Data Collection Protocol for Human Activity Recognition

1. Objective

This protocol provides detailed, step-by-step instructions for collecting high-quality sensor data (accelerometer and gyroscope) for human activity recognition tasks. The collected data will be used to build machine learning models such as LSTM, GRU, and CNN for activity classification.

2. Equipment Required

- Smartphone with accelerometer and gyroscope sensors (Android/iOS device)
- Data collection application (e.g., Physics Toolbox Sensor Suite, Sensor Kinetics, custom app)
- Secure strap or band to attach the phone to the body (waist or thigh preferred)
- Notebook or digital log for manual annotations

3. Setup

- Attach the smartphone securely to the participant's waist, chest, or thigh.
- Ensure the phone orientation remains consistent across participants.
- Start the sensor recording before beginning activities.
- Collect accelerometer (x, y, z) and gyroscope (x, y, z) data at consistent sampling rates (e.g., 20–50Hz).

4. Activities to Record

Participants should perform the following activities:

Activity	Duration (Recommended)
Walking	3–5 minutes
Jogging	3–5 minutes
Sitting	3–5 minutes
Standing	3–5 minutes
Upstairs Walking	2–3 minutes
Downstairs Walking	2–3 minutes

- Each activity should be clearly annotated in the data (or manually logged).
- Repeat each activity 2–3 times for better coverage.

5. Data Logging

If your app supports labeling:

- Label data during recording (preferred).

If manual:

- Log the start and end times of each activity manually.
- Use a stopwatch if needed for accuracy.

Sample log example:

Activity	Start Time	End Time
Walking	10:02:00	10:07:00
Sitting	10:08:00	10:13:00

6. Data Format

Data should be saved in CSV format with the following columns:

user_id	activity	timestamp	acc_x	acc_y	acc_z	gyro_x	gyro_y	gyro_z

- Timestamp should be in milliseconds if possible.
- Activity should match the activity list.

7. Quality Assurance

- Check for missing sensor data (NaN values).
- Confirm that each activity duration matches the log.
- Remove or fix data segments where the device was loose or removed

8. Ethical Considerations

- Inform participants about the study purpose.
- Obtain written consent if needed.
- Anonymize user IDs to protect privacy.

9. Common Issues to Avoid

- Unstable attachment causing noisy data.
- Inconsistent phone orientation between users.
- Missing activity labels.
- Gaps or jumps in timestamp data.

Conclusion

Following this protocol ensures the collection of clean, structured, and reliable sensor data for developing human activity recognition models. Consistency and careful logging are the keys to building high-quality datasets for AI projects.a