

JUNE 2025

# Real-Time Human Activity Detection and Classification System

Final Project - Artificial Intelligence I

PRESENTED TO

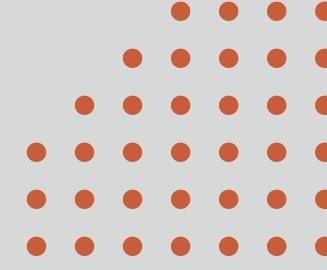
Milton Orlando Sarria  
Paja

PRESENTED BY

Andrés Pino  
Jhonatan Castaño



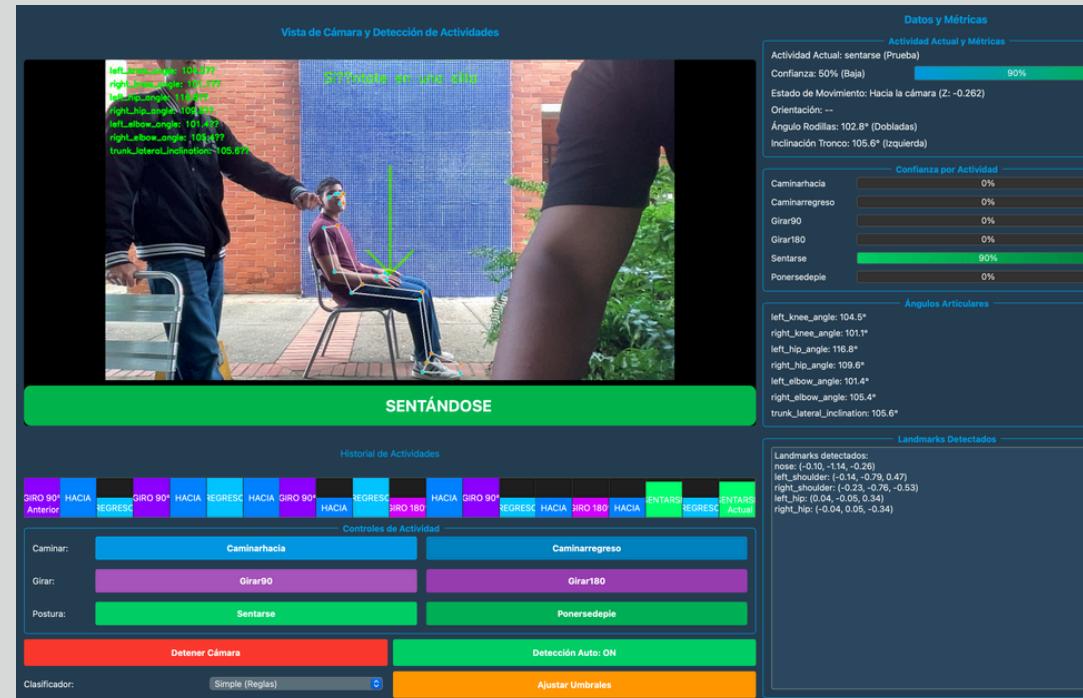
# What is this project?



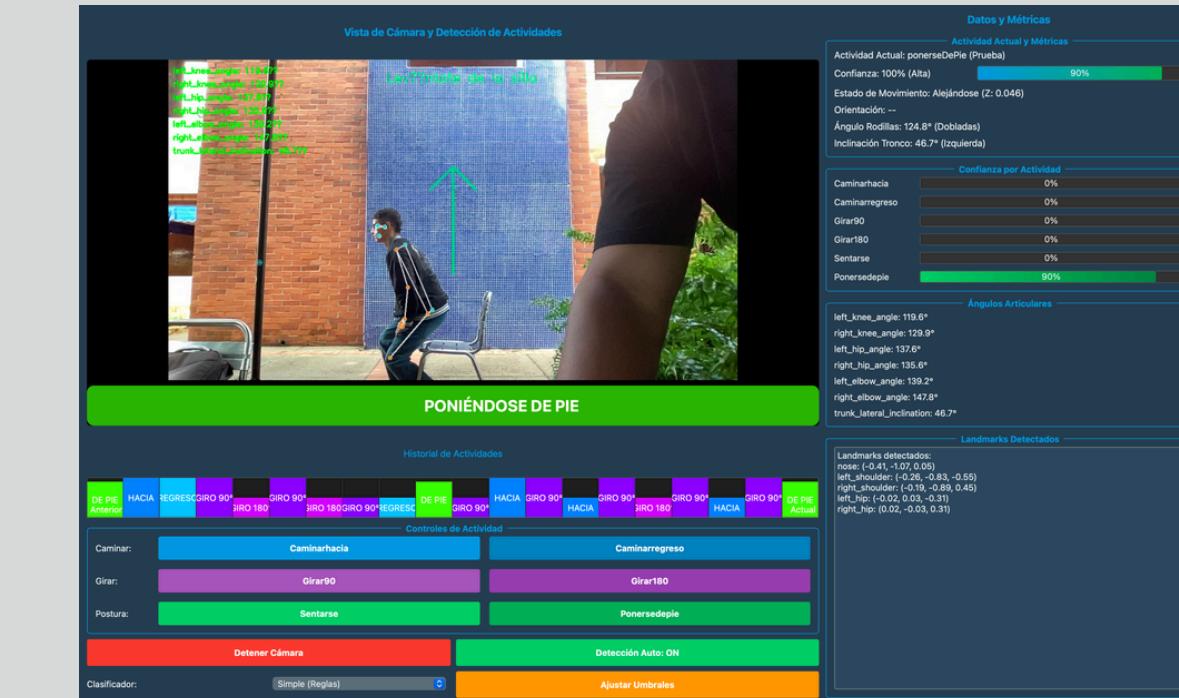
Teaching a Machine to Understand Human Action

We have built a system that uses a standard webcam to identify 6 basic human activities in real time.

One of the most representative screenshots of the system in action



Sitting Down



Standing up

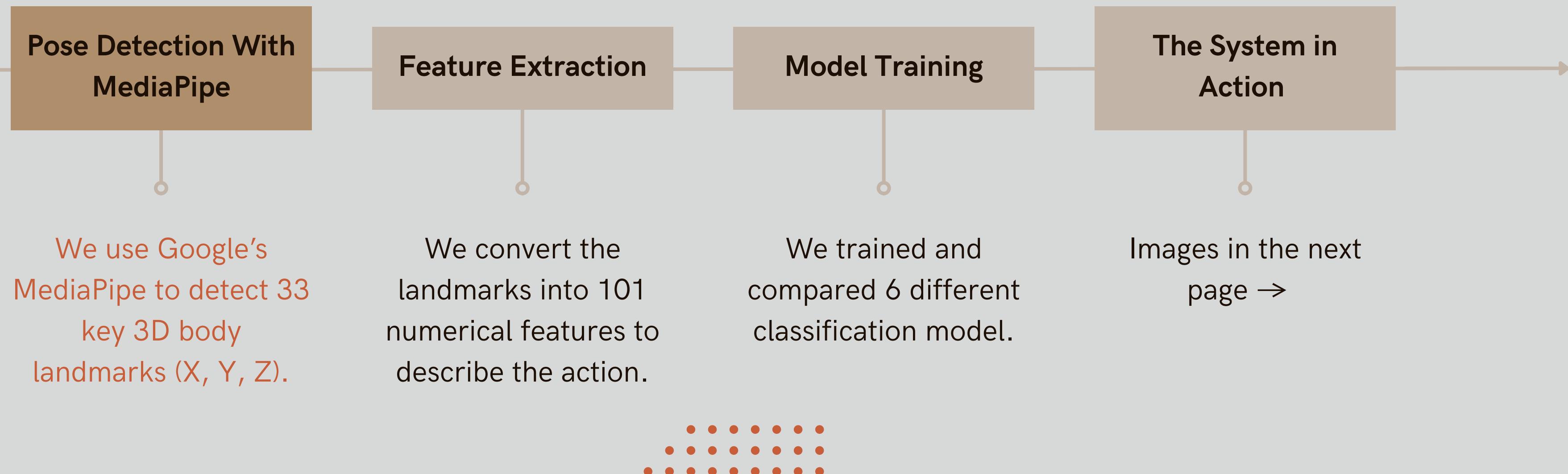
# Our Goal: The 6 core Activities

- Walking Towards 
- Walking Away 
- Turning 90° 
- Turning 180° 
- Sitting Down 
- Standing Up 



# System Architecture: How Does It Work?

Our Methodology: From Pixels to Predictions → A 4-step workflow to process video and classify the activity



# Pose Detection With MediaPipe

The foundational step: where the system converts raw video pixels into structured information.

1. A single frame is captured from the video stream (either from a file or a live webcam).
2. This frame is fed into Google's MediaPipe Pose model.
3. MediaPipe uses a pre-trained ML model to detect and return 3D coordinates (x, y, z) and a visibility score for 33 specific landmarks on the human body.



# Pose Detection With MediaPipe

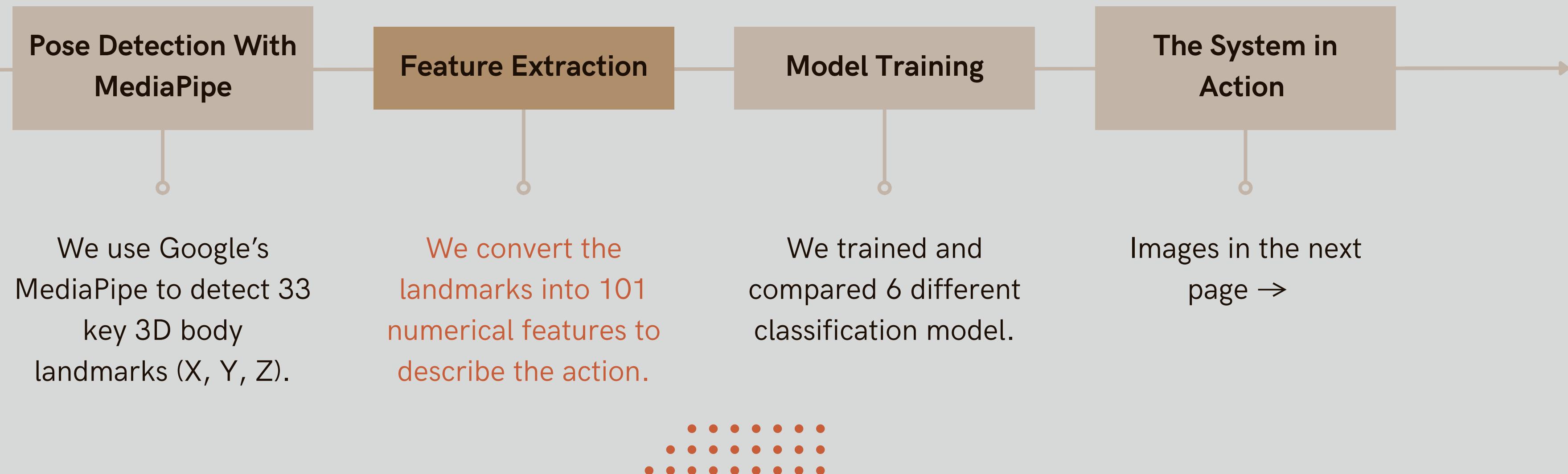
These are the standalone Python scripts:

- **extract\_video\_features.py**: This is the first step.
- **generate\_report\_plots.py**: To understand the data that we've just extracted before training a model.
- **train\_classifier.py**: This is the core of the machine learning process.
- **run\_pipeline.py**: This script is an orchestrator.



# System Architecture: How Does It Work?

Our Methodology: From Pixels to Predictions → A 4-step workflow to process video and classify the activity



# Feature Extraction

How does it work?

**Geometric Features:** This step takes the raw 3D coordinates of the 33 landmarks between joints. For example, by using the coordinates of the shoulder, elbow and wrist landmarks, it calculates the elbow angle. Similarly, it computes angles for the knees, hips, and shoulders.

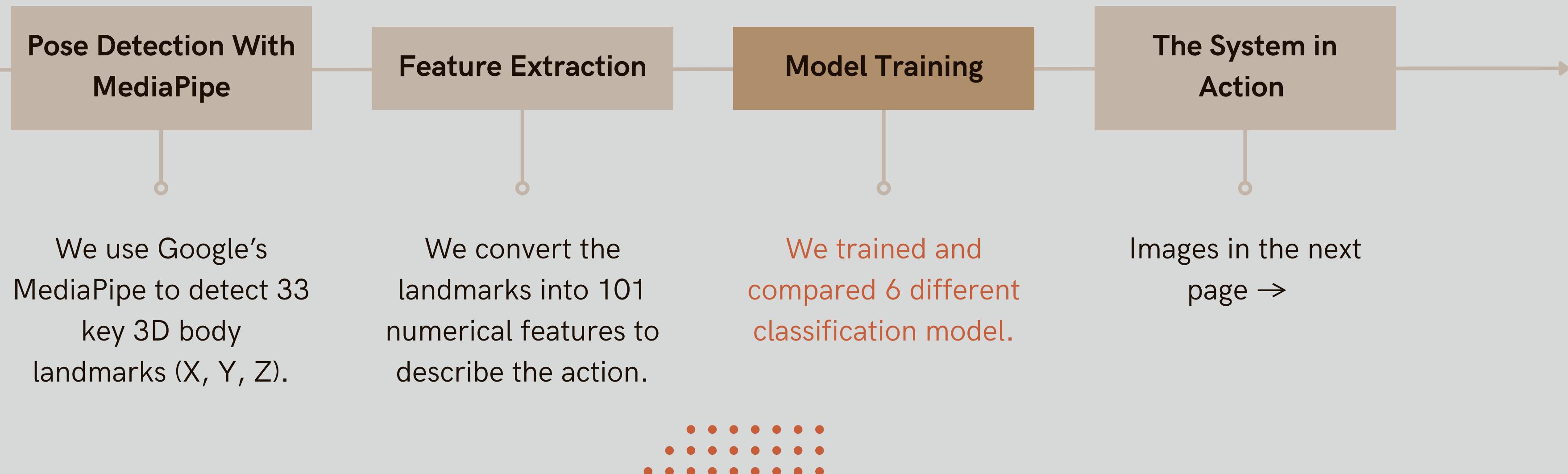
**Relational Features:** It calculates distances and proportions, like the width between the shoulders or the hips.

**Temporal (Dynamic) Features:** The system compares a landmark's position in the current frame to its position in the previous frame to calculate its velocity. This captures the dynamics of the movement.



# System Architecture: How Does It Work?

Our Methodology: From Pixels to Predictions → A 4-step workflow to process video and classify the activity



# Model Training

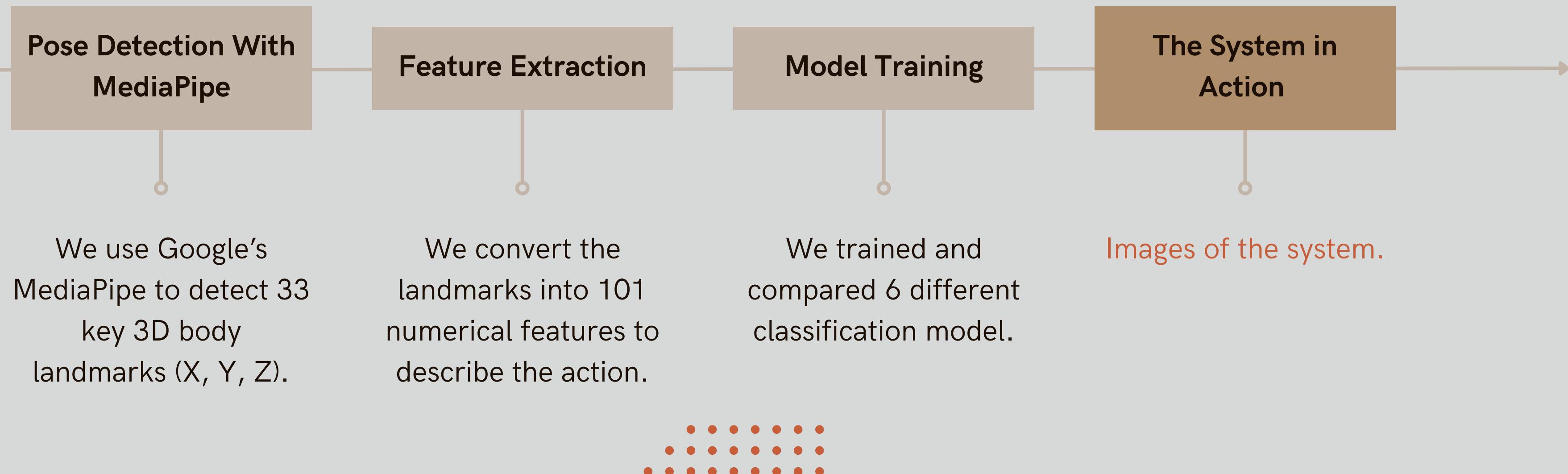
The “learning” phase

- 1. Data Collection:** This creates a large dataset where each row is a 101-feature vector, and each row is labeled with the correct activity (e.g., “sitting\_down”).
- 2. Training:** The labeled dataset is fed into various classification algorithms (like logistic Regression, SVM, Random Forest). Each algorithm analyzes the data and learns the statistical patterns.
- 3. Evaluation & Selection:** After training, the models are tested on a portion of the data they’ve never seen before. The model that performs the best is selected and saved to a file.



# System Architecture: How Does It Work?

Our Methodology: From Pixels to Predictions → A 4-step workflow to process video and classify the activity



# The System in Action

The live application where the trained model is used to make predictions on a real-time video from a webcam.

1. **Capture:** It grabs a new frame from the webcam.
2. **Detect & Extract:** Instantly performs Step 1 and Step 2 on that single frame to generate a 101-feature vector.
3. **Predict:** This new, unlabeled feature vector is fed into the loaded Logistic Regression Model (from Step 3).
4. **Display:** The model outputs a prediction. The application's GUI then displays this predicted activity on the screen, and immediately gets ready for the next frame.



# The System in Action

## Walking Towards

## Walking Away

## Turning 90°

Vista de Cámara y Detección de Actividades

Datos y Métricas

Actividad Actual: caminarHacia

Confianza: 100% (Alta) 100%

Estado de Movimiento: Hacia la cámara (Z: -2.042)

Orientación: --

Ángulo Rodillas: 119.7° (Dobladas)

Inclinación Tronco: 90.1° (Izquierda)

Confiabilidad: 115.577

left\_knee\_angle: 115.577  
right\_knee\_angle: 123.977  
left\_hip\_angle: 154.499  
right\_hip\_angle: 171.377  
left\_elbow\_angle: 122.079  
right\_elbow\_angle: 127.499  
trunk\_lateral\_inclination: 90.119

Confianza por Actividad

Caminarhacia 100%

Caminaregreso 0%

Girar90 0%

Girar180 0%

Sentarse 0%

Ponersedepie 0%

Ángulos Articulares

left\_knee\_angle: 115.5°  
right\_knee\_angle: 123.9°  
left\_hip\_angle: 154.4°  
right\_hip\_angle: 171.5°  
left\_elbow\_angle: 122.0°  
right\_elbow\_angle: 127.4°  
trunk\_lateral\_inclination: 90.1°

Historial de Actividades

Landmarks Detectados

Controles de Actividad

Camar: Caminarchacia Caminaregreso

Girar: Girar90 Girar180

Postura: Sentarse Ponersedepie

Detener Cámara Detención Auto: ON

Ajustar Umbrales

Clasificador: Simple (Reglas)

ACTIVIDAD: CAMINANDO HACIA CÁMARA

Vista de Cámara y Detección de Actividades

Datos y Métricas

Actividad Actual: caminarRegreso

Confianza: 100% (Alta) 100%

Estado de Movimiento: Alejándose (Z: 0.442)

Orientación: --

Ángulo Rodillas: 169.7° (Estendidas)

Inclinación Tronco: 101.4° (Izquierda)

Confiabilidad: 169.399

left\_knee\_angle: 169.399  
right\_knee\_angle: 156.699  
left\_hip\_angle: 175.773  
right\_hip\_angle: 153.899  
left\_elbow\_angle: 153.899  
right\_elbow\_angle: 154.579  
trunk\_lateral\_inclination: 101.477

Confianza por Actividad

Caminarhacia 0%

Caminaregreso 100%

Girar90 0%

Girar180 0%

Sentarse 0%

Ponersedepie 0%

Ángulos Articulares

left\_knee\_angle: 169.3°  
right\_knee\_angle: 156.6°  
left\_hip\_angle: 175.7°  
right\_hip\_angle: 153.9°  
left\_elbow\_angle: 153.8°  
right\_elbow\_angle: 154.5°  
trunk\_lateral\_inclination: 101.4°

Historial de Actividades

Landmarks Detectados

Controles de Actividad

Camar: Caminarchacia Caminaregreso

Girar: Girar90 Girar180

Postura: Sentarse Ponersedepie

Detener Cámara Detención Auto: ON

Ajustar Umbrales

Clasificador: Simple (Reglas)

ACTIVIDAD: CAMINANDO ALEJÁNDOSE

Vista de Cámara y Detección de Actividades

Datos y Métricas

Actividad Actual: girar90

Confianza: 100% (Alta) 100%

Estado de Movimiento: Alejándose (Z: 0.337)

Orientación: --

Ángulo Rodillas: 172.7° (Estendidas)

Inclinación Tronco: 63.8° (Izquierda)

Confiabilidad: 172.172

left\_knee\_angle: 172.172  
right\_knee\_angle: 172.172  
left\_hip\_angle: 156.5°  
right\_hip\_angle: 156.6°  
left\_elbow\_angle: 155.6°  
right\_elbow\_angle: 144.3°  
trunk\_lateral\_inclination: 63.8°

Confianza por Actividad

Caminarhacia 0%

Caminaregreso 0%

Girar90 100%

Girar180 0%

Sentarse 0%

Ponersedepie 0%

Ángulos Articulares

left\_knee\_angle: 172.1°  
right\_knee\_angle: 172.3°  
left\_hip\_angle: 156.5°  
right\_hip\_angle: 156.6°  
left\_elbow\_angle: 155.6°  
right\_elbow\_angle: 144.3°  
trunk\_lateral\_inclination: 63.8°

Historial de Actividades

Landmarks Detectados

Controles de Actividad

Camar: Caminarchacia Caminaregreso

Girar: Girar90 Girar180

Postura: Sentarse Ponersedepie

Detener Cámara Detención Auto: ON

Ajustar Umbrales

Clasificador: Simple (Reglas)

ACTIVIDAD: GIRANDO 90°

# The System in Action

## Turning 180°

Vista de Cámara y Detección de Actividades

Datos y Métricas

Actividad Actual: girar180

Confianza: 50% (Baja)

Estado de Movimiento: Hacia la cámara (Z: -0.634)

Orientación: --

Ángulo Rodillas: 132.7° (Dobladas)

Inclinación Tronco: 106.1° (Izquierda)

Confiabilidad: 100%

Historial de Actividades

Controles de Actividad

Landmarks Detectados

Actividad Actual: GIRANDO 180°

left\_knee\_angle: 132.4°  
right\_knee\_angle: 133.1°  
left\_hip\_angle: 162.8°  
right\_hip\_angle: 172.4°  
left\_elbow\_angle: 156.1°  
right\_elbow\_angle: 105.2°  
trunk\_lateral\_inclination: 106.1°

Landmarks detectados:  
nose: (-0.3, -1.61, -0.63)  
left\_shoulder: (0.41, -0.16, -0.48)  
right\_shoulder: (0.15, -1.13, 0.49)  
left\_hip: (0.11, -0.05, -0.30)  
right\_hip: (-0.11, 0.05, 0.30)

Caminar: Caminahacia Caminaregreso  
Girar: Girar90 Girar180  
Postura: Sentarse Ponersedepie  
Detener Cámara Detección Auto: ON  
Ajustar Umbral

Clasificador: Simple (Reglas)

## Sitting Down

Vista de Cámara y Detección de Actividades

Datos y Métricas

Actividad Actual: sentarse (Prueba)

Confianza: 90% (Alta)

Estado de Movimiento: Hacia la cámara (Z: -0.262)

Orientación: --

Ángulo Rodillas: 132.7° (Dobladas)

Inclinación Tronco: 105.6° (Izquierda)

Confiabilidad: 100%

Historial de Actividades

Controles de Actividad

Landmarks Detectados

SENTÁNDOSE

left\_knee\_angle: 104.5°  
right\_knee\_angle: 101.1°  
left\_hip\_angle: 116.8°  
right\_hip\_angle: 105.6°  
left\_elbow\_angle: 101.4°  
right\_elbow\_angle: 105.4°  
trunk\_lateral\_inclination: 105.6°

Landmarks detectados:  
nose: (-0.3, -1.61, -0.63)  
left\_shoulder: (0.41, -0.16, -0.48)  
right\_shoulder: (0.15, -1.13, 0.49)  
left\_hip: (0.04, 0.05, 0.34)  
right\_hip: (-0.04, 0.05, -0.34)

Caminar: Caminahacia Caminaregreso  
Girar: Girar90 Girar180  
Postura: Sentarse Ponersedepie  
Detener Cámara Detección Auto: ON  
Ajustar Umbral

Clasificador: Simple (Reglas)

## Standing Up

Vista de Cámara y Detección de Actividades

Datos y Métricas

Actividad Actual: ponerseDePie (Prueba)

Confianza: 90% (Alta)

Estado de Movimiento: Alejándose (Z: 0.046)

Orientación: --

Ángulo Rodillas: 124.8° (Dobladas)

Inclinación Tronco: 46.7° (Izquierda)

Confiabilidad: 100%

Historial de Actividades

Controles de Actividad

Landmarks Detectados

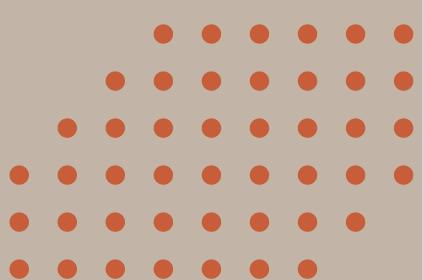
PONIÉNDOSE DE PIE

left\_knee\_angle: 119.6°  
right\_knee\_angle: 122.9°  
left\_hip\_angle: 137.6°  
right\_hip\_angle: 130.0°  
left\_elbow\_angle: 130.2°  
right\_elbow\_angle: 119.5°  
trunk\_lateral\_inclination: 46.7°

Landmarks detectados:  
nose: (-0.41, -1.07, 0.05)  
left\_shoulder: (-0.16, 0.23, -0.47)  
right\_shoulder: (0.18, -0.23, -0.48, -0.53)  
left\_hip: (0.02, 0.03, -0.31)  
right\_hip: (0.02, -0.03, 0.51)

Caminar: Caminahacia Caminaregreso  
Girar: Girar90 Girar180  
Postura: Sentarse Ponersedepie  
Detener Cámara Detección Auto: ON  
Ajustar Umbral

Clasificador: Simple (Reglas)



# Results: Model Comparison

Modelo	Accuracy	Precision	Recall	F1-Score
RandomForest	0.6541	0.6563	0.6541	0.6551
SVM	0.6649	0.6599	0.6649	0.6603
<b>LogisticRegression</b>	<b>0.6946</b>	<b>0.6930</b>	<b>0.6946</b>	<b>0.6933</b>
GradientBoosting	0.6757	0.6757	0.6757	0.6757
KNN	0.6649	0.6564	0.6649	0.6598
XGBoost	0.6703	0.6705	0.6703	0.6703

The Logistic Regression model outperformed the others across key metrics

- Mejor Modelo: **LogisticRegression**
  - Accuracy: ~69.46%
  - Precision: ~69.30%
  - Recall: ~69.46%
  - F1-Score: ~69.33%

# Main Achievements

- **End-to-End System:** From video capture to final prediction. ✓
- **Practical Application of AI:** Real-world use of Computer Vision and ML techniques. ✓
- **Interactive GUI:** A functional graphical interface for live demonstration. ✓
- **Automated Pipeline:** Scripts to facilitate reproducibility and experimentation. ✓





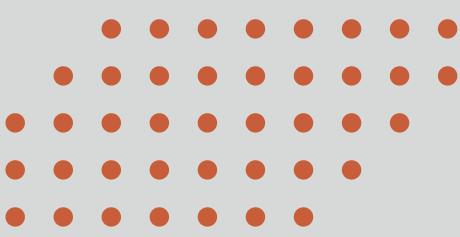
# Conclusions & Future Work

## Conclusions:

- Successfully built a functional, real-time HAR system.
- Logistic Regression proved to be the most effective model for our dataset.

## Future Work:

- Expand the dataset (more data, more variety).
- Explore Deep Learning models (LSTMs, Transformers).
- Add more activities to recognize.



# Thank You!

