Desk-Mate

Abdulrahman Shashtari & Mohammad Wahbeh

Jun 2025

Acknowledgment

All praise and gratitude are due to Allah, the Lord of the Worlds, and may peace and blessings be upon the most honorable of prophets and messengers. By the grace and mercy of Allah, we have successfully and distinctively completed our first graduation project. We extend our heartfelt thanks to Allah Almighty for the countless blessings He has bestowed upon us. This project's success is a testament to His generosity and favor.

We extend our sincere gratitude to our beloved parents for their continuous support and encouragement throughout this journey. Their motivation and unwavering belief in us has given us the strength to persevere, work hard, and achieve our dreams. Your love will remain eternally in our hearts.

We also extend our sincere thanks to our esteemed supervisor, Dr. Luai Malhis, for his valuable support and guidance throughout every stage of this project. His insightful advice and constructive comments were essential to the project's success and development. Words cannot express how grateful we are for his continuous assistance, which has greatly enriched our knowledge and experience. Thank you, Dr. Luai, for your constant encouragement and dedication.

Contents

Abstraction

This project presents a Self-Balancing Pet Desk Robot, an interactive and intelligent robot designed for desktop environments. The robot is designed to be a pet-like companion that brings you comfort and keeps you company.

The robot moves on two wheels and use edge detection to prevent falling off desk table, also it has multiple interactive features, including an AI-powered voice assistant that processes voice commands via a microphone and responds accordingly. The robot can speak through a built in speaker. The robot expresses emotions through an OLED screen that displays dynamic eye animations. The Robot also reacts to physical interactions like when hitting or tapping his head he stops the movement and behave well, and show happiness when patted. Also, as a functional assistant, it includes a built-in flashlight, allowing it to act as a table lamp wen requested.

In addition, the robot is accessible via a WebSocket based web app that's optimized for mobile devices, offering multiple control modes. In Remote Mode, you can drive and move the robot manually. In Desk Mode the robot can move or stay calm on your desk.

By combining robotics, artificial intelligence, and emotional expressions, mobile connectivity, the Pet Desk Robot serves as both a companion and a functional desk assistant.

Introduction

Recent years have seen amazing developments in the field of robotics, with a growing focus on intelligent systems that can interact with humans in more emotionally attractive and natural ways. Demand for small, responsive, and emotionally intelligent companions that exceed conventional automation is rising as robots become more prevalent in daily life. The Self-Balancing Pet Desk Robot is a novel robot that can function in desktop settings while retaining its ability to balance, move, and interact.

This robot integrates several complex actions into a single, small device, in contrast to traditional robotic systems that focus on a single core function, like movement or stability. In addition to being able to balance on two wheels, it uses edge detection to navigate safely on a small surface area without falling. With AI-powered voice recognition, physical reactions to touch, emotional expression via an OLED screen, and multi-modal control via a web-based mobile interface, this robot is unique in its ability to interact with users.

Designed to provide comfort, functionality, and a sense of presence on the user's desk, the robot acts as a practical assistant and pet-like companion. When tapped, it stops moving; when petted, it expresses happiness; and when requested for light, it turns on a flashlight. This self-balancing platform's integration of emotional intelligence, responsive behavior, and real-time control is a major step toward more interactive and intimate human-robot interactions.

The Self-Balancing Pet Desk Robot's design, development, and features are examined in this report, with an emphasis on how unique it is in comparison to other balancing robots and how it might be used as an intelligent, emotionally intelligent desk companion.

Literature Review

Research and development on self-balancing robots has been ongoing, with a primary focus on using two-wheeled platforms to achieve dynamic stability. The majority of earlier research in this area focused on preserving vertical balance without adding intricate behaviors like movement or interaction. Usually used as examples of control algorithms, these systems are frequently restricted to static balancing or basic forward and backward motion in extremely controlled environments.

However, balancing and carrying out other tasks like controlled movement and user interaction at the same time present a very different challenge than balancing alone. Since the robot's center of gravity, speed, and direction must be changed in real time while maintaining stability, the incorporation of motion adds another level of complexity. The literature that is currently available hardly ever addresses this degree of complexity, particularly in small spaces like a desktop surface.

Additionally, robotics typically treats emotional expression, voice command responsiveness, and physical interaction handling as distinct domains. These domains are typically found in mobile robots that do not require constant balance adjustments or stationary assistants. It hasn't been thoroughly investigated how to combine all of these components into a single, mobile, self-balancing system.

By combining dynamic balancing, intelligent mobility, and emotional interaction into a small package, our Self-Balancing Pet Desk Robot fills these gaps. In contrast to earlier systems, our robot can move intelligently on a small surface in response to physical interactions and user commands, and it is not restricted to static balance. Voice control, dynamic eye animations for visual expression, and adaptive behavior based on user engagement are all made possible by artificial intelligence.

This combined method of balancing, movement, emotion, and interaction in a single robot is a new development in the field and a big step toward making robotics more approachable, entertaining, and useful in daily life.

Hardware Components

1. NEMA 17 Stepper Motors (x2)

NEMA 17 stepper motors provide precise rotational movement, crucial for maintaining the robot's balance and enabling smooth navigation. They operate in discrete steps, allowing for accurate control of position and speed through the step signals from the motor driver.

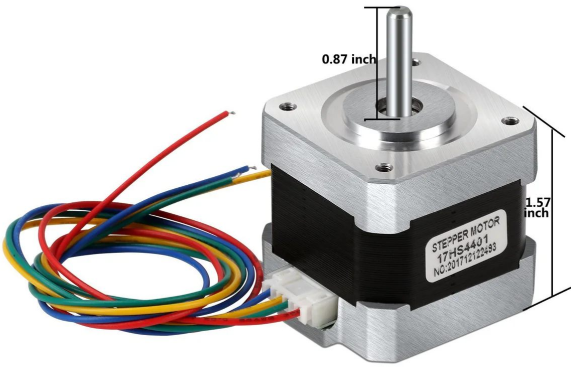

Figure 4.1: NEMA 17 Stepper Motor

2. A4988 Stepper Motor Drivers (x2)

Each A4988 driver controls one stepper motor by converting step and direction signals into precise voltage and current outputs. It supports microstepping for smoother motor rotation and is responsible for motor speed and torque control.

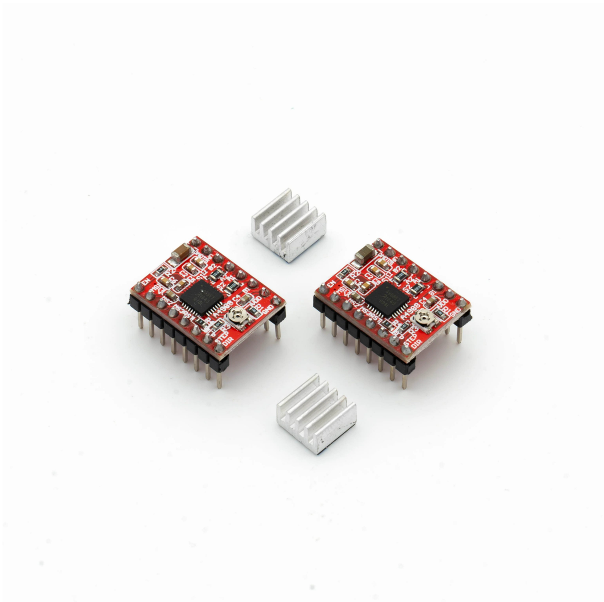

Figure 4.2: A4988 Motor Driver

3. MPU-6050 IMU Sensor

The MPU-6050 provides real-time tilt and motion detection using a 3-axis accelerometer and 3-axis gyroscope. It is essential for balance control in the robot through integration with the PID control algorithm.

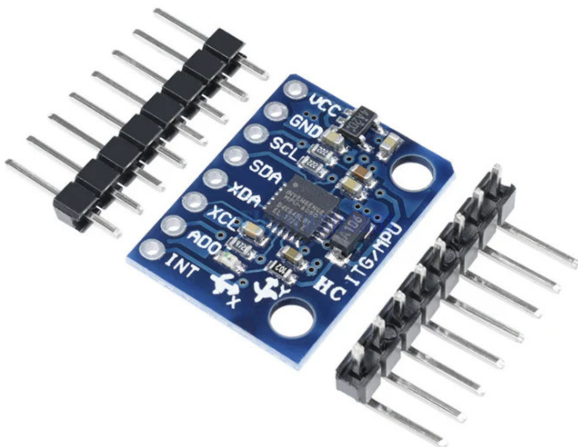

Figure 4.3: MPU-6050 IMU Sensor

4. OLED Screen

The OLED screen serves as the face of the robot, displaying animated eyes and emotional expressions. It visually responds to voice interaction and system states, enhancing user engagement.

Figure 4.4: OLED Display

5. IR Sensors (x2)

Two infrared sensors are mounted at the bottom edges of the robot to detect surface boundaries (like table edges) and prevent falls. They ensure the robot operates safely within its environment.

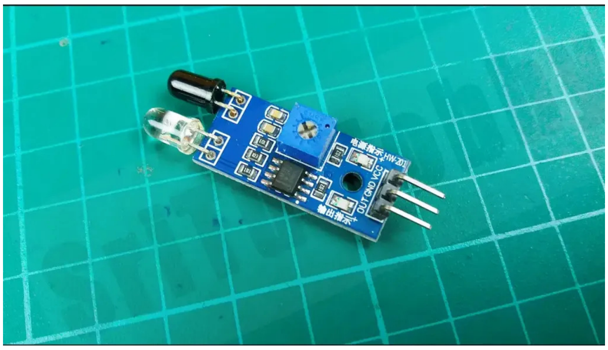

Figure 4.5: IR Edge Detection Sensor

6. Conductive Touch Sensor

A simple piece of conductive material acts as a touch sensor that detects when the user touches the robot. This input initiates recording or triggers responses, making interaction feel more intuitive.

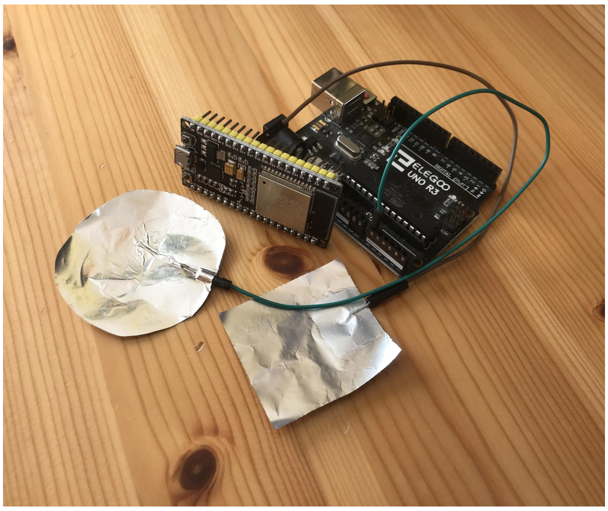

Figure 4.6: Touch Sensor Material

7. MAX9814 Microphone

This microphone module amplifies incoming audio and provides automatic gain control. It captures clear voice input for processing and streaming to the AI backend server.

Figure 4.7: MAX9814 Electret Microphone Amplifier

8. XY-SP5W Audio Amplifier

This 5W amplifier module boosts the DAC audio output to a level suitable for the 8-ohm speaker. It ensures clear and loud audio playback from the robot.

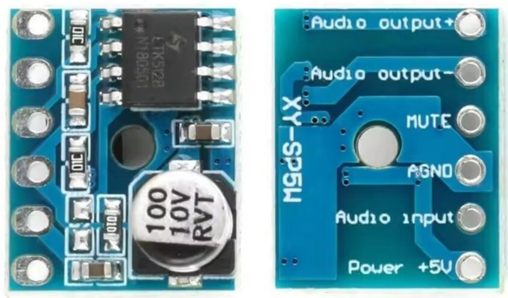

Figure 4.8: XY-SP5W Amplifier

9. 8ohm 5W Speaker

The speaker plays back voice responses or sound effects received from the AI engine. Connected through the DAC and amplifier, it provides clear and realistic audio output.

Figure 4.9: 8ohm 5W Speaker

10. LDR (Light Dependent Resistor)

The LDR sensor detects ambient light levels. The robot uses this data to adjust its eye brightness or respond to environmental changes, such as entering sleep mode in darkness.

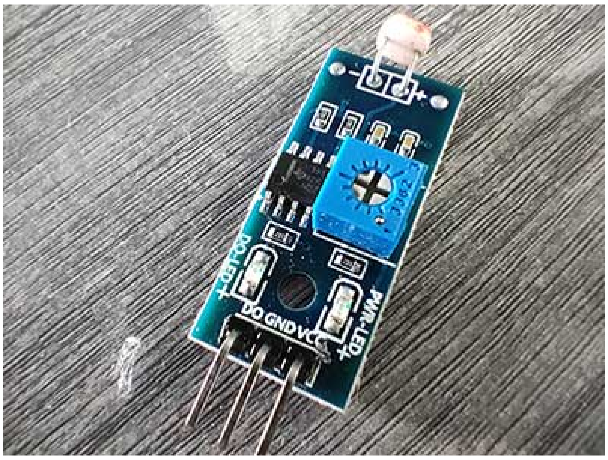

Figure 4.10: LDR Light Sensor

11. SD Card Module

This module stores audio recordings and TTS response files. It also contains the front-end interface of the mobile control app (HTML/CSS/JS), served directly from the robot.

Figure 4.11: SD Card Module

12. High-Power White LED

This LED is used as the robot's light source and can be turned on by user voice command. It adds both functional lighting and visual interactivity.

Figure 4.12: White High-Brightness LED

13. Voltage Regulator

The voltage regulator converts the total 12V input from the batteries into a stable 5V output required by the ESP32 and other modules, ensuring safe and reliable operation.

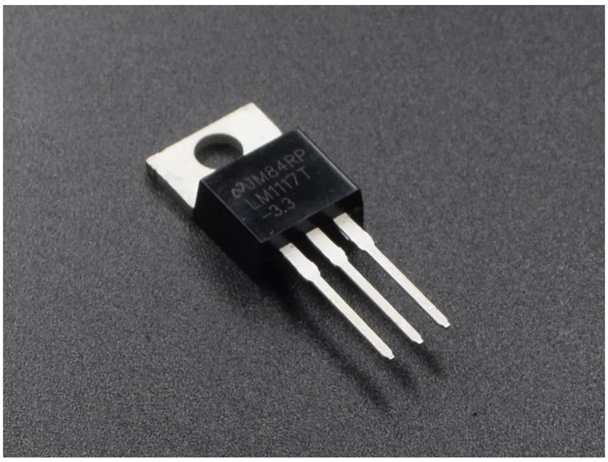

Figure 4.13: 5V Voltage Regulator

14. Batteries (3x 3.7V Li-ion)

Three 3.7V lithium-ion batteries provide a combined 11.1V power source. These are rechargeable and allow the robot to function wirelessly for extended periods.

Figure 4.14: 3.7V Li-ion Batteries

15. Power Input Socket

The power socket allows for external charging of the robot's batteries. It connects to a regulated charger and ensures safe charging without disassembly.

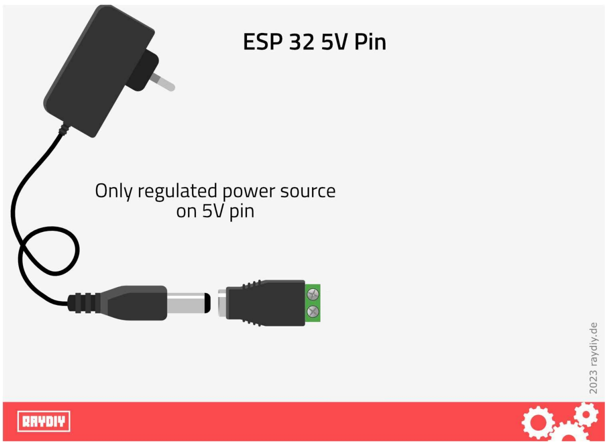

Figure 4.15: Charging Power Socket

Methodology

5.1 Wi-Fi Connectivity and App Communication

To establish a connection between the user and the Self-Balancing Pet Desk Robot, the ESP32 microcontroller first attempts to connect to a previously stored Wi-Fi network. If it fails to connect, it automatically switches to Access Point (AP) mode and displays the temporary network name and password on the OLED screen. The user connects their device to this network, after which the ESP32 serves a web interface that lists all available Wi-Fi networks. The user selects a desired network and enters the password. Upon successful connection, the ESP32 saves the credentials and generates a local IP address, which is displayed as a QR code on the OLED screen. When the user scans the QR code, the ESP32 serves the necessary HTML and JavaScript files to launch the web app. If the connection is unsuccessful, the user is prompted to try again. Once the app is opened, the user can choose between three main control modes: Remote Mode, Desk Mode, and Path Mode. In Remote Mode, the user manually controls the robot's movement. Desk Mode offers two sub-modes: Calm Mode, in which the robot remains stationary, and Normal Mode, where the robot performs small, gentle movements on the desk. This mode also includes edge detection using two upward-facing IR sensors (one at the front and one at the back) to prevent the robot from falling. In Path Mode, the user can send predefined commands such as move forward, move backward, rotate left, and rotate right, along with the number of repetitions (loops). The robot then executes the command sequence accordingly. This structured approach ensures both ease of connectivity and versatile interaction with the robot.

5.2 Balancing and Motion Control

A PID (Proportional–Integral–Derivative) controller is a feedback control algorithm widely used in real-time systems to maintain a desired output (setpoint) by minimizing the error between the target and actual values, and it achieves this by continuously calculating an error value—defined as the difference between a measured variable such as angle or speed and the desired setpoint—and then applying a correction based on three key components: the Proportional term, which reacts to the current error by applying a strong correction in proportion to how far off the system is; the Integral term, which reacts to the accumulation of past errors and helps eliminate persistent small deviations or biases; and the Derivative term, which responds to the rate of change of the error and anticipates future trends to smooth out corrections and prevent overshooting. The final output from the PID algorithm is a control signal sent to the motors, calculated using the formula: Output = (Kp × Error) + (Ki × Integral of Error) + (Kd × Derivative of Error),where Kp, Ki, and Kd are constants that must be finely tuned for the system to work effectively.

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau)\, d\tau + K_d \frac{de(t)}{dt}$$

In the Gemini AI Self-Balancing Pet Desk Robot, this algorithm is implemented in real-time on an ESP32 microcontroller, forming the core of its balance and motion system. The robot uses a single MPU-6050 sensor, which combines a 3-axis accelerometer and a 3-axis gyroscope to detect both the tilt angle relative to gravity and the rate of angular change. Through sensor fusion techniques such as complementary filtering, the ESP32 computes an accurate estimate of the robot's tilt angle. It then compares this angle to the upright setpoint, calculates the resulting error,and processes it through the PID controller to determine how much correction is required. This correction is translated into precise step and direction signals sent to the A4988 stepper motor drivers, which in turn control the two high-torque stepper motors attached to the robot's wheels. For instance, if the robot begins to tip forward, the wheels move forward to counteract the fall and restore balance, and the same applies in reverse for backward tipping. This entire sensing, calculating, and correcting process forms a fast, continuous feedback loop that runs hundreds of times per second, enabling the robot to maintain its upright position while stationary or in motion, and allowing it to dynamically respond to external disturbances or changes in direction with smooth and stable control.

5.3 voice assistant

The self-balancing, desk mate Office Robot delivers a rich, human-like interactive voice assistant experience through a harmonious blend of hardware and AI. When the user touches the robot, it immediately begins recording via its built-in MAX9814 microphone, which features automatic gain control to ensure clear audio. Meanwhile, the robot's OLED display animates expressive eyes, indicating that it is listening and interacting.

The captured audio is temporarily stored in WAV format on a microSD card, providing reliable local storage and freeing up internal memory for immediate processing. This raw audio data is then transmitted via Wi-Fi to a back-end server, where it is transcribed and passed to Google's Gemini AI model to generate a context-aware response. The generated response is converted into an audio file. The resulting audio—compressed as an 8-bit mono PCM WAV file (8kHz or 16kHz)—is streamed back to the robot and saved to an SD card in chunks. The ESP32 reads these audio chunks sequentially and sends them to the internal 8-bit DAC output (GPIO25 or GPIO26). The analog signal from the DAC is passed through a single XY-SP5W audio amplifier, which drives a 5W, 8-ohm speaker to deliver clear, natural sound.

Audio playback operates in a continuous, non-disruptive loop, ensuring smooth, continuous output even while new data is buffered or read from storage. This architecture enables seamless integration between audio input, processing, and output—enabling the robot to engage in seamless, interactive conversations with the user.

5.4 commands

As part of the robot's voice interaction system, users can issue specific commands through the onboard microphone, enabling hands-free control. When the user gives a voice command such as "dance," the robot responds by performing a predefined dance routine while simultaneously playing a song through its speaker, providing an engaging and entertaining experience. Additionally, users can control the built-in flashlight using voice commands like "open the light on" or "open the light off," upon which the robot turns the light on or off accordingly. These commands are processed through the voice assistant pipeline, which includes speech recognition, communication with the backend server, and command execution. This voice-based command feature enhances interactivity and adds both functional and expressive capabilities to the robot.

5.5 movement

The robot is capable of moving while maintaining its balance by leveraging adjustments within the PID algorithm. To initiate movement, the setpoint (target angle) in the PID controller is intentionally offset from the upright position, causing the robot to lean slightly in the desired direction. This controlled tilt prompts the robot to move forward or backward while still stabilizing itself. For rotation, the robot activates one stepper motor while keeping the other motor stationary, enabling it to turn left or right. By combining these methods, the robot achieves smooth, controlled motion while continuously balancing on two wheels.

5.6 Emotional Feedback Through Eye Animations

The robot uses an OLED display to simulate expressive eye animations that reflect its emotional state in response to user interactions. When the robot is hit or tapped suddenly, the eyes display a dissatisfied or upset expression, providing visual feedback that discourages harsh handling. Conversely, when the user gently touches or pats the robot for an extended duration, the eyes shift to a happy and friendly expression. This emotional feedback mechanism enhances the robot's personality, making interactions feel more lifelike and engaging. It also encourages users to treat the robot with care, reinforcing its role as a responsive and emotionally aware companion.

5.7 Ambient Light Detection Using LDR Sensor

The robot is equipped with an LDR (Light Dependent Resistor) sensor to monitor ambient light levels in its environment. When the surrounding light decreases significantly—such as when the room lights are turned off—the LDR sensor detects the change and triggers the robot to display a "sleeping" eye animation on its OLED screen. This feature gives the impression that the robot is going to sleep in response to darkness, enhancing its lifelike behavior and emotional presence. The use of ambient light detection adds another layer of interactivity, allowing the robot to respond naturally to environmental changes without user input.

5.8 Power Supply and Charging System

The robot is powered by a 12V rechargeable battery, providing sufficient voltage to drive the motors and other onboard components. To simplify the charging process, the robot is designed to charge automatically when connected to a charger. This is achieved by plugging the ESP32's USB interface into the charging system, which distributes power to the battery without requiring manual intervention. This setup ensures continuous operation with minimal user effort, making the robot more user-friendly and efficient for daily use.

Result

The Self-Balancing Pet Desk Robot effectively combines a number of intricate features that come together to create a companion that is intelligent, interactive, and sensitive to emotions. The robot uses two stepper motors and A4988 drivers to efficiently move forward, backward, and rotate while maintaining its balance with the help of a PID-controlled feedback loop that receives data from the MPU6050 sensor. ESP32 is used to connect the robot to the user's device in both Station and Access Point modes, providing dependable communication and flexible Wi-Fi configuration.

Users can switch between Remote Mode, Desk Mode, and Path Mode with ease thanks to the web-based control interface. When in Desk Mode, the robot is successfully kept from falling off the desk by the edge detection system that uses infrared sensors. The robot can respond to commands like "dance" or "turn on the light" thanks to voice command functionality, which offers both entertainment and practicality.

Additionally functional are emotion-based responses: the robot's OLED eyes display discontent when struck or tapped, and prolonged, gentle touches produce joyful expressions. The robot's lifelike personality is enhanced by displaying a sleeping animation in low-light conditions in response to changes in ambient lighting detected by the LDR sensor.

Conclusion

In conclusion, the Desk mate Robot successfully combines advanced voice assistant functionality with real-time motion control and interactive environmental awareness to create a rich, emotionally engaging user experience. The robot captures user voice input through the onboard MAX9814 microphone and displays active listening using expressive OLED-based eye animations, signaling its responsiveness and emotions. Voice data is saved to an SD card and streamed to a backend server where Google's Gemini AI processes it and returns a contextual, natural-language response. This response is converted to speech and smoothly played back through an 8-ohm, 5-watt speaker via the ESP32's 8-bit DAC and an XY-SP5W amplifier, enabling high-quality audio output. The robot maintains its balance using a precise PID control algorithm with input from an MPU-6050 IMU sensor and stepper motors driven by A4988 modules, allowing dynamic movement and upright stability. Additionally, the system includes an LDR sensor for ambient light detection, enabling the robot to adjust its visual feedback and behavior based on environmental brightness. Through this sensor and OLED screen, the robot can visually express different emotional states such as curiosity, sleepiness, or excitement, making the interaction more lifelike. Voice commands allow users to trigger specific actions, ask questions, or change the robot's behavior, enhancing interactivity. Altogether, this robot is a sophisticated integration of AI, embedded electronics, control systems, and emotional expression, representing a cutting-edge advancement in personal robotics.