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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **RNABert component** | **Submodule** | | | **Number of layers** | **Frozen** |
| **BertModel** | **BertEmbeddings** | | **word\_embeddings** | **4** | **Yes** |
| **position\_embeddings** | **Yes** |
| **token\_type\_embeddings** | **Yes** |
| **LayerNorm** | **Yes** |
| **BertEncoder** | **BertLayer**  **(transformer )** | **BertAttention** | **5** | **Yes** |
| **BertIntermediate** | **1** | **Yes** |
| **BertOutput** | **2** | **Yes** |
| **BertLayer** | | **8** | **Yes** |
| **BertLayer** | | **8** | **Yes** |
| **BertLayer** | | **8** | **Yes** |
| **BertLayer** | | **8** | **Yes** |
| **BertLayer** | **BertAttention** | **5** | **Yes** |
| **BertIntermediate** | **1** | **Yes** |
| **\*BertOutput** | **2** | **No** |
| **\*BertPooler** | | | **1** | **No** |
| **\*BertPreTrainingHeads** | **\*MaskedWordPredictions** | | **\*nn.Linear** | **3** | **No** |
| **\*nn.Linear** | **No** |
| **\*LayerNorm** | **No** |
| **\*MaskedWordPredictions** | | | **3** | **No** |
| **\*nn.Linear** | | | **1** | **No** |