1. Hyper-parameter selection

The hyper-parameters of our model are mainly divided into two parts: the ones in the neural network and the ones of the topology extraction module. The hyper-parameters in the neural network include the kernel size in the convolutional layer, the window size of the pooling layer, the number of output channels, the size of dilation rate and learning rate. The dimension of embedding matrix for the drug-target pair is  or , so the selection range of the convolution kernel and pooling window in the dilation convolution and multi-layer convolutional neural networks is {2×1, 2×2, 2×3, 2×4}. The number of output channels is selected from {8, 16, 32, 64}. The dilation rate and learning rate are selected from the {1, 2, 3, 4, 5} and {10-4, 10-3, 10-2, 10-1} respectively. In the topology extraction and embedding module, the hyper-parameters include low-dimensional feature dimension *k*, and optimization item parameters and . The variation range of the feature dimension *k* is {100, 150, 200, 250}. and are selected from {10-2, 10-1, 1,10,100}. According to our previous research experience about DTI prediction, we set the initial values of these parameters. Then we adjusted one of parameters in its variation range when fixing the remaining parameters. We selected the final value of the parameter when the model obtains the best performance by setting it to the specific value. In this way, we determined the values of all the hyper-parameters.