* 最新的技术：Collaborative Knowledge Base、Neural Factorization Machine
* Meta-path
* KG-enhanced recommendation
* Collaborative filtering(CF)：协同过滤
* Embedding layer：做矩阵乘法，实现降维或升维。
* LSTM layer：
* 负采样：每次让一个训练样本仅仅更新一小部分的权重，这样就会降低梯度下降过程中的计算量。
* MF
* NFM
* CKE
* FMG
* 非自回归解码器：可以并行处理，不依赖顺序
* 二分图匹配：
* 实体查询：一个可学习向量的集合；实体查询的一个集合；Qspan表示的N个实体查询；实体查询间的自注意力，实体间的依赖；交叉注意力，上下文信息。
* 自注意力和交叉注意力：self-attention的Q，K和V都是同一个输入, 即当前序列由上一层输出的高维表达。cross-attention的Q代表当前序列；而K和V是同一个输入，对应的是encoder最后一层的输出结果。Q，K和V是经过卷积后得到的特征，其形状为（batch\_size，seq\_length，num\_features）。将查询（Q）和键（K）相乘会得到（batch\_size，seq\_length，seq\_length）特征，这大致告诉我们序列中每个元素的重要性，确定我们“注意”哪些元素。 注意数组使用softmax标准化，因此所有权重之和为1。 最后，注意力将通过矩阵乘法应用于值（V）数组。[从头开始实现Transformer - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/299409721)
* 多头注意力机制：参照Transformer模型中的解释。类别不平衡问题：指分类任务中不同类别的训练样本数目差别很大的情况。一般来说，不平衡样本会导致训练模型侧重样本数目较多的类别，而“轻视”样本数目较少类别，这样模型在测试数据上的泛化能力就会受到影响。一个例子，训练集中有99个正例样本，1个负例样本。在不考虑样本不平衡的很多情况下，学习算法会使分类器放弃负例预测，因为把所有样本都分为正便可获得高达99%的训练分类准确率。
* Focal Loss：主要是为了解决one-stage目标检测中正负样本比例严重失衡的问题，修改了交叉熵loss，使模型关注难以区分的样本来减小loss，lambda调节简单样本权重降低的速率，当lambda为0时即为交叉熵损失函数，当lambda增加时，调整因子的影响也在增加。实验发现lambda为2是最优。同时增加了一个平衡因子alpha，用来平衡正负样本本身的比例不均问题。[类别不平衡问题 - 深夜十二点三十三 - 博客园 (cnblogs.com)](https://www.cnblogs.com/CJT-blog/p/10223157.html" \l ":~:text=%E7%B1%BB%E5%88%AB%E4%B8%8D%E5%B9%B3%E8%A1%A1%E9%97%AE%E9%A2%98%E6%8C%87%E5%88%86%E7%B1%BB%E4%BB%BB%E5%8A%A1%E4%B8%AD%E4%B8%8D%E5%90%8C%E7%B1%BB%E5%88%AB%E7%9A%84%E8%AE%AD%E7%BB%83%E6%A0%B7%E6%9C%AC%E6%95%B0%E7%9B%AE%E5%B7%AE%E5%88%AB%E5%BE%88%E5%A4%A7%E7%9A%84%E6%83%85%E5%86%B5%E3%80%82,%E4%B8%80%E8%88%AC%E6%9D%A5%E8%AF%B4%EF%BC%8C%E4%B8%8D%E5%B9%B3%E8%A1%A1%E6%A0%B7%E6%9C%AC%E4%BC%9A%E5%AF%BC%E8%87%B4%E8%AE%AD%E7%BB%83%E6%A8%A1%E5%9E%8B%E4%BE%A7%E9%87%8D%E6%A0%B7%E6%9C%AC%E6%95%B0%E7%9B%AE%E8%BE%83%E5%A4%9A%E7%9A%84%E7%B1%BB%E5%88%AB%EF%BC%8C%E8%80%8C%E2%80%9C%E8%BD%BB%E8%A7%86%E2%80%9D%E6%A0%B7%E6%9C%AC%E6%95%B0%E7%9B%AE%E8%BE%83%E5%B0%91%E7%B1%BB%E5%88%AB%EF%BC%8C%E8%BF%99%E6%A0%B7%E6%A8%A1%E5%9E%8B%E5%9C%A8%E6%B5%8B%E8%AF%95%E6%95%B0%E6%8D%AE%E4%B8%8A%E7%9A%84%E6%B3%9B%E5%8C%96%E8%83%BD%E5%8A%9B%E5%B0%B1%E4%BC%9A%E5%8F%97%E5%88%B0%E5%BD%B1%E5%93%8D%E3%80%82 %E4%B8%80%E4%B8%AA%E4%BE%8B%E5%AD%90%EF%BC%8C%E8%AE%AD%E7%BB%83%E9%9B%86%E4%B8%AD%E6%9C%8999%E4%B8%AA%E6%AD%A3%E4%BE%8B%E6%A0%B7%E6%9C%AC%EF%BC%8C1%E4%B8%AA%E8%B4%9F%E4%BE%8B%E6%A0%B7%E6%9C%AC%E3%80%82)
* N-gram：N-Gram是一种基于统计语言模型的算法。将文本里面的内容按照字节进行大小为N的滑动窗口操作，形成了长度是N的字节片段序列。每一个字节片段称为gram，对所有gram的出现频度进行统计，并且按照事先设定好的阈值进行过滤，形成关键gram列表，也就是这个文本的向量特征空间，列表中的每一种gram就是一个特征向量维度。
* layer disorientation：
* 假阳性预测：True positive (TP)：真实为P，预测为P；True negative (TN): 真实为N，预测为N； False positive (FP)：真实为N，预测为P；False negative (FN)：真实为P，预测为N
* Dropout layer：节点过多防止过拟合，随机丢弃一些节点
* inductive bias：
* Layer normalization：[NLP中 batch normalization与 layer normalization - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/74516930)
* CNN with zero paddings：
* SmoothL1 loss：
* IoU：目标检测中的常用算法，交并比，预测框和真值框的重叠区域和并集区域的比值，选择匹配较好（重叠度较高）的框
* NMS 非极大化抑制：NMS主要就是通过迭代的形式，不断的以最大得分的框去与其他框做IoU操作，并过滤那些IoU较大（即交集较大）的框。[NMS、 soft-nms、softer-nms - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/89426063)
* the Soft-NMS algorithm：NMS算法略显粗暴，因为NMS直接将删除所有IoU大于阈值的框。soft-NMS吸取了NMS的教训，在算法执行过程中不是简单的对IoU大于阈值的检测框删除，而是降低得分。算法流程同NMS相同，但是对原置信度得分使用函数运算，目标是降低置信度得分.
* bi-affine classifier：

# 一 NER

NLP模型合集：[一文读懂BERT(原理篇)\_程序猿废柴的博客-CSDN博客\_bert](https://blog.csdn.net/jiaowoshouzi/article/details/89073944)

### Transformer: [一文理解 Transformer 的工作原理-InfoQ](https://www.infoq.cn/article/QBloqM0Rf*SV6v0JMUlF)、

### [transfomer里面self-attention的Q, K, V的含义 - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/158952064)

### FLAT: [FLAT：中文NER屠榜之作！ - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/225238667)

### BiLSTM-CRF: [序列标注：Bi-LSTM + CRF - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/42096344)

### GCN: [如何理解 Graph Convolutional Network（GCN）？ - 知乎 (zhihu.com)](https://www.zhihu.com/question/54504471)

**类别不平衡问题和Focal loss**：[类别不平衡问题 - 深夜十二点三十三 - 博客园 (cnblogs.com)](https://www.cnblogs.com/CJT-blog/p/10223157.html" \l ":~:text=%E7%B1%BB%E5%88%AB%E4%B8%8D%E5%B9%B3%E8%A1%A1%E9%97%AE%E9%A2%98%E6%8C%87%E5%88%86%E7%B1%BB%E4%BB%BB%E5%8A%A1%E4%B8%AD%E4%B8%8D%E5%90%8C%E7%B1%BB%E5%88%AB%E7%9A%84%E8%AE%AD%E7%BB%83%E6%A0%B7%E6%9C%AC%E6%95%B0%E7%9B%AE%E5%B7%AE%E5%88%AB%E5%BE%88%E5%A4%A7%E7%9A%84%E6%83%85%E5%86%B5%E3%80%82,%E4%B8%80%E8%88%AC%E6%9D%A5%E8%AF%B4%EF%BC%8C%E4%B8%8D%E5%B9%B3%E8%A1%A1%E6%A0%B7%E6%9C%AC%E4%BC%9A%E5%AF%BC%E8%87%B4%E8%AE%AD%E7%BB%83%E6%A8%A1%E5%9E%8B%E4%BE%A7%E9%87%8D%E6%A0%B7%E6%9C%AC%E6%95%B0%E7%9B%AE%E8%BE%83%E5%A4%9A%E7%9A%84%E7%B1%BB%E5%88%AB%EF%BC%8C%E8%80%8C%E2%80%9C%E8%BD%BB%E8%A7%86%E2%80%9D%E6%A0%B7%E6%9C%AC%E6%95%B0%E7%9B%AE%E8%BE%83%E5%B0%91%E7%B1%BB%E5%88%AB%EF%BC%8C%E8%BF%99%E6%A0%B7%E6%A8%A1%E5%9E%8B%E5%9C%A8%E6%B5%8B%E8%AF%95%E6%95%B0%E6%8D%AE%E4%B8%8A%E7%9A%84%E6%B3%9B%E5%8C%96%E8%83%BD%E5%8A%9B%E5%B0%B1%E4%BC%9A%E5%8F%97%E5%88%B0%E5%BD%B1%E5%93%8D%E3%80%82 %E4%B8%80%E4%B8%AA%E4%BE%8B%E5%AD%90%EF%BC%8C%E8%AE%AD%E7%BB%83%E9%9B%86%E4%B8%AD%E6%9C%8999%E4%B8%AA%E6%AD%A3%E4%BE%8B%E6%A0%B7%E6%9C%AC%EF%BC%8C1%E4%B8%AA%E8%B4%9F%E4%BE%8B%E6%A0%B7%E6%9C%AC%E3%80%82)

## nested NER嵌套命名实体识别

### 什么是嵌套实体？

一个实体内部包含其他实体。

### FLAT 模型

### BiLSTM-CRF

1. BiLSTM 可以获取到上下文信息，但是输出标签过程得到的是局部最优解，可能会出现BI0I、BB这样的错误。
2. CRF 加入了标签转移的概率，可以解决上述标签序列不匹配的错误，但是没有全局信息。
3. 将两者结合，取长补短。