

## A Kinematic Feature Extraction

To characterize the spatial and temporal properties of joint movements, we extract a set of well-established kinematic features from the preprocessed gait data. These features are formally defined below, along with their corresponding mathematical formulations.

*Mean Height* The average vertical position of a joint during the gait cycle, providing insight into elevation patterns [14].

$$\mu = \frac{1}{N} \sum_{i=1}^N h_i$$

*Range of Motion (ROM)* The difference between maximum and minimum joint positions, reflecting joint flexibility [15].

$$\text{ROM} = h_{\max} - h_{\min}$$

*Standard Deviation* Measures the variability of joint movement, indicating dispersion around the mean [14].

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (h_i - \mu)^2}$$

*Skewness* Quantifies the asymmetry of the joint movement distribution [16].

$$\text{Skewness} = \frac{1}{N} \sum_{i=1}^N \left( \frac{h_i - \mu}{\sigma} \right)^3$$

*Kurtosis* Measures the peakedness of the movement distribution, indicating the presence of outliers [16].

$$\text{Kurtosis} = \frac{1}{N} \sum_{i=1}^N \left( \frac{h_i - \mu}{\sigma} \right)^4$$

*Time to Peak Height* Represents the normalized time at which the joint reaches its maximum vertical position [15].

$$T_{\text{peak}} = \frac{t_{\text{peak}}}{T_{\text{total}}}$$

*Velocity RMS* The root mean square (RMS) of joint velocity, indicating average movement speed [17].

$$\text{Velocity}_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{i=1}^N v_i^2}$$

*Acceleration RMS* RMS of joint acceleration, representing the average rate of change of velocity [17].

$$\text{Acceleration}_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{i=1}^N a_i^2}$$

*Jerk RMS* RMS of joint jerk, commonly associated with movement smoothness [18].

$$\text{Jerk}_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{i=1}^N j_i^2}$$

## B Kinetic Feature Extraction

To capture the dynamics of ground reaction forces and loading patterns, we extract the following kinetic features from the force plate data. These features quantify both the magnitude and variability of forces involved in gait.

*Mean Force* The average ground reaction force (GRF) during the stance phase, indicating overall loading characteristics [15].

$$\mu_F = \frac{1}{N} \sum_{i=1}^N F_i$$

*Range of Force* The difference between the maximum and minimum recorded ground reaction forces, reflecting variability in loading [14].

$$\text{Range}_F = F_{\max} - F_{\min}$$

*Standard Deviation* Measures the variability of ground reaction forces across the stance phase [14].

$$\sigma_F = \sqrt{\frac{1}{N} \sum_{i=1}^N (F_i - \mu_F)^2}$$

*Skewness* Quantifies the asymmetry of the GRF distribution, which may indicate irregular loading patterns [16].

$$\text{Skewness}_F = \frac{1}{N} \sum_{i=1}^N \left( \frac{F_i - \mu_F}{\sigma_F} \right)^3$$

*Kurtosis* Measures the peakedness of the GRF distribution, highlighting extreme force events [16].

$$\text{Kurtosis}_F = \frac{1}{N} \sum_{i=1}^N \left( \frac{F_i - \mu_F}{\sigma_F} \right)^4$$

*Time to Peak Force* Represents the normalized time at which the peak ground reaction force occurs during the stance phase [15].

$$T_{\text{peak}}^F = \frac{t_{\text{peak}}^F}{T_{\text{stance}}}$$

*Velocity RMS of Force Signal* The RMS of the first derivative of the force signal, reflecting the rate of change in applied forces [17].

$$\text{Velocity}_{\text{RMS}}^F = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{dF_i}{dt} \right)^2}$$

*Acceleration RMS of Force Signal* RMS of the second derivative of the force signal, capturing rapid variations in applied forces [17].

$$\text{Acceleration}_{\text{RMS}}^F = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{d^2 F_i}{dt^2} \right)^2}$$

*Jerk RMS of Force Signal* RMS of the third derivative of the force signal, used to assess abrupt transitions and loading smoothness [18].

$$\text{Jerk}_{\text{RMS}}^F = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{d^3 F_i}{dt^3} \right)^2}$$

## C LLM Explainability Configuration

### C.1 Prompt Template

The following prompt was used with the GPT-4o model via the OpenAI API to generate human-readable explanations of learning progression. This configuration was applied throughout the experiments unless otherwise stated.

You are an expert in AI-assisted cognitive and motor function monitoring.

For the following subject, you are given the results of three machine learning models designed to detect signs of learning over 20 movement trials. Additionally, a combined Differential Learning Index (DLI) has been computed as a reference metric based on:

- 40% Autoencoder reconstruction error reduction,
- 30% Dynamic Time Warping (DTW) final distance evolution,
- 30% Forecasting accuracy (small gap between real and predicted slopes).

**Differential Learning Index (DLI):** {dli\_score} *Interpretation: Higher DLI indicates stronger learning potential based on objective metrics.*

**Autoencoder:** - Trial-wise reconstruction errors: {rec\_errors} - Learning index (error delta from first to last trial): {learning\_index}

**Forecasting Transformer:** - Real slope of gait evolution: {real\_slope} - Predicted slope from model: {pred\_slope} - Gap between real and predicted slopes: {slope\_gap}

**Dynamic Time Warping (DTW):** - Final DTW distance from trial 1 to trial 20: {final\_dtw} - DTW curve over trials: {dtw\_curve}

**Instructions:**

Based on the above data and the DLI score, answer the following: 1. On a scale from **0.0 to 1.0**, rate the subject's learning potential across these trials. *0.0 = No observable learning behavior, 1.0 = Strong, clear evidence of learning and adaptation*

2. Provide a brief reasoning behind your score, specifying which models influenced your judgment the most.

Be precise, but concise. Output the score on a separate line in the following format: **Learning Potential Score: [value between 0.0 and 1.0]**

## C.2 Example Output

**Subject S07 (DLI: 0.3052) LLM Verdict:**

Based on the provided data, the subject's learning potential can be evaluated as follows:

**1. Autoencoder:** The reconstruction error shows a slight decrease from the first to the last trial, with a learning index of -0.01004. This suggests a minimal improvement in the subject's ability to perform the task as expected by the model. However, the reduction is not substantial, indicating limited learning.

**2. Forecasting Transformer:** The gap between the real and predicted slopes is 0.01558, which is relatively large. This indicates that the model's ability to predict the subject's gait evolution is not closely aligned with the actual performance, suggesting limited learning or adaptation.

**3. Dynamic Time Warping (DTW):** The final DTW distance of 40.85 shows some improvement from the initial distance of 48.75, but the curve over trials is inconsistent, with fluctuations indicating variability in the subject's performance. While there is some reduction in distance, the inconsistency suggests that learning is not robust.

The Differential Learning Index (DLI) of 0.3052 reflects these observations, indicating moderate learning potential. The DTW model, with its final distance and curve, provides the most insight into the subject's performance variability, while the autoencoder and forecasting transformer suggest limited learning.

**Learning Potential Score: 0.3**