**The Chinese character in the manuscript：**

1. "江" (river)
2. "河" (river)
3. "湖" (lake)
4. "海" (sea)
5. "氵" (water)
6. "老师" (teacher)
7. "老鼠" (mouse)
8. "老虎" (tiger)
9. "老"（The word "老" often means "old," but in the three words listed before, it is just a phonetic symbol and does not convey the semantic meaning of "old."）
10. "骼" (skeleton)
11. "骨" (bone)
12. "左右结构 "(structure: left-right composition)
13. "膝关节康复训练"(Knee joint rehabilitation training)
14. "身体" (body)
15. "康复方法" (rehabilitation methods)

**The Chinese character in the figures：**

1. "髋" in fig.1 means "hip"
2. "骨" in fig.1 means "bone"
3. "左右" in fig.1 means "left-right composition"
4. "形声" in fig.1 means "pictophonetic compounds"
5. "名词" in fig.1 means "noun"
6. "骼" in fig.1 means "skeleton"
7. "肩外展屈髋以及跌倒" : The phrase "肩外展屈髋以及跌倒" is a part of the Chinese medical text in Fig.4 and Fig.5, and its English meaning is "Shoulder abduction with external rotation, hip flexion, and falling."

**The Chinese character in the code files in:**

<https://github.com/AaabbB-quick/A-Named-Entity-Recognition-model-embedding-rehabilitation-medicine-Chinese-character-knowledge-graph.>

（DOI: 10.6084/m9.figshare.29466521）

"初始化" means "Initialization"

"正常训练" means "Normal training"

"反向传播，得到正常的" means "Backpropagation to obtain regular gradients"

"对抗训练" means "Adversarial training"

"在embedding上添加对抗扰动" means "Add adversarial perturbations to embeddings"

"反向传播，并在正常的grad基础上，累加对抗训练的梯度" means "Backpropagate and accumulate adversarial gradients on top of regular gradients"

"恢复embedding参数" means "Restore embedding parameters"

"梯度下降，更新参数" means "Gradient descent to update parameters"

"这个参数要换成你模型中embedding的参数名" means "Replace this parameter with your model's embedding parameter name"

"自定义学习率变化机制" means "Custom learning rate scheduling mechanism"

"Bert模型内定的学习率变化机制" means "Built-in learning rate scheduling mechanism for BERT models"

"线性预热方式" means "Linear warmup method"

"主要参考论文<< Attention Is All You Need>>中的学习更新方式" means "Primarily references the learning update method from the paper <<Attention Is All You Need>>"

"模型保存，两种模式：" means "Model saving, two modes:"

"1. 直接保存最好模型" means "1. Directly save the best model"

"2. 按照epoch频率保存模型" means "2. Save models at epoch frequency"

"计算模式" means "Computation mode"

"这里主要重新加载模型时候" means "Primarily when reloading models here"

"对best重新赋值" means "Reassign the 'best' value"

"正常模型" means "Regular model"

"需要保存的信息" means "Information to be saved"

"当前判断指标" means "Current evaluation metric"

"是否保存最好模型" means "Whether to save the best model"

"每隔几个epoch保存下模型" means "Save model every few epochs"

"适合bert类型模型，适合pytorch\_transformer模块" means "Suitable for BERT-type models and pytorch\_transformer module"

"重新开始训练的epoch点" means "Epoch point to resume training"

"np.float32会报错" means "np.float32 will cause errors"

"写入文件" means "Write to file"

"保存train图像" means "Save training plots"

"指标变化" means "Metric changes"

"曲线" means "Curves"

"需要成对出现" means "Need to appear in pairs"

"康复方法" means "Rehabilitation methods"

"功能障碍及表现" means "Dysfunctions and manifestations"

"身体" means "Body"

"康复评定" means "Rehabilitation assessment"

"药物" means "Pharmacotherapy"

"康复设备" means "Rehabilitation equipment"

"功能障碍及表现康复评定" means "Dysfunctions-manifestations rehabilitation assessment"

"直接模型保存，如果分布式有另外表达方式" means "Direct model saving (use distributed-specific method if applicable)"

"保存定义的args参数" means "Save defined args parameters"

"保存tokenizer，如果tokenizer改变了就加载保存的，而不是直接从s3加载" means "Save tokenizer; if modified load saved version instead of directly from S3"

"优化器和schdule保存" means "Save optimizer and scheduler"

"训练完毕" means "Training completed"

"给它转成和原始的json格式相同，将源文本输入进去更好" means "Convert to match original JSON format, preferably with source text input"

"定义 NER task" means "Define NER task"

"加载模型，注意这里的Token和bert是基于继承原始模型的。token要加载save\_pretrained才行，否则报错。" means "Load model (note: Token/Bert inherit from original; token requires save\_pretrained load to avoid errors)"

"如果是接着训练的话也可以加载保存的" means "Load saved checkpoint if resuming training"

"模型加载完毕" means "Model loading completed"

"global\_step作用其实是在保存模型的时候起一个tag作用" means "global\_step serves as a tag during model saving"

"预测的结果转成json格式" means "Convert prediction results to JSON format"

**The Chinese character in the code files in:**

<https://github.com/CBLUEbenchmark/CBLUE>

"无关系" means "Unrelated"  
"主语" means "Subject"  
"从sequence中寻找子串pattern" means "Find substring pattern in sequence"  
"如果找到，返回第一个下标；否则返回0" means "Return first index if found; otherwise 0"  
"sub在text的起点" means "sub is at text start position"  
"不是起点" means "Not a starting point"  
"起点和终点重合" means "Start and end positions coincide"  
"终点前遇到新的起点，停止搜索" means "Encounter new start before end → stop search"  
"没找到终点->孤立点" means "End not found → isolated point"  
"起点" means "Start position"  
"终点" means "End position"  
"获取token的“词干”" means "Extract token stem"  
"如果是##开头，则自动去掉" means "Remove '##' prefix automatically"  
"控制类字符判断" means "Control character detection"  
"判断是不是有特殊含义的符号" means "Identify semantically significant symbols"  
"给出原始的text和tokenize后的tokens的映射关系" means "Map raw text to tokenized tokens"

"国际疾病分类 ICD-10北京临床版" means "International Classification of Diseases ICD-10 Beijing Clinical Edition"

"词典" means "Dictionary"

"语料库" means "Corpus"

"建立模型" means "Model training"

"相似度分数" means "Similarity score"

"疾病表述" means "Disease description"

"指标解读" means "Indicator interpretation"

"医疗费用" means "Medical costs"

"治疗方案" means "Treatment plan"

"功效作用" means "Efficacy effects"

"病情诊断" means "Condition diagnosis"

"其他" means "Other"

"注意事项" means "Precautions"

"病因分析" means "Etiology analysis"

"就医建议" means "Medical consultation advice"

"后果表述" means "Outcome description"

"阴性" means "Negative"

"阳性" means "Positive"

"不标注" means "Not labeled"

"主诉" means "Chief complaint"

"现病史" means "Present illness history"

"既往史" means "Past medical history"

"诊断" means "Diagnosis"

"建议" means "Recommendation"

"预防阶段" means "Prevention stage"

"就诊科室" means "Clinical department"

"同义词" means "Synonym"

"辅助治疗" means "Adjuvant therapy"

"化疗" means "Chemotherapy"

"放射治疗" means "Radiotherapy"

"手术治疗" means "Surgical treatment"

"实验室检查" means "Laboratory test"

"影像学检查" means "Imaging examination"

"辅助检查" means "Auxiliary examination"

"组织学检查" means "Histological examination"

"内窥镜检查" means "Endoscopy"

"筛查" means "Screening"​

"多发群体" means "High-risk population"

"发病率" means "Incidence rate"

"发病年龄" means "Age of onset"

"多发地区" means "Endemic regions"

"发病性别倾向" means "Gender predisposition"

"死亡率" means "Mortality rate"

"多发季节" means "Seasonal prevalence"

"传播途径" means "Transmission route"

"并发症" means "Complication"

"病理分型" means "Pathological subtype"

"相关（导致）" means "Related (causing)"

"鉴别诊断" means "Differential diagnosis"

"相关（转化）" means "Related (transformation)"

"相关（症状）" means "Related (symptom)"

"临床表现" means "Clinical manifestation"

"治疗后症状" means "Post-treatment symptoms"

"侵及周围组织转移的症状" means "Symptoms of local invasion/metastasis"

"病因" means "Pathogenesis"

"高危因素" means "Risk factors"

"风险评估因素" means "Risk assessment factors"

"病史" means "Medical history"

"遗传因素" means "Genetic factors"

"发病机制" means "Pathogenic mechanism"

"病理生理" means "Pathophysiology"

"药物治疗" means "Pharmacotherapy"

"发病部位" means "Site of onset"

"转移部位" means "Metastatic site"

"外侵部位" means "Site of invasion"

"预后状况" means "Prognosis status"

"预后生存率" means "Prognostic survival rate"

"治疗药物" means "Therapeutic drugs"

"用法用量" means "Dosage and administration"

"禁用药物" means "Contraindicated drugs"​

"将符合诊疗决策树的节点前序序列转化为代表诊疗决策树结构的节点矩阵" means "Convert qualifying diagnostic decision tree node sequences into node adjacency matrices"