**Table S-6**. RNN Model architecture, hyperparameters, and training details

|  |  |
| --- | --- |
|  **Parameter** |  **Value**  |
|  Number of LSTM layers |  2 |
|  Hidden units in LSTM layers |  (375, 375) |
|  Initial learning rate |  1e-5 |
|  Learning rate decay |  0.5 |
|  Loss function |  patient-averaged binary cross entropy |
|  Batch size |  32 |
|  Optimizer  |  RMSprop |
|  Dropout of sequences |  0.25 |
|  Dropout of input elements  |  0.2 |
|  Recurrent dropout |  0.2 |
|  Weights initializer |  Glorot uniform (inputs), orthogonal (recurrent) |
|  Weights Regularizer, parameter |  L2, 1e-4 |
|  Output activation |  sigmoid |