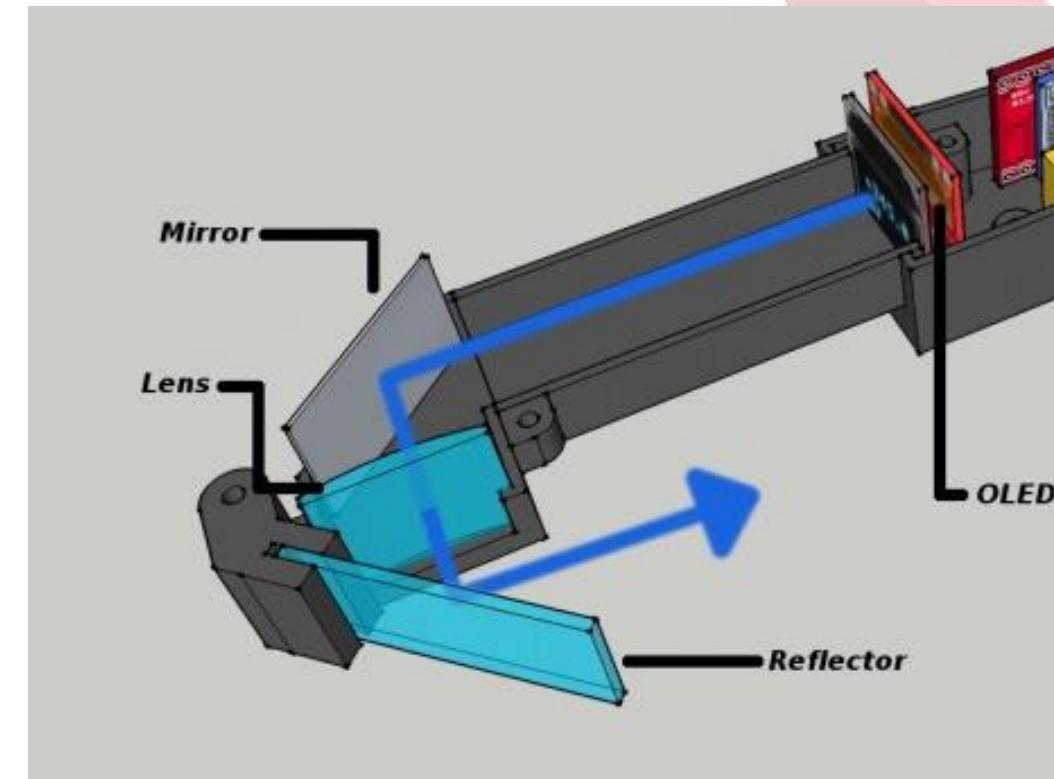
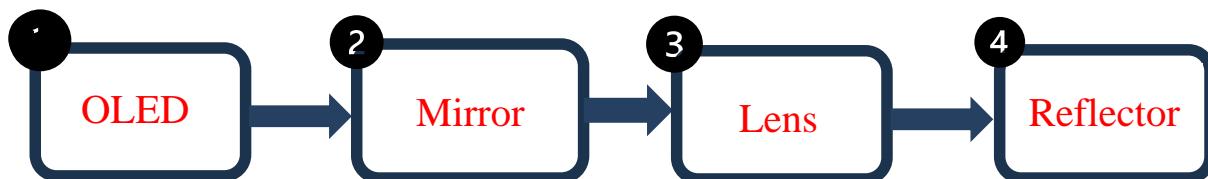
- Display texts so that users able to read
- Integration with other system blocks

## ❖ KPIs

- Small display size
- Avoid blurry vision

# Displaying Process

- Our system for displaying consists of **OLED , mirror , lens and reflector**



# Displaying

## 1 OLED (Organic light emitting diode)

### ❖ Size

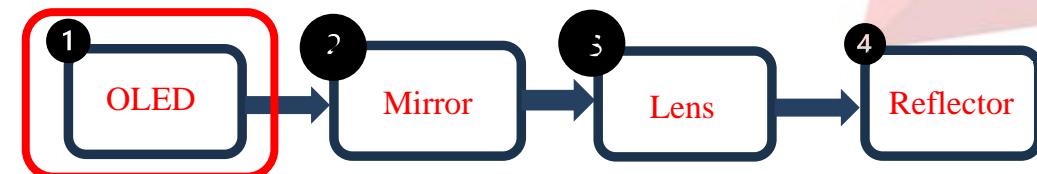
- in size ranging from 75 to 1 inch (**small size**)

### ❖ Low Power Consumption

- Since OLED pixels only use power when lit, they are very efficient, especially when displaying mostly static text like transcripts. This extends battery life in our transcribe glasses.

### ❖ High Contrast

- pixels light up individually, resulting in perfect blacks. This creates excellent contrast for text, making it easier to read transcripts on the glasses

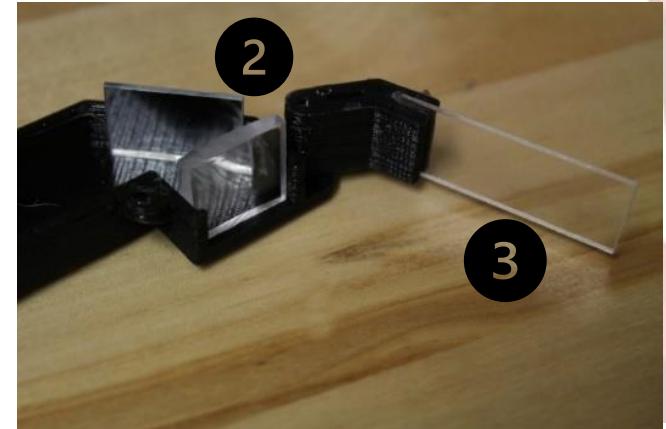


# Displaying

## 2 ♦ Mirror

Reflects light or image from the display (such as OLED) towards the lens.

The mirror is fixed at an angle of 45 degrees, bringing the image to the lens.



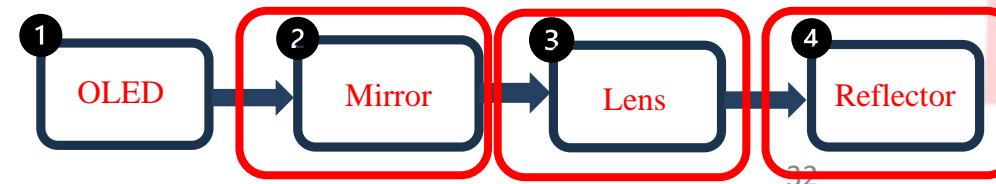
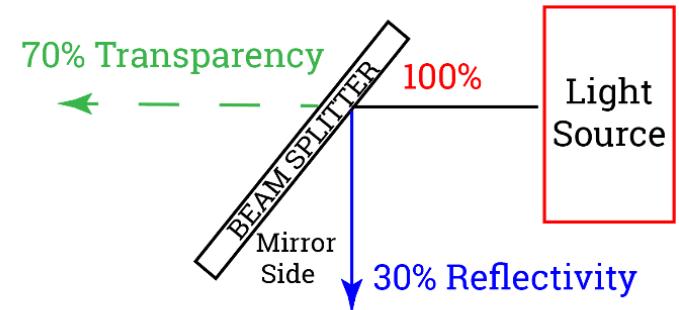
## 3 ♦ Lens

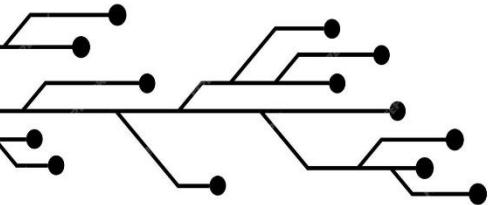
- Focus Range (25cm+) for human
- Acrylic Plano-Convex Lens 27-30cm
- Magnification (x3).

## 4 ♦ Beam splitter

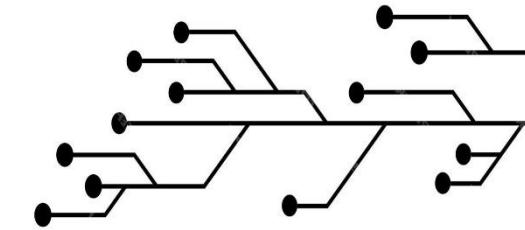
(The Bridge Between Screen and Eye)

- **Function:** This critical component acts like a partial mirror. It splits light rays coming from the display.
  - **Light Path:** One portion of the light passes through to the user's eye directly.
- Virtual Image Creation: The other portion reflects off the beam splitter, creating a virtual image of the magnified text.





# Supercapacitors



# Supercapacitors Overview

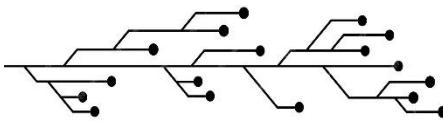
➤ **Supercapacitor** : is an electro chemical capacitor that can store significantly more energy compared to traditional capacitors.



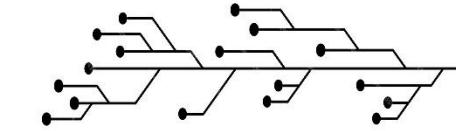
## ➤ Requirements for powering the project

- Fast charging Time (High power Density)
- Slow Discharge Time (High energy Density)
- Light Weight
- Long Life Cycle

Parameter	Supercapacitors	Capacitors	Batteries
<b>Energy Storage</b>	W-sec of energy	W-sec of energy	W-Hr of energy
<b>Charge Method</b>	voltage across terminals i.e. from a battery	voltage across terminals i.e. from a battery	current & voltage
<b>Power Delivered</b>	rapid discharge, linear or exponential voltage decay	rapid discharge, linear or exponential voltage decay	constant voltage over long time period
<b>Charge/Discharge Time</b>	msec to sec	psec to msec	1 to 10 hrs
<b>Form Factor</b>	small	small to large	large
<b>Weight</b>	1-2g	1g to 10kg	1g to >10kg
<b>Energy Density</b>	1 to 5Wh/kg	0.01 to 0.05Wh/kg	8 to 600Wh/kg
<b>Power Density</b>	High, >4000W/kg	High, >5000W/kg	Low, 100-3000W/kg
<b>Operating Voltage</b>	2.3V – 2.75V/cell	6V – 800V	1.2V - 4.2V/cell
<b>Lifetime</b>	>100k cycles	>100k cycles	150 to 1500 cycles
<b>Operating Temp</b>	-40 to +85°C	-20 to +100°C	-20 to +65°C

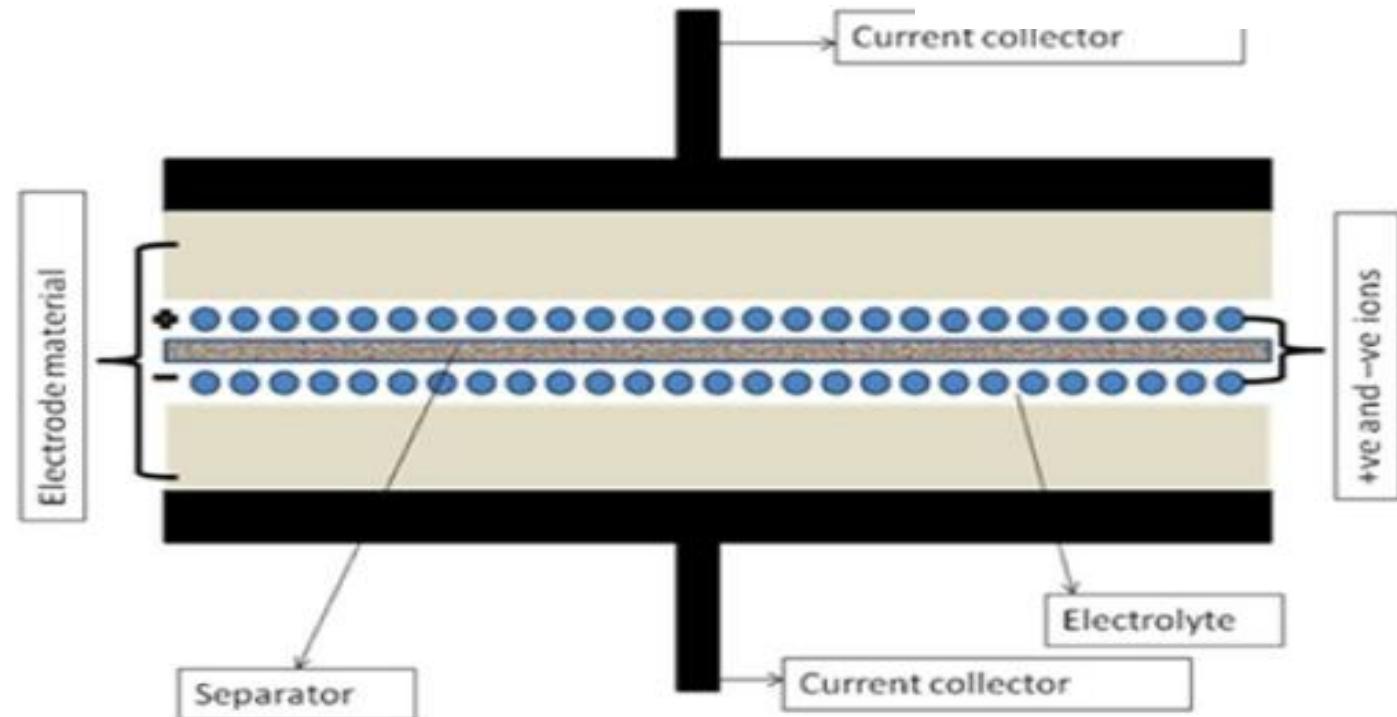


# Supercapacitors Overview

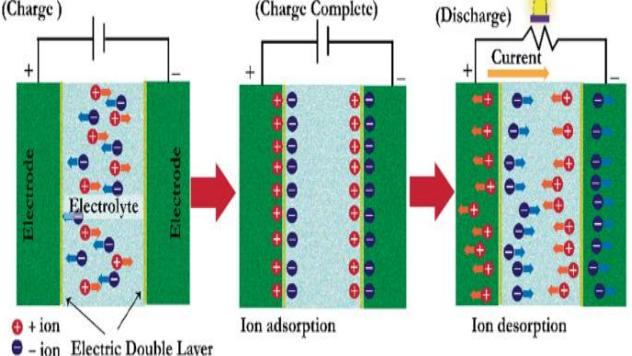
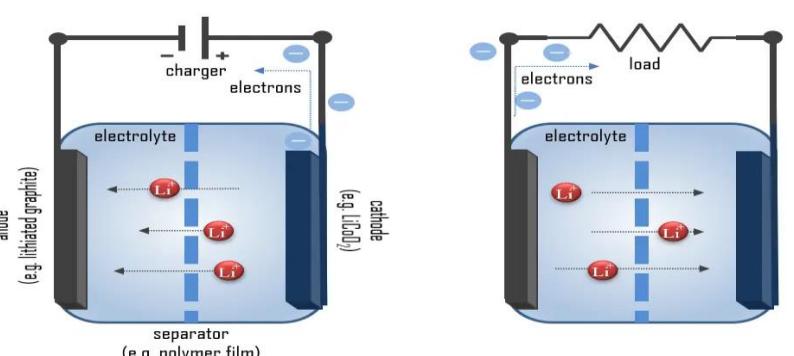
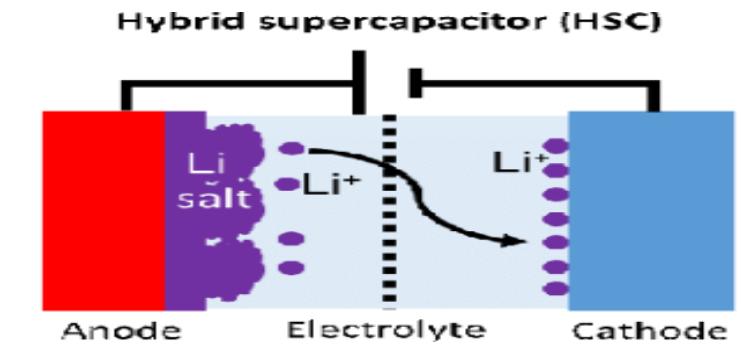


## □ Component Of Supercapacitor

1. Electrodes
2. Electrolyte
3. Separator
4. Current Collectors



## Types of Supercapacitors

	<b>Electric Double Layer Capacitor</b>	<b>Pseudocapacitor</b>	<b>Hybrid Supercapacitors</b>
<b>Definition</b>	stores energy in an electric field formed by the separation of charge	A type of electrochemical capacitor that stores electrical through a faradaic processes	Capacitor act like electrode (EDLC) is anode. Battery act like electrode (PC) is cathode
<b>Types</b>	Types of material : <ul style="list-style-type: none"><li>• Activated carbon.</li><li>• Graphene</li></ul>	Types of material : <ul style="list-style-type: none"><li>• Conducting polymer.</li><li>• Metal oxide.</li></ul>	<ul style="list-style-type: none"><li>• <b>Composite electrodes</b></li><li>• <b>Asymmetric electrodes</b></li><li>• <b>Battery-capacitor electrodes</b></li></ul>
<b>Adv.</b>	<ul style="list-style-type: none"><li>• High power density.</li><li>• Fast charging.</li><li>• Long cycle life.</li></ul>	<ul style="list-style-type: none"><li>• Higher energy density.</li><li>• Longer cycle life.</li><li>• Environmentally friendly</li></ul>	<ul style="list-style-type: none"><li>• High power density.</li><li>• Higher energy density.</li><li>• Longer cycle life.</li></ul>
<b>Disadv.</b>	<ul style="list-style-type: none"><li>• Low energy density</li><li>• High cost</li></ul>	<ul style="list-style-type: none"><li>• Lower energy density than batteries.</li><li>• Higher cost</li></ul>	<ul style="list-style-type: none"><li>• Higher cost</li></ul>
<b>Charge mechanism</b>	<ul style="list-style-type: none"><li>• Electrostatic.</li></ul> 	<ul style="list-style-type: none"><li>• electrochemical reaction.</li></ul> 	<ul style="list-style-type: none"><li>• electrostatic and electrochemical.</li></ul> 

# Choosing supercapacitor for our project

## Hybrid Supercapacitors

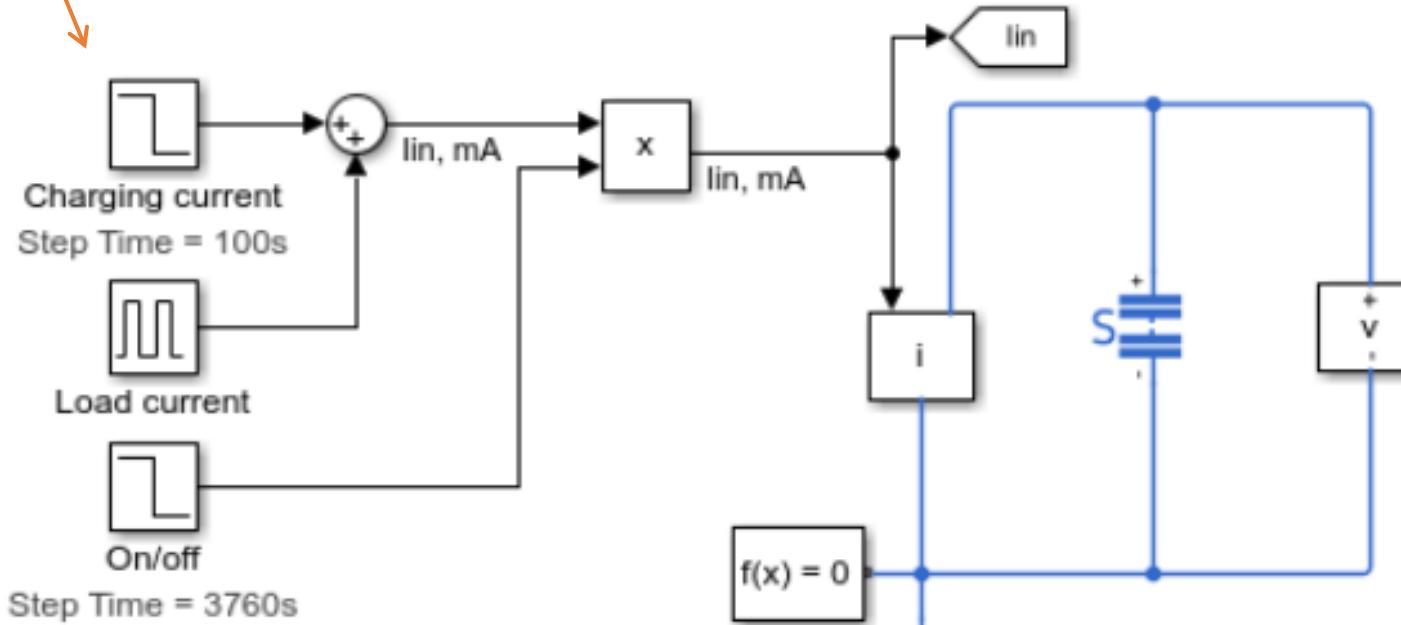
Feature	Composite Electrode	Asymmetric Electrode	Battery-Capacitor Hybrid
Electrode composition	Both use combinations	Different for positive and negative	Similar to asymmetric, often with "battery-like" material
Primary storage mechanism	Combined EDLC and pseudocapacitance	Primarily pseudocapacitance (potentially with some EDLC)	Similar to asymmetric
Complexity	Moderate	High	High
Potential energy density	Moderate-high	High	High
Power density	High	High	High
Cost	Moderate	Potentially higher	Potentially higher

➤ **Best solution :**

Battery-capacitor electrodes in hybrid super capacitor

# Supercapacitor Charging and Discharging Simulation

# 100mA for 100sec



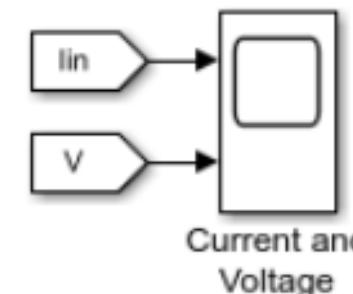
Step Time = 3760s

**50mA for 1sec**

Every 50 sec  
For

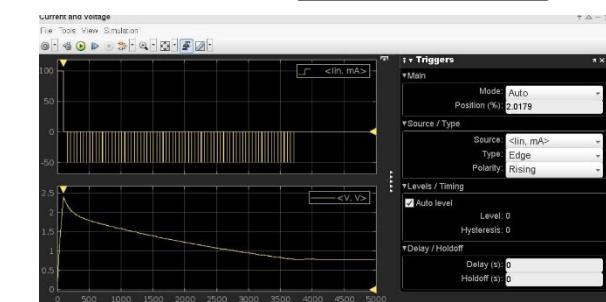
For 1 hour

## Model 1



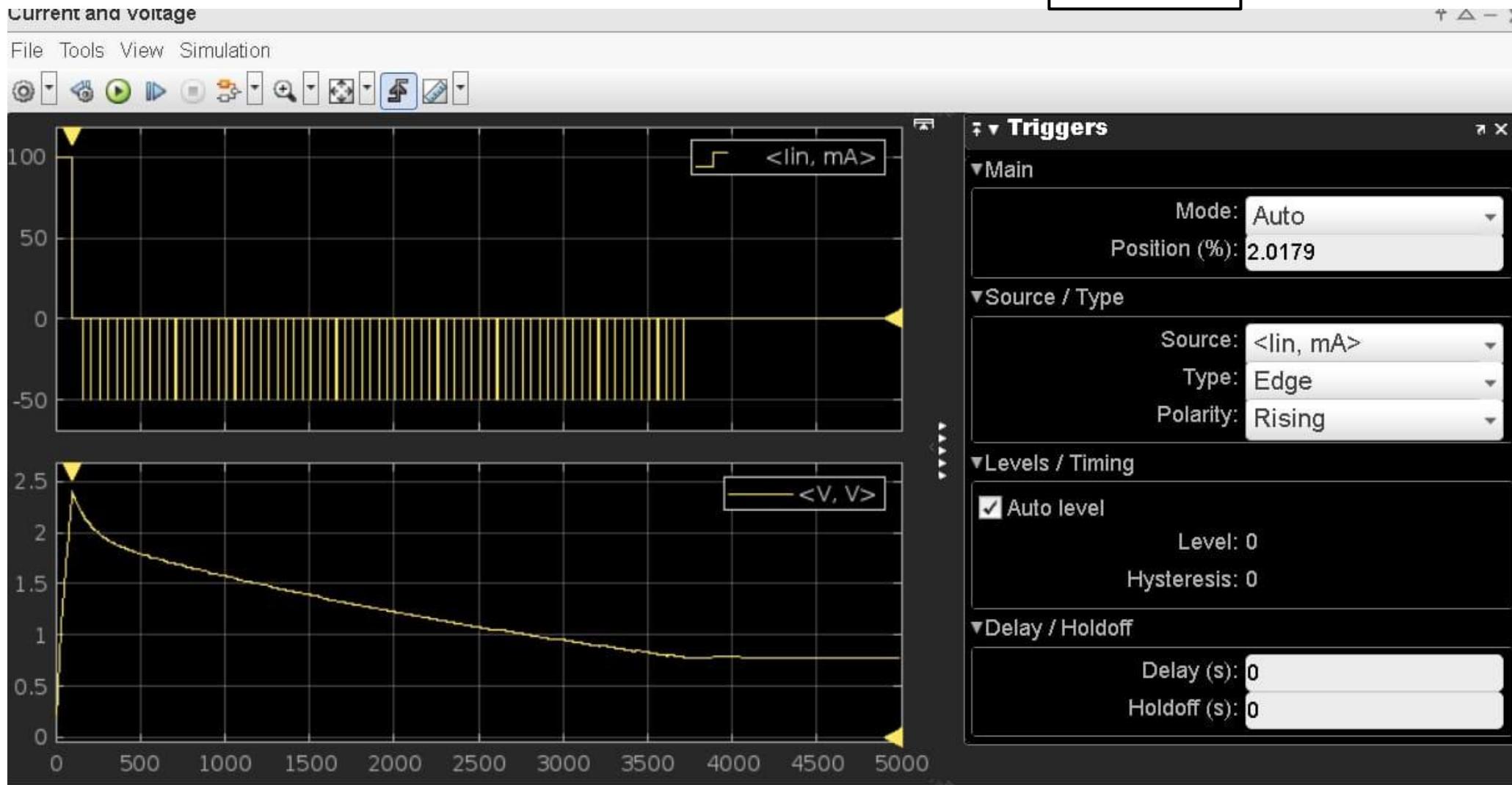
## Current and Voltage

output1

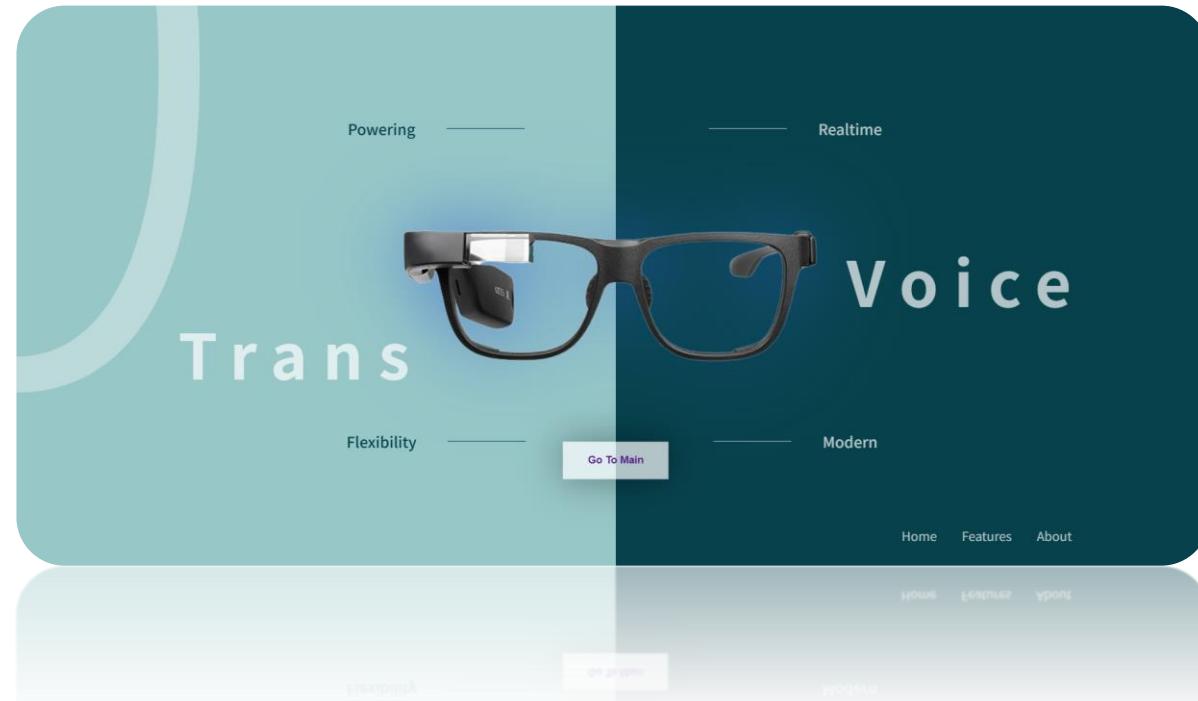


# Supercapacitor Charging and Discharging Behavior

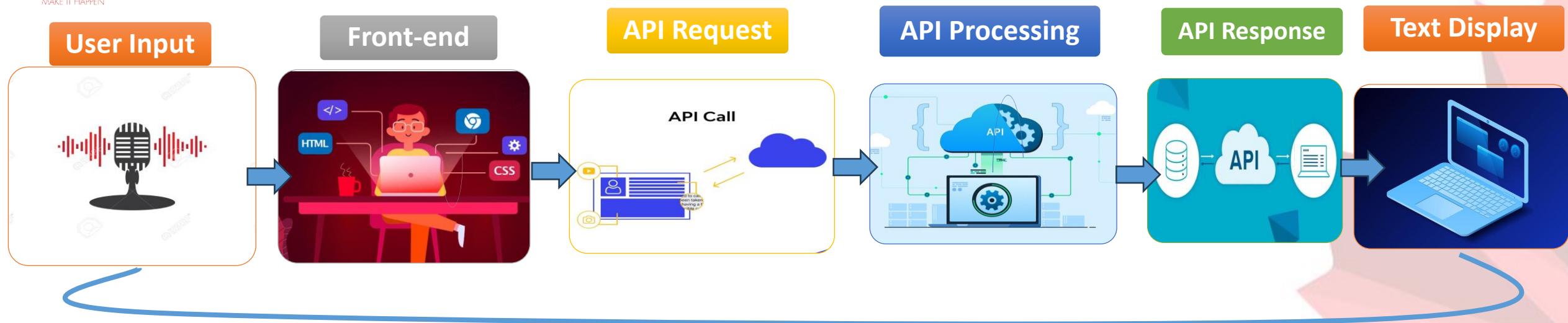
## output1



# Interface



# Interfacing



## Website Architecture

**Front-End:** HTML, CSS, and JavaScript .

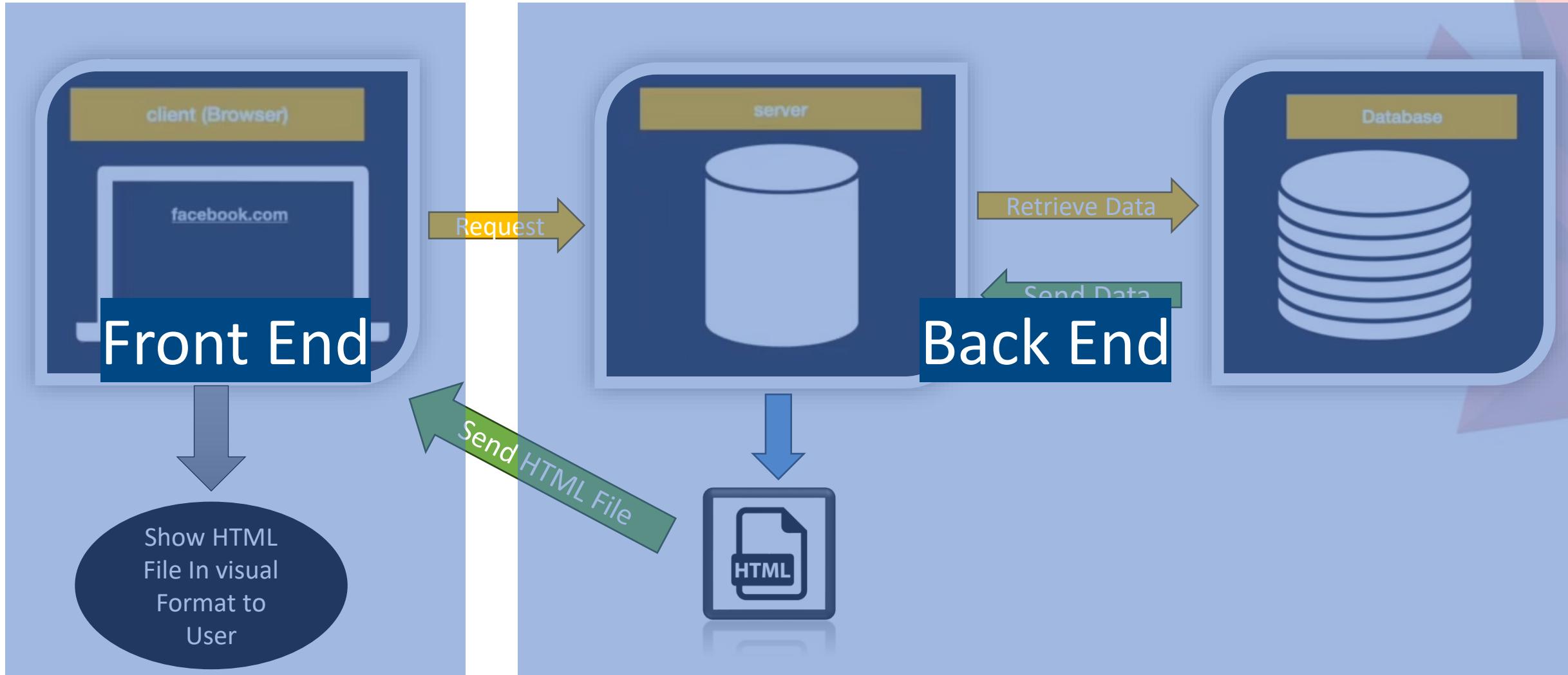
**Back-End:** Robust server-side scripting languages .

**Whisper API:** Integration with a leading speech recognition API.

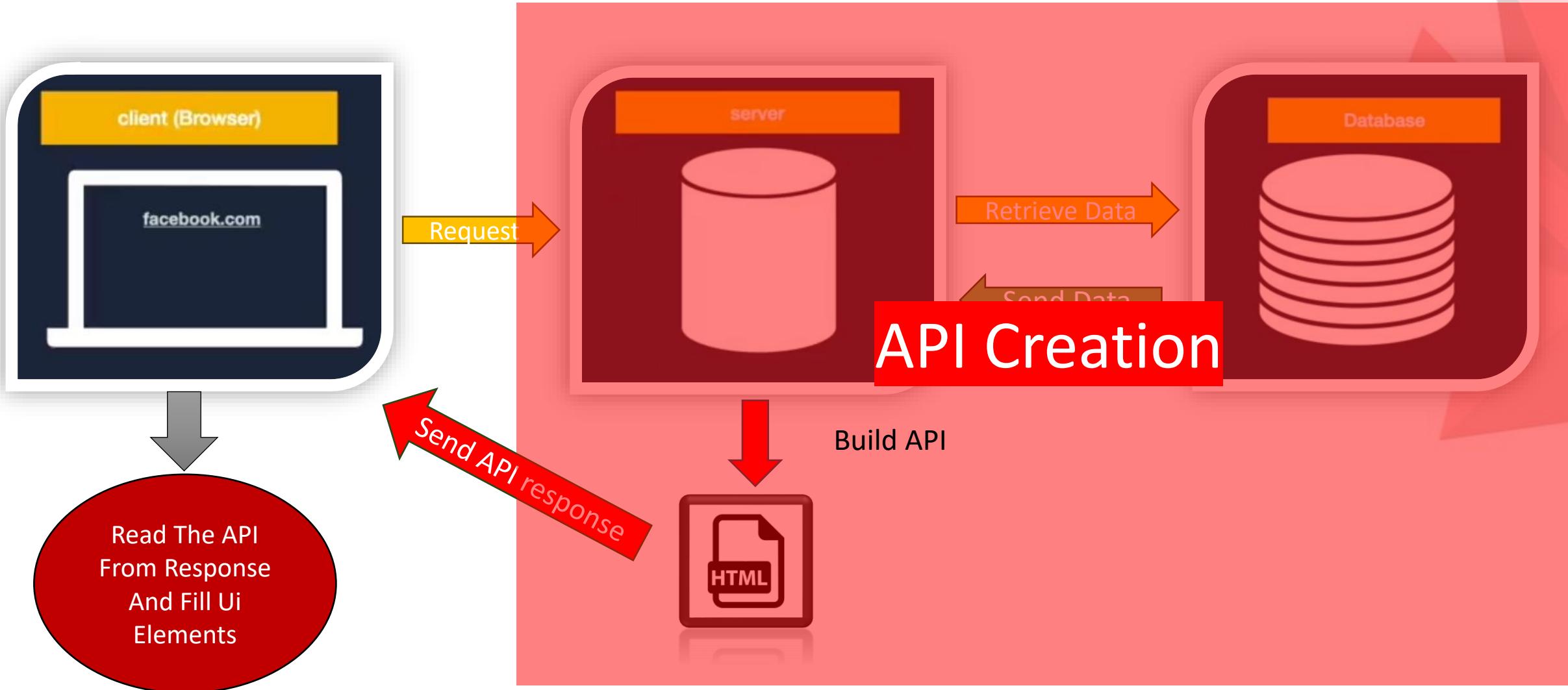
**Google Translation API:** Integration with a reliable machine translation API (e.g., Google Translate API)

**Database:** Secure database (e.g., MySQL, PostgreSQL) to store user data and website preferences (optional).

# Site Communication Technique



# API Generation



# Whisper & Translation Architecture

## Integrate Whisper & Google Translate API:

### 1. Obtain API Key:

1. Register on the Whisper API platform to get your API key.
2. Keep your API key secure as it authenticates your requests.

### 2. Set Up Backend:

1. Use Node.js and Express.js to set up the server.
2. Install required packages: express, axios for making HTTP requests.

### 3. Create API Route: By Postman

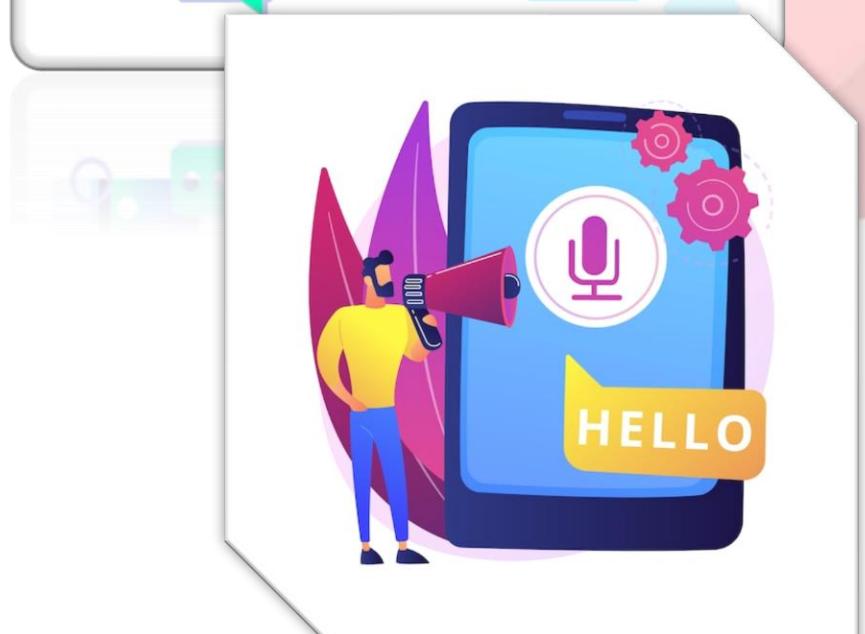
1. Whisper :Define a route in Express to handle audio file uploads and processing.
2. Google API :Define a route in Express to handle text translation requests.

### 4. Handle Audio File Upload:

1. Use multer middleware to handle audio file uploads from the frontend.

### 5. Send Audio to Whisper API:

1. Make an HTTP POST request to Whisper API with the audio file.
2. Receive and process the text response.

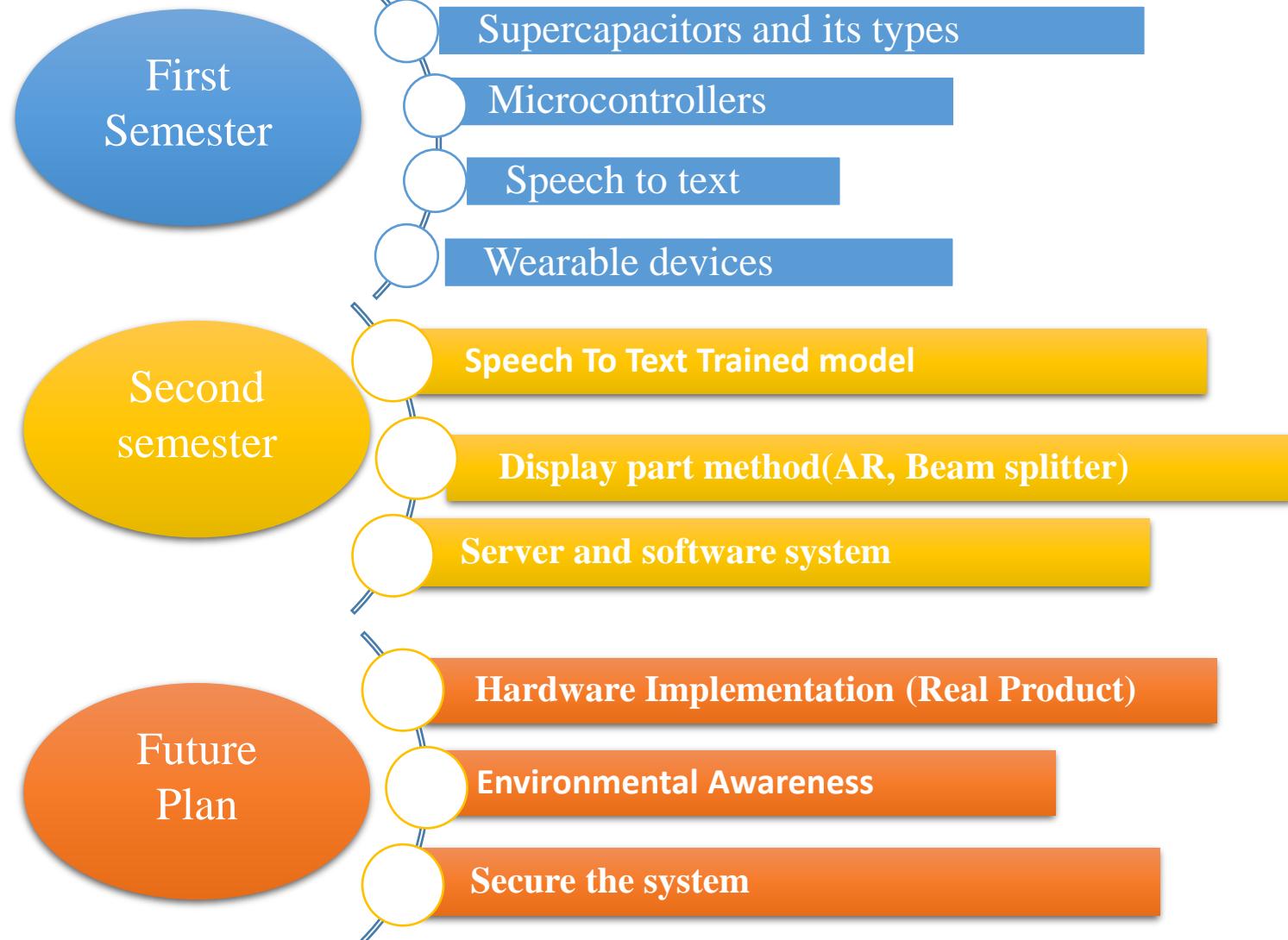


Website



# TIMELINE AND FUTURE PLAN

Collecting survey paper and research about:





CANADIAN  
INTERNATIONAL COLLEGE  
MAKE IT HAPPEN

---

Thank you

