Determining Athlete Next-Day Readiness (via continuous and binary)

Project Aim

This is a readiness determination project, for which given a set of physiological, sleep, training, and recovery features, the model determines an athlete's readiness to compete the following day. 'Readiness' is quantified into brackets such as 'optimal performance ready', 'normal, baseline ready', and 'severely fatigued and unready'.

Essentially, we monitor athletes' physiology, sleep, training, and recovery over a set period of time and then determine their readiness for an upcoming (next-day) activity.

Note that days' intensities will differ across weeks; a typical week will be of a steady intensity, whereas some weeks may have extra intensity (e.g. extra match) and some lighter intensity (e.g. recovery).

Train/Val/Eval Setup

We will take a leave-one-athlete-out (LOAO) cross-validation approach, meaning that we will run 100 separate evaluations for all 100 athletes, for which every run will be using a model trained and validated on the remaining 99 athletes. We will take averages to determine performance.

- Train: ~80% of remaining 99 athletes
- Validate: ~20% of remaining 99 athletes
- Test: the single held-out athlete

Model Architecture

The model architecture will be a Transformer Encoder, in which every data point (day) within a window will simultaneously take the preceding and subsequent days' information (from within that window) to determine temporal patterns to forecast readiness at t+1.

During self-attention, weights will be assigned to both previous and subsequent days, for which high weights of past days hint at an emerging pattern, and high weights of future days hint at future outcomes validating said pattern.

With a window slide of 1, each athlete will have 103 windows ((149-47) + 1). This means 103×100 = 10300 windows. A batch size of 64 means ~ 161 (10300÷64) tensors of shape [64, 14, 17] per epoch.

Note that every window (14 days) will have one corresponding readiness label (CRI, binary, & ordinal), for which patterns between days can be identified to determine prediction.

<u>Linear Projection and Positional Encoding</u>

Linear projection increases the column dimension of the input tensor via feature embedding, learning relationships between features on the same day. [B, 14, 17] → [B, 14, d_model]. The output is a latent physiological state representation for each day.

Positional encoding gives the model temporal awareness; it can now recognise temporal relationships between days (e.g. day 5 is 3 days before day 8).

Transform Encoder

Every layer in the Transformer Encoder contains two crucial elements: self-attention and feed-forward network. The self-attention mechanism uses backpropagation to assign weights to every day's (datapoint) contextual days (surrounding datapoints). Contextual days that result in lower loss (CRI, binary, and ordinal) will be assigned higher weights. Context-aware vectors for each day are sent to the feed-forward network, where temporal patterns are identified.

There will be 3 layers within the Transformer Encoder. Every self-attention head in every layer will compose a context-aware vector of weights that is more intricate and detailed than their corresponding head in the previous layer, while still addressing the same learning. For example, with day 8 as the focus, SA-head 1 in layer 1 identifies that days 5-7 minimise loss (short-term effect). Then, SA-head 1 in layer 2 understands that it's the HRV and RHR trends within days 5-7 that are driving that effect. Finally, SA-head 1 in layer 3 identifies nuance; this effect is only applicable to those individuals with a particular baseline.

Self-attention and Feed-forward networks

Self-attention has 4 heads, meaning that there are 4 Query, Key, Value projections for which different assortments of contextual weights are applied. This means that for a single day, there will be multiple context-aware vectors, each giving rise to specific inter-day patterns determining readiness (e.g. short-term patterns, medium-range recovery cycles, long-term workload trends). Self-attention learns relationships between days over the 14-day window.

There is one feed-forward network per layer, where the FFN processes all the context-aware vectors into non-linear feature interactions. Early-layer FFNs learn basic feature combinations (e.g. high load + low sleep → strain), whereas later-layer FFNs learn more abstract, context-dependent rules (e.g. strain pattern only matters if chronic load is already high).

Residual connection and Layer normalization

Residual connection is the addition of the original input to the output of the attention and FFN to prevent loss of original information and the gradual change of weights.

Layer normalization rescales layers to ensure stable training.

Attention Pooling

The output vector from the Encoder is still of shape [B, 14, d_model], where d_model is the embedding size. To predict readiness, we must have one value (per feature interaction) representing all 14 days. Hence [B, 14, d_model] → [B, d_model]. This can be done via attention pooling, where the most important days in determining readiness are weighted more heavily.

After attention pooling, the pooled vector feeds 3 output heads: a CRI regression head ([B, 1]) trained with MSE, a binary readiness head ([B, 1] with sigmoid) trained with BCE, and a CORAL ordinal head ([B, K-1] logits) trained with ordinal loss.

Transformer Architecture (Diagram)

```
Inputs (per batch)
[B, 14, 17]   ← 14-day window, 17 features/day (z-scored; includes CRI_{t-1})

   │ Linear Projection (feature embedding)
   ▼
[B, 14, d_model]

   │ + Positional Encoding (sin/cos over day index)
   ▼
[B, 14, d_model]

   │ Encoder Layer × L
   │  ┌─────────────────────────────────────────────┐
   │  │ Multi-Head Self-Attention (MHA)             │
   │  │  - Training: bidirectional attention        │
   │  │  - Inference: causal mask (attend to ≤ t only)  │
   │  │ Residual + LayerNorm                        │
   │  │ Position-wise FeedForward (MLP)             │
   │  │ Residual + LayerNorm                        │
   │  └─────────────────────────────────────────────┘
   ▼
[B, 14, d_model]

   │ Sequence Pooling → one vector per window
   │  (choices: [CLS] token, attention pooling, or mean pooling)
   ▼
[B, d_model]

   │ Readiness Heads (multitask)
   ├── CRI head: Linear → Sigmoid           → CRI_{t+1} ∈ [0,1]
   ├── Binary head: Linear → Sigmoid        → Ready / Not (0/1)
   └── Ordinal head: Ordinal/Cornell/CE loss → Class 1..5

Loss = w_reg * MSE(CRI) + w_bin * BCE + w_ord * OrdinalLoss
```

Loss and Performance Metrics

Loss for:

- CRI: Mean Squared Error (average difference between predicted CRI and true CRI).
- ReadyBin: Binary Cross Entropy (average difference between predicted probability of 0/1 and the true bin).

- ReadyOrd: Ordinal (CORAL) Loss (average difference between predicted thresholds and true thresholds).

Performance Metrics for CRI:
- Root Mean Squared Error
- Mean Absolute Error.

Performance Metrics for ReadyBin:
- Area Under ROC Curve (probability a random positive, 1, is scored higher than a random negative, 0); the higher the value the better.
- F1 Score (combines model's precision and recall); the higher value the better.

Performance Metrics for ReadyOrd:
- Quadratic Weighted Kappa (measures how well predicted classes match true ones, -1 to 1, where 1 is perfect, 0 is indifferent to randomness, and -1 is complete disagreement). Predictions further from true value are punished more.
- MAE of class index.
- Macro-F1, where all classes are treated equally.

Feature Data

The following data will be used as features for determining athlete readiness. Features will be contained in windows, where each window holds data for 14 days. Therefore, each window will hold ~238 (14×17) values. Feature data will be normalised (n) to reflect the variation seen amongst subjects in the PPG-Dalia dataset.

Physiological:
- Resting Heart Rate (RHR) [bpm] (n)
- Heart Rate Variability (HRV) [bpm] (n)
- HRV delta (vs 7-day baseline)
- RHR delta (vs 7-day baseline)

HRV and RHR delta will measure the difference in HRV and RHR between the present day and the baseline (average) for the past 7 days.

Sleep:
- Sleep duration [hours] (n)
- Sleep efficiency [%] (n)

A low sleep efficiency indicates that an individual was frequently disrupted during their sleep, meaning that sleep cycles were likely unfulfilled and hence worse recovery. Irrespective of actual sleep duration.

Training Load:

- Session duration [min]
- Session RPE (Rating of Perceived Exertion) [0-10]
- Internal Load (= RPE x duration)
- Acute Load (7-day sum)
- Chronic Load (28-day average of weekly sums)
- ACWR (Acute Chronic Workload Ratio) (= acute / chronic)
- Day-of-week cycle (encoded as sine)
- Day-of-the-week cycle (encoded as cosine)

The ACWR allows for comparison of the last week's exertion to the average of the previous 4 weeks' exertion. A day-of-the-week cycle will allow for the model to recognise different durations between different days of the week, enabling the identification of patterns between days (e.g. how events on Monday influence metrics reported on Wednesday compared to Sunday).

Recovery:
- Fatigue (Likert 1-7) (n)
- Soreness (Likert 1-7) (n)
- Previous-day CRI (from t-1)

Readiness has temporal dependence; an individual who isn't ready today will likely not be ready tomorrow. This feature allows for the model to understand recovery dynamics, and hence make predictions based on trajectories.

Raw Feature Data

The intricacy of this project means that feature data will have to be synthetically created. The procedures for determining these feature values are illustrated below. Feature data will be realistic (reflect physiological patterns seen in reality).

The feature data for physiology and sleep will have PPG-Dalia and Sleep-EDF distributions respectively. This is to ensure inter-individual variation, which is present in humans.

As for training load and recovery, their distributions will be similar to that of PPG-Dalia and Sleep-EDF's. This will be executed through custom coefficients, per athlete, for all variables influencing RPE, fatigue, and soreness.

HRV and RHR dynamics (physiology)

HRV dynamics (z-scored)

$$\begin{aligned} HRV_{z,t} = \rho_h\, HRV_{z,t-1} \;-\; a_1\, Load_{z,t-1} \;-\; a_2\, ACWR_t \;-\; a_3\, RHR_{z,t-1} \\ +\; b_1\, SleepEff_{z,t-1} \;+\; b_2\, SleepDur_{z,t-1} \\ -\; a_4\, Load_{z,t-2} \;-\; a_5\, ACWR_{t-1} \\ +\; e_1\left(Load_{z,t-1} \times SleepEff_{z,t-1}\right) \;+\; \epsilon_t \end{aligned}$$

RHR dynamics (z-scored)

$$\begin{aligned} RHR_{z,t} = \rho_r\, RHR_{z,t-1} \;+\; c_1\, Load_{z,t-1} \;+\; c_2\, ACWR_t \;+\; c_3\left(-HRV_{z,t-1}\right) \\ -\; d_1\, SleepEff_{z,t-1} \;-\; d_2\, SleepDur_{z,t-1} \\ +\; c_4\, Load_{z,t-2} \;+\; c_5\, ACWR_{t-1} \\ +\; e_2\left(Load_{z,t-1} \times SleepEff_{z,t-1}\right) \;+\; \eta_t \end{aligned}$$

HRV and RHR values will be determined by the previous period's (yesterday) HRV, Load, ACWR, RHR, and sleep. These relationships are physiologically sound. Lag (hysteresis) was included through Loadt-2 and ACWRt-1, as well as the synergistic effect of sleep efficiency on Load's effect (Loadt-1 x SleepEff). Noise also included.

Sleep Duration and Sleep Efficiency (sleep)

Sleep duration

$$\begin{aligned} Dur_{z,t} = \rho_d\, Dur_{z,t-1} \;+\; \gamma_1\, Def_{t-1} \;+\; \gamma_2\, Load_{z,t-1} \;-\; \gamma_3\, ACWR_t \\ +\; \kappa_{d1} \sin DOW_t \;+\; \kappa_{d2} \cos DOW_t \\ +\; \gamma_4\left(Load_{z,t-1} \times Def_{t-1}\right) \;+\; \epsilon_t^{(d)} \end{aligned}$$

Sleep efficiency

$$\begin{aligned} Eff_{z,t} = \rho_e\, Eff_{z,t-1} \;-\; \beta_1\, Load_{z,t-1} \;-\; \beta_2\, ACWR_t \;-\; \beta_3\, Fatigue_{z,t-1} \\ +\; \beta_4\, Def_{t-1} \;+\; \kappa_{e1} \sin DOW_t \;+\; \kappa_{e2} \cos DOW_t \\ +\; \beta_5\left(Load_{z,t-1} \times Eff_{z,t-1}\right) \;+\; \epsilon_t^{(e)} \end{aligned}$$

Sleep Duration values will be determined by the previous period's (yesterday) Duration, Deficit, Load, ACWR, and day of the week. An individual who has worked hard and is holding sleep debt will likely sleep longer that night (Loadt-1 x Deft-1). Noise also included.

Sleep Efficiency values will be determined by the previous period's (yesterday) Efficiency, Load, ACWR, Fatigue, Deficit, and day of the week. An individual who slept strong last night will likely handle today's load better during tonight's sleep (Loadt-1 x Efft-1). Noise also included.

$$Def_{t-1} = \max\left(0,\; \frac{TargetDur_i - Dur_{raw,t-1}}{SD_{Dur,i}}\right)$$

TargetDur is simply the individual's baseline sleep duration. If Dur > TargetDur, then Deficit is simply 0.

sinDOWt and cosDOWt shape a 7-day oscillation for which it oscillates around the average sleep duration and efficiency. Their coefficients determine what day the wave aligns on (e.g. wave peaks on Saturday as individual sleeps better on weekends).

Rating of Perceived Exertion (RPE) (training load)

$$\begin{aligned}RPE^{(i)}_{z,t} = \rho^{(i)}_p\, RPE^{(i)}_{z,t-1} \;+\; \phi^{(i)}_1\, Load^{(i)}_{z,t-1} \;+\; \phi^{(i)}_2\, ACWR^{(i)}_{t-1} \;+\; \phi^{(i)}_3\, Fatigue^{(i)}_{z,t-1} \\ -\; \phi^{(i)}_4\, Readiness^{(i)}_{t-1} \;-\; \phi^{(i)}_5\, SleepEff^{(i)}_{z,t-1} \;-\; \phi^{(i)}_6\, SleepDur^{(i)}_{z,t-1} \\ +\; \lambda^{(i)}_1 \sin DOW_t \;+\; \lambda^{(i)}_2 \cos DOW_t \\ +\; \phi^{(i)}_7 (Fatigue^{(i)}_{z,t-1} \times (1 - Readiness^{(i)}_{t-1})) \;+\; \mu^{(i)}_{RPE} \;+\; \epsilon^{(i,p)}_t\end{aligned}$$

RPE values will be determined by the previous period's (yesterday) RPE, Load, ACWR, Fatigue, Readiness (CRI), Sleep Efficiency, Sleep Duration, and day of the week. An individual who has high fatigue and low readiness (CRI) will likely experience a higher RPE than usual. Noise also included.

Notice the superscript (i) and additional variable μ to reflect individuality in variables and individuality in the subject variable respectively. This contributes to distribution in training load.

🏋 Microcycle pattern (example)

Day	Typical focus	Relative load
Mon	Training load begins to build	Medium
Tue	Hard session	High
Wed	Hard or moderate	High
Thu	Taper begins	Moderate
Fri	Light / activation	Low
Sat	Game / test day	High (short but intense)
Sun	Recovery	Very low

While the RPE equation determines variability in RPE within an individual between the same days in different weeks, the above table demonstrates the variability in RPE between days of the week. A tapered microcycle has been selected, whereby exercise intensity (RPE; load) rises from Mon-Wed, then decreases Thu-Fri, and then spikes on Saturday (game day), before being very low on Sunday (recovery).

Fatigue and Soreness (recovery)

Fatigue measures systemic tiredness: the body's overall physiological and neurological depletion. Soreness measures local muscular discomfort or stiffness from tissue microdamage and inflammation.

$$
\begin{aligned}
Fatigue^{(i)}_{z,t} = \rho^{(i)}_{f}\, Fatigue^{(i)}_{z,t-1} \;+\; \psi^{(i)}_{1}\, Load^{(i)}_{z,t-1} \;+\; \psi^{(i)}_{2}\, ACWR^{(i)}_{t-1} \;-\; \psi^{(i)}_{3}\, SleepDur^{(i)}_{z,t-1} \\
-\; \psi^{(i)}_{4}\, SleepEff^{(i)}_{z,t-1} \;-\; \psi^{(i)}_{5}\, Readiness^{(i)}_{t-1} \;+\; \psi^{(i)}_{6}\, RHR^{(i)}_{z,t-1} \;-\; \psi^{(i)}_{7}\, HRV^{(i)}_{z,t-1} \\
+\; \psi^{(i)}_{8}\, \sin DOW_t \;+\; \psi^{(i)}_{9}\, \cos DOW_t \\
+\; \psi^{(i)}_{10}\big(Load^{(i)}_{z,t-1} \times (1 - Readiness^{(i)}_{t-1})\big) \;+\; \mu^{(i)}_{Fatigue} \;+\; \epsilon^{(i,f)}_{t}
\end{aligned}
$$

Fatigue values will be determined by the previous period's (yesterday) Fatigue, Load, ACWR, Sleep Duration, Sleep Efficiency, Readiness (CRI), RHR, HRV, and the day of the week. An individual who has high load and low readiness (CRI) will likely experience a higher fatigue than usual. Noise also included.

Notice the superscript (i) and additional variable μ to reflect individuality in variables and individuality in the subject variable respectively. This contributes to distribution in fatigue.

$$
\begin{aligned}
Soreness^{(i)}_{z,t} = \rho^{(i)}_{s}\, Soreness^{(i)}_{z,t-1} \;+\; \omega^{(i)}_{1}\, Load^{(i)}_{z,t-1} \;+\; \omega^{(i)}_{2}\, ACWR^{(i)}_{t-1} \;+\; \omega^{(i)}_{3}\, Fatigue^{(i)}_{z,t-1} \\
-\; \omega^{(i)}_{4}\, SleepDur^{(i)}_{z,t-1} \;-\; \omega^{(i)}_{5}\, SleepEff^{(i)}_{z,t-1} \;-\; \omega^{(i)}_{6}\, Readiness^{(i)}_{t-1} \\
+\; \omega^{(i)}_{7}\, \sin DOW_t \;+\; \omega^{(i)}_{8}\, \cos DOW_t \\
+\; \mu^{(i)}_{Soreness} \;+\; \epsilon^{(i,s)}_{t}
\end{aligned}
$$

Soreness values will be determined by the previous period's (yesterday) Soreness, Load, ACWR, Fatigue, Sleep Duration, Sleep Efficiency, Readiness (CRI), and the day of the week. An individual who has high load and low readiness (CRI) will likely experience a higher fatigue than usual. Noise also included.

Notice the superscript (i) and additional variable μ to reflect individuality in variables and individuality in the subject variable respectively. This contributes to distribution in soreness.

Scale	z-range	Fatigue meaning	Soreness meaning
1	z ≤ −1.5	Fully fresh, energetic	No soreness
2	−1.5 < z ≤ −1.0	Very low fatigue	Minimal soreness
3	−1.0 < z ≤ −0.3	Slight fatigue	Mild muscle tightness
4	−0.3 < z ≤ +0.3	Moderate / baseline	Typical post-training soreness
5	+0.3 < z ≤ +1.0	Noticeable fatigue	Moderate soreness, training impact
6	+1.0 < z ≤ +1.5	High fatigue	Pronounced soreness, reduced comfort
7	z > +1.5	Extreme exhaustion	Severe soreness, needs recovery

As explained under *Feature Data*, the normalised fatigue and soreness values will take a Likert (1-7) scale. 1 being extremely positive and 7 being extremely negative.

Raw Feature Data Continued

Below showcases the influence of different predictors on different subject predictors, alongside their average baseline for all 100 athletes. The absolute value of influence (e.g. -3.0 ms) is per 1 SD increase in a predictor (relative to that athlete's baseline).

HRV and RHR dynamics (physiology)

Average HRV Baseline:
- mean 80 ms
- within-athlete SD 15 ms

Persistence: **ρ_h = 0.65** (autocorrelation of HRV_z)
Absolute effects (HRV change in **ms**):

- **Load_{t−1}: a_1 = 0.20 ⇒ −3.0 ms**
- **ACWR_{t}: a_2 = 0.15 ⇒ −2.25 ms**
- **RHR_{t−1}: a_3 = 0.10 ⇒ −1.5 ms** (per +1 SD RHR)
- **SleepEff_{t−1}: b_1 = 0.18 ⇒ +2.7 ms**
- **SleepDur_{t−1}: b_2 = 0.08 ⇒ +1.2 ms**
- **Load_{t−2}: a_4 = 0.08 ⇒ −1.2 ms**
- **ACWR_{t−1}: a_5 = 0.06 ⇒ −0.9 ms**
- **Interaction (Load×SleepEff): e_1 = 0.05 ⇒ +0.75 ms** (per 1 SD × 1 SD)

Average RHR Baseline:
- mean 48 bpm
- within-athlete SD 5 bpm

Persistence: **ρ_r = 0.70** (autocorrelation of RHR_z)
Absolute effects (RHR change in **bpm**):

- **Load_{t−1}: c_1 = 0.18 ⇒ +0.90 bpm**
- **ACWR_{t}: c_2 = 0.12 ⇒ +0.60 bpm**
- **(−HRV_{t−1}): c_3 = 0.10 ⇒ −0.50 bpm** (i.e., higher HRV lowers RHR)
- **SleepEff_{t−1}: d_1 = 0.15 ⇒ −0.75 bpm**
- **SleepDur_{t−1}: d_2 = 0.06 ⇒ −0.30 bpm**
- **Load_{t−2}: c_4 = 0.07 ⇒ +0.35 bpm**
- **ACWR_{t−1}: c_5 = 0.05 ⇒ +0.25 bpm**
- **Interaction (Load×SleepEff): e_2 = 0.05 ⇒ +0.25 bpm** (per 1 SD × 1 SD; acts as a moderator alongside the negative sleep main effects)

Sleep Duration and Sleep Efficiency (sleep)

Average Sleep Duration Baseline:

- mean 8.0 h
- within-athlete SD 0.8 h

Persistence ρ_d = 0.55
Absolute effect per +1 SD predictor:

- **Sleep deficit (t−1): γ_1 = 0.30 → +0.24 h** (rebound effect)
- **Load (t−1): γ_2 = 0.12 → +0.10 h** (tired → sleep longer)
- **ACWR (t): γ_3 = 0.08 → −0.06 h** (sustained overload → shorter sleep)
- **Interaction (Load×Def): γ_4 = 0.06 → +0.05 h** (big rebound after heavy load)
- **Weekly rhythm (sin/cos DOW): κ ≈ 0.05 → ±0.04 h**

Average Sleep Efficiency Baseline:

- mean 90%
- within-athlete SD 5%

Persistence ρ_e = 0.60
Absolute effect per +1 SD predictor:

- **Load (t−1): β_1 = 0.10 → −0.5 %**
- **ACWR (t): β_2 = 0.10 → −0.5 %**
- **Fatigue (t−1): β_3 = 0.12 → −0.6 %**
- **Sleep deficit (t−1): β_4 = 0.06 → +0.3 %** (sleepier → fall asleep faster)
- **Interaction (Load×Eff t−1): β_5 = 0.05 → +0.25 %** (good sleepers buffer load)
- **Weekly rhythm (sin/cos DOW): κ ≈ 0.05 → ±0.25 %**

Rating of Perceived Exertion (RPE) (training load)

Average RPE Baseline:

- mean 6.0 (on a 0-10 Borg CR10 scale)
- within-athlete SD 1.0

Persistence: ρ□ = **0.60** → about **60 %** of yesterday's RPE carries into today.

Absolute effect per +1 SD predictor (≈ 1 RPE unit):

- **Load$_{(\square-1)}$:** φ_1 = 0.10 → **+0.10 RPE** (more training stress → next session feels harder)
- **ACWR$_{(\square-1)}$:** φ_2 = 0.10 → **+0.10 RPE** (overload phase → higher perceived effort)
- **Fatigue$_{(\square-1)}$:** φ_3 = 0.20 → **+0.20 RPE** (tired body → sessions feel tougher)
- **Readiness$_{(\square-1)}$:** φ_4 = 0.25 → **−0.25 RPE** (high readiness → training feels easier)
- **SleepEff$_{(\square-1)}$:** φ_5 = 0.10 → **−0.10 RPE** (better sleep → lower effort)

- **SleepDur$_{(\square-1)}$:** φ_6 = 0.08 → **−0.08 RPE** (longer sleep → lower effort)
- **Interaction (Fatigue × (1 − Readiness))$_{(\square-1)}$:** φ_7 = 0.10 → **+0.10 RPE** (low readiness + high fatigue → big RPE spike)
- **Weekly rhythm (sin/cos DOW):** λ ≈ 0.05 → **±0.05 RPE** (harder mid-week, lighter weekend)

Fatigue and Soreness (recovery)

Average Fatigue Baseline:
- mean 4.0 (on a 1-7 Likert scale)
- within-athlete SD 1.0

Persistence: ρ_f = **0.60** → about **60 %** of yesterday's fatigue carries into today.

Absolute effect per +1 SD predictor:

- **Load$_{(\square-1)}$:** ψ_1 = 0.25 → **+0.25 fatigue** (higher load → more systemic tiredness)
- **ACWR$_{(\square-1)}$:** ψ_2 = 0.15 → **+0.15 fatigue** (recent overload → increased fatigue)
- **SleepDur$_{(\square-1)}$:** ψ_3 = 0.10 → **−0.10 fatigue** (longer sleep → less tiredness)
- **SleepEff$_{(\square-1)}$:** ψ_4 = 0.10 → **−0.10 fatigue** (better sleep → improved recovery)
- **Readiness$_{(\square-1)}$:** ψ_5 = 0.25 → **−0.25 fatigue** (higher readiness → lower fatigue)
- **RHR$_{(\square-1)}$:** ψ_6 = 0.10 → **+0.10 fatigue** (elevated RHR → fatigue signal)
- **HRV$_{(\square-1)}$:** ψ_7 = 0.10 → **−0.10 fatigue** (higher HRV → better recovery)
- **Weekly rhythm (sin/cos DOW):** ψ_8, ψ_9 ≈ 0.05 → **±0.05 fatigue** (mid-week peaks, weekend dips)
- **Interaction (Load × (1 − Readiness)):** ψ_{10} = 0.10 → **+0.10 fatigue** (low readiness + high load → fatigue spike)

Average Soreness Baseline:
- mean 4.0 (on a 1-7 Likert scale)
- within-athlete SD 1.0

Persistence: ρ_s = **0.50** → about **50 %** of yesterday's soreness carries into today.

Absolute effect per +1 SD predictor:

- **Load$_{(\square-1)}$:** ω_1 = 0.30 → **+0.30 soreness** (training stress → more muscle soreness)
- **ACWR$_{(\square-1)}$:** ω_2 = 0.20 → **+0.20 soreness** (sustained overload → elevated soreness)
- **Fatigue$_{(\square-1)}$:** ω_3 = 0.20 → **+0.20 soreness** (overall tiredness increases perceived soreness)
- **SleepDur$_{(\square-1)}$:** ω_4 = 0.08 → **−0.08 soreness** (longer sleep aids tissue repair)
- **SleepEff$_{(\square-1)}$:** ω_5 = 0.08 → **−0.08 soreness** (better sleep quality → lower soreness)
- **Readiness$_{(\square-1)}$:** ω_6 = 0.20 → **−0.20 soreness** (higher readiness → less muscle discomfort)
- **Weekly rhythm (sin/cos DOW):** ω_7, ω_8 ≈ 0.05 → **±0.05 soreness** (e.g., higher mid-week, lower on rest day)

Label Set

Our main label will be synthetically created, which will be of continuous nature and can be classified in two ways (binary and multi-class). These 3 measures of readiness will have their own loss which will guide models' performance.

Continuous Readiness Index (CRI)

$$\mathrm{Index}_t = \sigma\{+w_1\,\mathrm{HRV}_z - w_2\,\mathrm{RHR}_z + w_3\,\mathrm{SleepEff}_z - w_4\,\mathrm{ACWR} + w_5\,(\mathrm{HRV}_z \times \mathrm{SleepEff}_z) - w_6\,(\mathrm{ACWR} \times \mathrm{RHR}_z) + \mathrm{lags} + \epsilon\}$$

The CRI encompasses all the physiological metrics that would indicate an athlete's next-day 'readiness', in which the metrics' influence (weights and +/-) is determined by sport science. 'Lags' accounts for hysteresis among some events and the error term accounts for other uncontrollable factors. Some factors, in unity, are more telling of one's readiness state (e.g. HRV x SleepEfficiency) - hence their own variable. All predictors will be normalized to allow CRI to interpret the values relative to athletes' baseline. CRI will range from 0-1.

The CRI will be used to determine 2 more classifications:

- Binary Readiness (Ready / Not Ready)
- Ordinal Readiness (5 bands; very low → low → moderate → high → very high)

Ready: CRI ≥ 0.5 | Not Ready: CRI < 0.5

Very low (0-0.2):

- State: significant fatigue or overload; body not recovered
- Physiology: HRV well below baseline (> 1 SD down); RHR well above baseline (> 1 SD up)
- Interpretation: marked autonomic imbalance - parasympathetic suppression, sympathetic dominance, high systemic stress
- Performance: high injury risk, poor power output, slow reaction, reduced motivation; rest or active recovery strongly advised

Low (0.2-0.4):

- State: mild-to-moderate fatigue; partial recovery but still under residual stress
- Physiology: HRV moderately below baseline (0.5-1 SD down); RHR slightly above baseline (0.5 SD up)
- Interpretation: body still adapting or carrying fatigue from prior workload; not fully restored yet
- Performance: can train with caution - focus on technique, lighter loads, or recovery work; avoid maximal efforts

Moderate (0.4-0.6):

- State: normal, recovered enough to train safely
- Physiology: HRV and RHR near baseline (±0.5 SD)

- Interpretation: body is maintaining equilibrium - no excessive stress or overrecovery
- Performance: can train and compete effectively, but not necessarily peak output

High (0.6-0.8):
- State: strong recovery and good physiological balance
- Physiology: HRV moderately above baseline (+0.5–1 SD); RHR slightly below (−0.5 SD)
- Interpretation: Adaptations are trending positively - body primed to absorb training or perform well
- Performance: enhanced responsiveness, good energy, low fatigue

Very high (0.8-1):
- State: optimal physiological state - body fully recovered, stress well-managed
- Physiology: HRV well above baseline (>+1 SD); RHR clearly below baseline (−1 SD or more).
- Interpretation: exceptional recovery - strong parasympathetic nervous system, low systemic fatigue
- Performance: peak capacity - best days for high-intensity training, competition, or testing

TLDR:
Very low: severely fatigued (0-0.2)
Low: below normal readiness (0.2-0.4)
Moderate: typical / baseline readiness (0.4-0.6)
High: good recovery, ready to train (0.6-0.8)
Very high: peak readiness / optimal performance (0.8-1)

Label Set Continued

The CRI will have a sigmoid function applied to it so values are between 0 and 1, allowing for a more interpretable score.

Noise, ϵ, differences amongst athletes are normally distributed for which mean is 0 and standard deviation is 0.2. These values indicate athletes' spread of noise, indicating their volatility of unobserved factors affecting readiness. Unobserved factors include: stress, anxiety, weather, imperfect sensor readings.

$$\mathrm{Index}_t = \sigma\{+w_1\,\mathrm{HRV}_z - w_2\,\mathrm{RHR}_z + w_3\,\mathrm{SleepEff}_z - w_4\,\mathrm{ACWR} + w_5\,(\mathrm{HRV}_z \times \mathrm{SleepEff}_z) - w_6\,(\mathrm{ACWR} \times \mathrm{RHR}_z) + \mathrm{lags} + \epsilon\}$$

Rule of thumb: near CRI ≈ 0.5: a +1 SD change in a predictor shifts CRI by roughly 0.25 × weight after the sigmoid. CRI ≈ 0.5 is where the function is most reactive to a predictor's change, hence the sigmoid's 'S' shape. A 0.2 change on the logit scale roughly shifts CRI by ~0.05 (post-sigmoid).

Main effects (on z_t; all predictors are within-athlete z-scores)

- **HRV_z (w_1): +0.80** → increases readiness
- **RHR_z ($-w_2$): −0.60** → decreases readiness
- **SleepEff_z (w_3): +0.50** → increases readiness
- **ACWR_z ($-w_4$): −0.70** → decreases readiness

Interaction (synergy) terms

- **HRV_z × SleepEff_z (w_5): +0.30** → good sleep amplifies HRV's benefit
- **ACWR_z × RHR_z ($-w_6$): −0.25** → high load with high RHR compounds risk

Lag terms (delayed effects; all z-scored, from t−1)

- **HRV_z(t−1): +0.30**
- **RHR_z(t−1): −0.20**
- **Load_z(t−1): −0.25**
- **ACWR_z(t−1): −0.20**

Evaluation Results

After running 100 instantiations of the ReadinessTransformer() model for all 100 LOAO folds, below are the averaged results for evaluation (4dp):

- CRI_RMSE: **0.1203**
- CRI_MAE: **0.0903**

Strong accuracy with CRI predictions; model has captured inter-individual variability.

- BIN_AUROC: **0.9761**
- BIN_F1: **0.9054**

Strong accuracy determining ready versus not ready days.

- ORD_QWK: **0.9173**
- ORD_MAE: **0.3344**
- ORD_MacroF1: **0.6229**

ORD_QWK > 0.90 is very strong given the increased penalties with higher misclassifications. ORD_MAE being within a ⅓ of a category is still productive (if thinking of practical applications). ORD_MacroF1 is reasonable given the sharp thresholds and domination of mid-range categories.