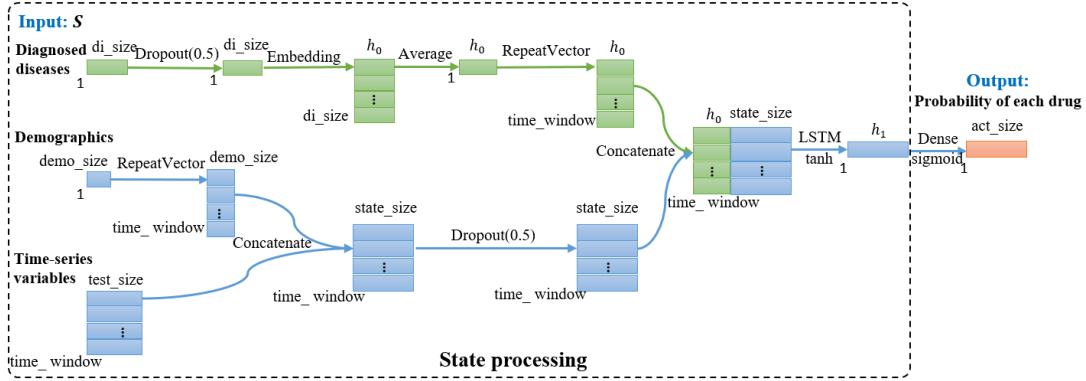
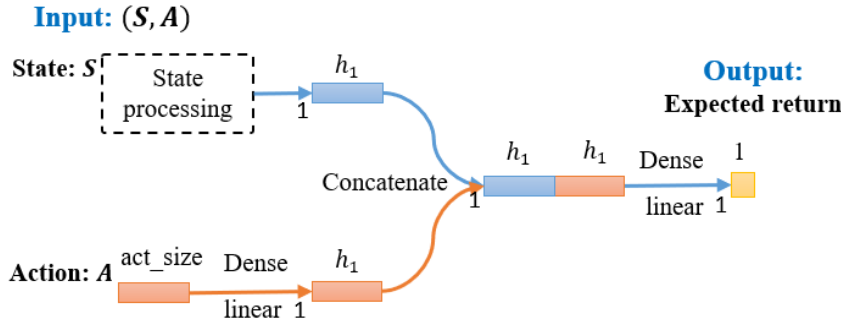


A. Actor network $\mu(S|\theta^\mu)$ and $\mu'(S|\theta^{\mu'})$



B. Critic network $Q(S, A|\theta^Q)$ and $Q'(S, A|\theta^{Q'})$



C. Training process of the SRL-LSTM model

