1 onehot\_cnn\_one\_branch

Single branch CNN model: First, perform one-hot encoding on Sequence in Interaction Region1 and Sequence in Interaction Region2, and then merge the two one-hot encoded data. Use KERNEL\_SIZE = 24, strides = 4 CNN layer, RELU activation layer and BatchNormalization layer as a CNN unit, repeat this CNN unit three times, then flatten the data and send it to the dense layer of size 2048, and then enter the dropout degree of 0.5 The dropout layer prevents the model from overfitting, and then enters the dense layer with a size of 2048, then enters the dropout layer with a dropout degree of 0.5, and then enters the dense layer with a size of 1000. Finally, output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax). This model imitates the well-known image processing model VGG16, which has more parameters and can obtain better results.

2 onehot\_cnn\_two\_branch

Two-branch CNN model: First, perform one-hot encoding on Sequence in Interaction Region1 and Sequence in Interaction Region2 respectively, and then send them to different CNN units. Use the CNN layer with KERNEL\_SIZE = 24 and strides = 4, the RELU activation layer and the BatchNormalization layer as a CNN unit, repeat this CNN unit three times, and then use the Flatten layer for the Sequence in Interaction Region 1 branch and Sequence in Interaction Region 2 branch of the CNN unit After merging, the data is sent to the dense layer with a size of 2048, and then it enters the dropout layer with a dropout degree of 0.5 to prevent the model from overfitting, and then enters the dense layer with a size of 2048, and then enters the dropout layer with a dropout degree of 0.5, and then Enter the dense layer with a size of 1000. Finally, output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).

3 onehot\_embedding\_dense

One-hot encoding and embedding model: First, perform one-hot encoding on Sequence in Interaction Region1 and Sequence in Interaction Region2 respectively, and then merge the two one-hot encoded data. Then use the 6mer embedding layer for embedding, then flatten the data and send it to the 2048 dense layer, then enter the dropout layer with a dropout degree of 0.5 to prevent the model from overfitting, then enter the 2048 dense layer, and then enter the dropout The dropout layer with a degree of 0.5, then the dense layer with a size of 1000. Finally, output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).

4 onehot\_embedding\_cnn\_one\_branch

Use one-hot encoding and embedding and then use CNN to process the model: First, perform one-hot encoding on Sequence in Interaction Region1 and Sequence in Interaction Region2 respectively. Then use the 6mer embedding layer for embedding, then merge the two embedding data, and then send it to the CNN unit, use the CNN layer with KERNEL\_SIZE = 24, strides = 4, the RELU activation layer and the BatchNormalization layer as a CNN unit, repeat this CNN unit three times, then flatten the data and send it to the dense layer with a size of 2048, then enter the dropout layer with a dropout degree of 0.5 to prevent the model from overfitting, then enter the dense layer with a size of 2048, and then enter the dropout with a dropout degree of 0.5 Layer, then enter the dense layer with a size of 1000. Finally, output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).

5 onehot\_embedding\_cnn\_two\_branch

Use one-hot encoding and embedding and then use CNN to process the model: First, perform one-hot encoding on Sequence in Interaction Region1 and Sequence in Interaction Region2 respectively. Then use the 6mer embedding layer for embedding, and then send it to the CNN unit, use the CNN layer with KERNEL\_SIZE = 24, strides = 4, the RELU activation layer and the BatchNormalization layer as a CNN unit, repeat the CNN unit three times, and then merge the Sequence in CNN results of Interaction Region1 and CNN results of Sequence in Interaction Region2. Then flatten the data and send it to the dense layer with a size of 2048, then enter the dropout layer with a dropout degree of 0.5 to prevent the model from overfitting, then enter the dense layer with a size of 2048, then enter the dropout layer with a dropout degree of 0.5, and then enter A dense layer with a size of 1000. Finally, output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).

6 onehot\_dense

Dense model: First, perform one-hot encoding on Sequence in Interaction Region1 and Sequence in Interaction Region2, and then merge the two one-hot encoded data. Then flatten the data and send it to the dense layer. The dense layer has five layers, which contain 2048, 2048, 2048, 2048, and 1000 parameters. Insert a dropout layer with a dropout of 0.5 between every two dense layers to prevent the model from overfitting. Finally, output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).

7 onehot\_resnet18

Resnet model: First, perform one-hot encoding on Sequence in Interaction Region1 and Sequence in Interaction Region2, and then merge the two one-hot encoded data. Use the resnet18 model with KERNEL\_SIZE = 16, strides = 4, and finally output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).

8 onehot\_resnet34

Resnet model: First, perform one-hot encoding on Sequence in Interaction Region1 and Sequence in Interaction Region2, and then merge the two one-hot encoded data. Use the resnet34 model with KERNEL\_SIZE = 16, strides = 4, and finally output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).

9 embedding\_cnn\_one\_branch

Single-branch CNN model: First embed the Sequence in Interaction Region1 and Sequence in Interaction Region2 respectively, and then merge the two-embedding data. The embedding matrix uses the hg19 matrix trained by DNA2vec, and the size is 6mer (4097, where the first item is all Is 0) x100 dimension. Then the data is sent to the CNN unit. Use KERNEL\_SIZE = 24, strides = 4 CNN layer, RELU activation layer and BatchNormalization layer as a CNN unit, repeat this CNN unit three times, then flatten the data and send it to the dense layer of size 2048, and then enter the dropout degree of 0.5 The dropout layer prevents the model from overfitting, and then enters the dense layer with a size of 2048, then enters the dropout layer with a dropout degree of 0.5, and then enters the dense layer with a size of 1000. Finally, output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).

10 embedding\_cnn\_one\_branch

Two-branch CNN model: First, perform embedding on Sequence in Interaction Region1 and Sequence in Interaction Region2 respectively, and then merge the two-embedding data. The embedding matrix uses the hg19 matrix trained by DNA2vec, and the size is 6mer (4097, where the first item is all Is 0) x100 dimension. Then the data is sent to the CNN unit. Use the CNN layer with KERNEL\_SIZE = 24 and strides = 4, the RELU activation layer and the BatchNormalization layer as a CNN unit, repeat this CNN unit three times, and then send the Sequence in Interaction Region 1 branch and Sequence in Interaction Region 2 branch of the CNN unit to flatten The layers are merged, and then the data is sent to the dense layer with a size of 2048, and then into the dropout layer with a dropout degree of 0.5 to prevent the model from overfitting, and then into the dense layer with a size of 2048, and then into the dropout layer with a dropout degree of 0.5 , And then enter the dense layer of size 1000. Finally, output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).

11 embedding\_dense

Dense model: First, perform embedding on Sequence in Interaction Region1 and Sequence in Interaction Region2 respectively, then merge the two embedding data, where the embedding matrix uses the hg19 matrix trained by DNA2vec, the size is 6mer (4097, where the first item is all 0) x100 dimensions. Then flatten the data and send it to the dense layer. The dense layer has five layers, which contain 2048, 2048, 2048, 2048, and 1000 parameters. Insert a dropout layer with a dropout of 0.5 between the two dense layers to prevent the model from overfitting. Finally, Output the result in the form of one-hot encoding (2 values, normalized to 01 using softmax).