问句中的第一个mention去除，将问句焦点词的复述添加到问句后，去除“？”号，使用LTP分词，将每个词的embedding加和

Svm svm.SVC(C=4, kernel="rbf", gamma=0.004)

accuracy: 0.920384951881 2104 182

accuracy: 0.951443569554 2175 111

sigmoid

**accuracy: 0.5 1143 1143**

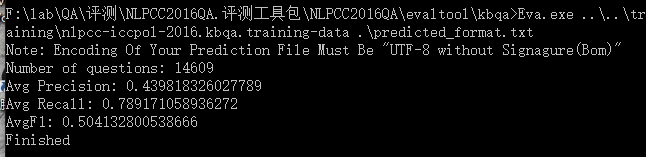
**去除问句中的疑问词，如果焦点词有复述，则使用其中第一个，使用sogou100维向量**

**accuracy:** **0.933508311461 2134 152**

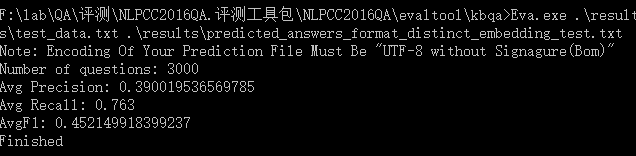
**去除问句中的疑问词，如果焦点词有复述，则使用其中第一个，使用sogou100维向量，添加了错误关系训练**

**accuracy: 0.891197691198 3088 377**

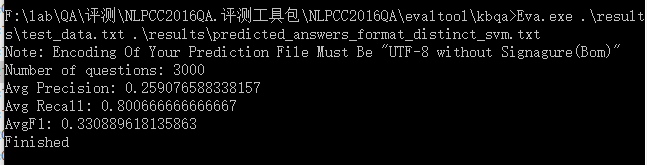
**仅仅使用embedding相似性，每个mention保留一个最终结果**



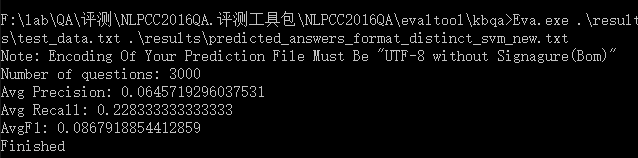
**在测试集上仅仅使用问句和关系的相似性，每个id只保留一项**



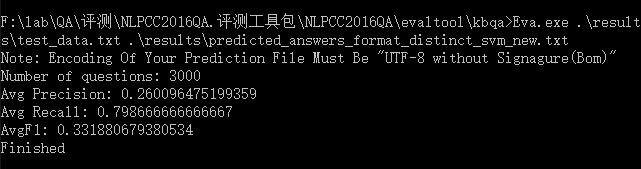
**svm 使用复述训练的模型跑的结果，跑结果中没有使用复述，没有删除第一个mention**



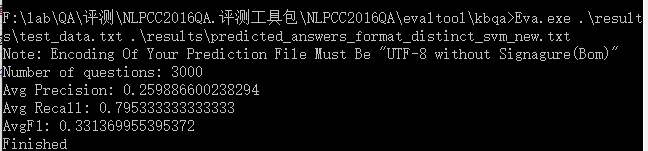
**svm 使用复述训练的模型跑的结果，跑结果中没有使用复述，但是删除第一个mention，从结果中可以看出分类准确率急剧下降**



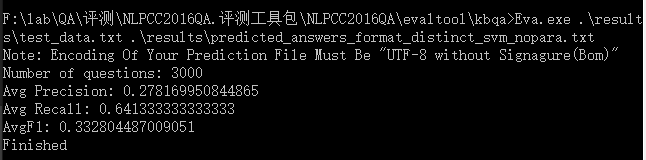
**svm 使用复述训练的模型跑的结果，跑结果中使用复述，没有删除第一个mention**



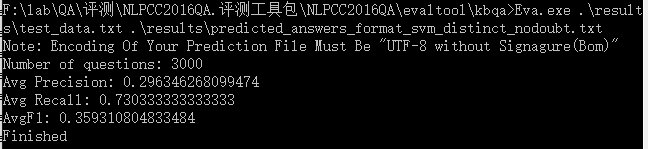
**svm 使用复述训练的模型跑的结果，跑结果中使用焦点词全部复述，没有删除第一个mention**



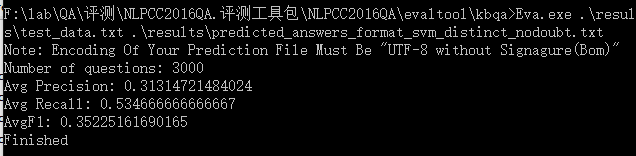
**Svm训练没有使用复述资源，分类中也没有使用焦点词复述，没有删除mention**



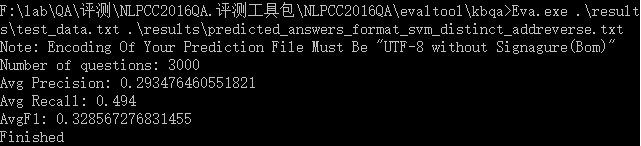
**svm 使用复述训练的模型跑的结果，跑结果中使用复述，没有删除第一个mention，在训练和跑结果过程中去除疑问词。通过结果发现，recall值下降了，但是precision和f1值上升了。**

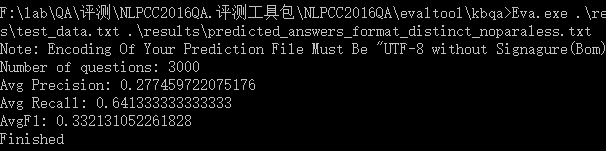


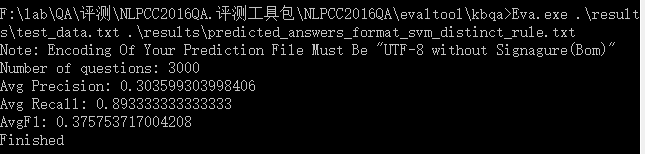
**增加了错误类的训练数据**



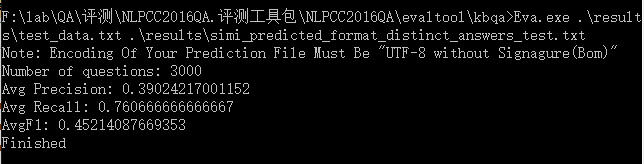
**使用添加反例的分类模型，没有移除问句中疑问词，添加复述词**



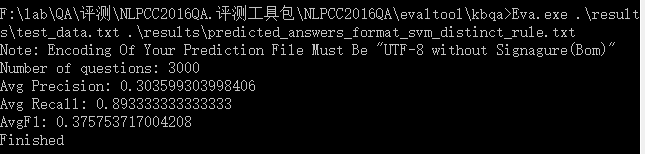


****

**问句中去除mention、疑问词，添加复述，使用问句和关系的相似性**



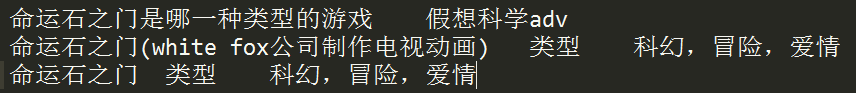
**使用去除第一个mention，使用一个复述的模型， 添加规则结果**

****

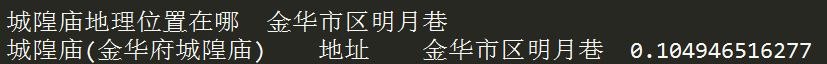
**使用复述训练的model，没有删除疑问词，使用问句焦点词的复述（第一个）。如果关系出现在问句中，则添加该关系，其它不包含关系的三元组全丢弃。**

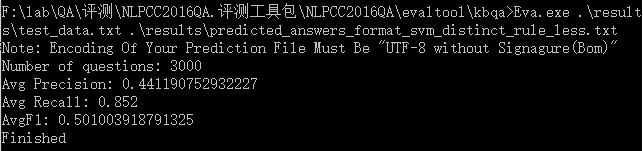
**从结果可以看出，precision和f1值有大幅提升，recall值有下降（下降了4个点），因为存在这种**











如果关系整个名字出现在问句中，则只保留那些关系出现在问句中的

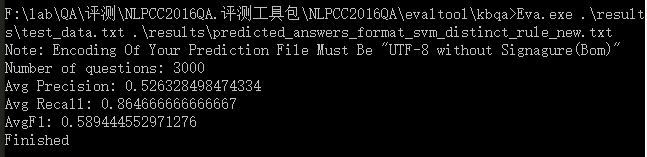
判断关系中是否有出现在问句中的字，如果有则保留

如果这个问句的所有实体没有一个关系中的字出现在问句中，则使用分类器进行分类

如果这个问句还是没有一个答案，则使用embedding相似性

从结果中看出precision值和f1值有较大提升

Recall值比上面的那个结果截图也有提升，只是比89%要低了3个点



如果关系整个名字出现在问句中，则只保留那些关系出现在问句中的

判断关系中是否有出现在问句中的字，如果有则保留

如果这个问句的所有实体没有一个关系中的字出现在问句中，则使用分类器进行分类

如果这个问句还是没有一个答案，则使用embedding相似性

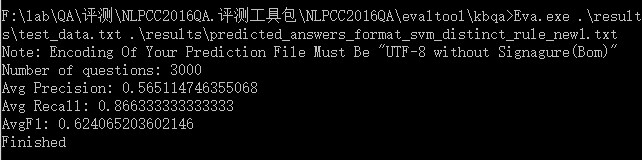
从结果中看出precision值和f1值有较大提升

Recall值比上面的那个结果截图也有提升，只是比89%要低了3个点

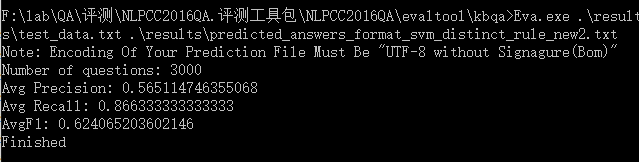
判断关系是否和问句有重复字，这里比较的问句是去除问句中抽取的第一个实体的

三个指标都有提升，recall有提升证明之前有三元组中关系有字出现在实体中，这影响了其它正确答案的抽取，这种情况使用分类或者embedding相似度可以一定程度获取正确答案。

使用model svm\_embedding\_model.txt



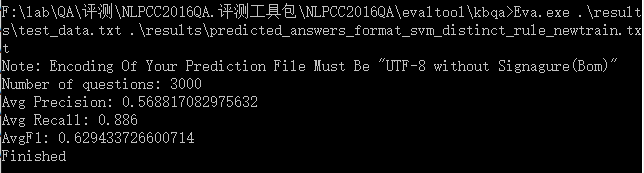
**使用**svm\_embedding\_model.txt，**当没有回答时使用embedding similarity获取答案处代码也有修改**

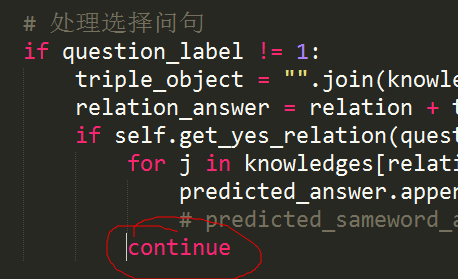


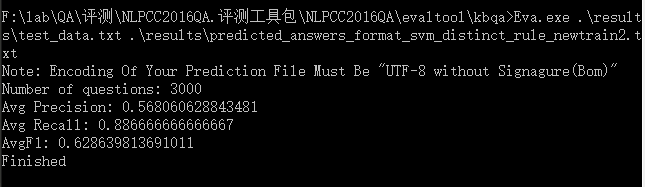
**使用新训练数据训练，****当没有回答时使用embedding similarity获取答案处代码也有修改**

**使用model**

**svm\_sogou\_newtrain.model**







**if (relation in i) and (len(list(relation)) != 1):**

**continue** 