***Table R1. The composition of the RNABert module***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **RNABert component** | **Submodule** | | | **Number of layers** |
| **BertModel** | **BertEmbeddings** | | **word\_embeddings** | **4** |
| **position\_embeddings** |
| **token\_type\_embeddings** |
| **LayerNorm** |
| **BertEncoder** | **BertLayer**  **(transformer )** | **BertAttention** | **5** |
| **BertIntermediate** | **1** |
| **BertOutput** | **2** |
| **BertLayer** | | **8** |
| **BertLayer** | | **8** |
| **BertLayer** | | **8** |
| **BertLayer** | | **8** |
| **BertLayer** | **BertAttention** | **5** |
| **BertIntermediate** | **1** |
| **\*BertOutput** | **2** |
| **\*BertPooler** | | | **1** |
| **\*BertPreTrainingHeads** | **\*MaskedWordPredictions** | | **\*nn.Linear** | **3** |
| **\*nn.Linear** |
| **\*LayerNorm** |
| **\*MaskedWordPredictions** | | | **3** |
| **\*nn.Linear** | | | **1** |