from fasttext import train\_unsupervised

**训练模型:**

classifier=fasttext.train\_supervised(

训练集路径,

label\_prefix = '\_\_label\_\_', #文本中标签前缀

dim = 100, #词向量维度

epoch = 5, #迭代次数

lr = 0.05, #学习率，一般0.05-0.1

lr\_update\_rate = 100, #学习率更新速率

min\_count = 3, #有效词的最小频数

loss = 'softmax', #损失函数类型，”ns”:归一化的sofmax; ”hs”:hierarchical softmax; “softmax”

word\_ngrams = 2, #n-gram数，一般取1、2、3

ws=5, #n-gram的窗口大小

neg=5, #负采样个数

thread=12, #线程个数

model=”cbow”或”skipgram”,

bucket = 1000000))

classifier.words #输出分词结果

**测试模型：**

result=classifier.test(测试集路径)

result[0]为数据量；result[1]为准确率；result[2]为召回率

**做预测：**

predict=classifier.predict(待预测文本,k=topk标签预测)

predict输出两个元组，第一个元组为类别，第二个元组为对应概率