7 Implementation Details

MIMIC-III physiologic deterioration prediction For patient graph structure learning, the dimension size of hidden states in GRUs is 17, the number of attention heads is 4, and the initial value of the learnable threshold ξ is 0.15. For patient graph structure refining, the number of disjoint clusters K is 11, and the add and delete ratios γ_{add} and γ_{del} are 0.35 and 0.25, respectively. For contrastive learning, the sampling ratio r is 0.4, the walk length l_{walk} is 2, and the dimension sizes of d_E and d_S are 23 and 19, respectively. The temperature parameter τ is 0.5. The dropout method is applied to the Softmax output layer, and the dropout rate is 0.2. For supervised learning setting, the scaling parameters α_1, α_2, and λ_2 are 1.65, 0.92 and 0.75, respectively.

MIMIC-III LOS prediction (3 days) For patient graph structure learning, the dimension size of hidden states in GRUs is 17, the number of attention heads is 4, and the initial value of the learnable threshold ξ is 0.15. For patient graph structure refining, the number of disjoint clusters K is 2, and the add and delete ratios γ_{add} and γ_{del} are 0.4 and 0.4, respectively. For contrastive learning, the sampling ratio r is 0.4, the walk length l_{walk} is 2, and the dimension sizes of d_E and d_S are 32 and 15, respectively. The temperature parameter τ is 0.35. The dropout method is applied to the Softmax output layer, and the dropout rate is 0.15. For supervised learning setting, the scaling parameters α_1, α_2, and λ_2 are 0.65, 0.27 and 0.87, respectively.

MIMIC-III LOS prediction (7 days) For patient graph structure learning, the dimension size of hidden states in GRUs is 17, the number of attention heads is 4, and the initial value of the learnable threshold ξ is 0.15. For patient graph structure refining, the number of disjoint clusters K is 3, and the add and delete ratios γ_{add} and γ_{del} are 0.4 and 0.25, respectively. For contrastive learning, the sampling ratio r is 0.4, the walk length l_{walk} is 2, and the dimension sizes of d_E and d_S are 30 and 14, respectively. The temperature parameter τ is 0.45. The dropout method is applied to the Softmax output layer, and the dropout rate is 0.2. For supervised learning setting, the scaling parameters α_1, α_2, and λ_2 are 1.8, 0.71 and 0.93, respectively.

MIMIC-III Self-supervised Learning Setting For patient graph structure learning, the dimension size of hidden states in GRUs is 17, the number of attention heads is 4, and the initial value of the learnable threshold ξ is 0.1. For patient graph structure refining, the number of disjoint clusters K is 9, and the add and delete ratios γ_{add} and γ_{del} are 0.4 and 0.05, respectively. For contrastive learning, the sampling ratio r is 0.5, the walk length l_{walk} is 4, and the dimension sizes of d_E and d_S are 31 and 18, respectively. The temperature parameter τ is 0.05. The dropout method is applied to the Softmax output layer, and the dropout rate is 0.35. The scaling parameters α_1, α_2, β_1, β_2 and λ_1 are 1.15, 0.98, 0.8, 1.9 and 0.9 respectively.

eICU physiologic deterioration prediction For patient graph structure learning, the dimension size of hidden states in GRUs is 17, the number of attention heads is 4, and the initial value of the learnable threshold ξ is 0.1. For patient graph structure refining, the number of disjoint clusters K is 6, and the

add and delete ratios γ_{add} and γ_{del} are 0.45 and 0.45, respectively. For contrastive learning, the sampling ratio r is 0.4, the walk length l_{walk} is 2, and the dimension sizes of d_E and d_S are 29 and 18, respectively. The temperature parameter τ is 0.6. The dropout method is applied to the Softmax output layer, and the dropout rate is 0.05. For supervised learning setting, the scaling parameters α_1, α_2, and λ_2 are 0.65, 0.47 and 0.8, respectively.

eICU LOS prediction (3 days) For patient graph structure learning, the dimension size of hidden states in GRUs is 17, the number of attention heads is 4, and the initial value of the learnable threshold ξ is 0.1. For patient graph structure refining, the number of disjoint clusters K is 7, and the add and delete ratios γ_{add} and γ_{del} are 0.25 and 0.3, respectively. For contrastive learning, the sampling ratio r is 0.5, the walk length l_{walk} is 4, and the dimension sizes of d_E and d_S are 21 and 17, respectively. The temperature parameter τ is 0.5. The dropout method is applied to the Softmax output layer, and the dropout rate is 0.05. For supervised learning setting, the scaling parameters α_1, α_2, and λ_2 are 0.5, 0.25 and 0.79, respectively.

eICU LOS prediction (7 days) For patient graph structure learning, the dimension size of hidden states in GRUs is 17, the number of attention heads is 4, and the initial value of the learnable threshold ξ is 0.15. For patient graph structure refining, the number of disjoint clusters K is 4, and the add and delete ratios γ_{add} and γ_{del} are 0.2 and 0.15, respectively. For contrastive learning, the sampling ratio r is 0.3, the walk length l_{walk} is 3, and the dimension sizes of d_E and d_S are 32 and 13, respectively. The temperature parameter τ is 0.05. The dropout method is applied to the Softmax output layer, and the dropout rate is 0.05. For supervised learning setting, the scaling parameters α_1, α_2, and λ_2 are 0.7, 0.7 and 0.25, respectively.

eICU Self-supervised Learning Setting For patient graph structure learning, the dimension size of hidden states in GRUs is 17, the number of attention heads is 4, and the initial value of the learnable threshold ξ is 0.15. For patient graph structure refining, the number of disjoint clusters K is 3, and the add and delete ratios γ_{add} and γ_{del} are 0.15 and 0.35, respectively. For contrastive learning, the sampling ratio r is 0.45, the walk length l_{walk} is 3, and the dimension sizes of d_E and d_S are 31 and 22, respectively. The temperature parameter τ is 0.15. The dropout method is applied to the Softmax output layer, and the dropout rate is 0.05. The scaling parameters α_1, α_2, β_1, β_2 and λ_1 are 1.95, 0.14, 0.85, 0.95 and 0.35 respectively.