**There are two ways to add regular expression in NER**

1. Gazette
   1. add entries in a gazette file in the format – ‘CLASS this is an example’.
   2. two ways to match the entry, cleanGazette, sloppyGazette.
   3. It is not strictly necessary that classes in gazette must represent the classes the model is trained to recognize.
   4. gazettes only provide another feature for CRF classifier, it don’t guarantee to label all words in the gazette with expected class. And if the gazette features are triggered, words out of the gazette might be also labeled as classes in gazette.
   5. Only the words in gazette show up in the train file, the gazette would help recognize entities.
   6. Usage: Provided with gazette files in the format described in ‘a’ and declare to use gazette features in prop file - gazette=austen.gaz.txt, then train a new model.
   7. ‘PERS-GAZ2|C PERS-GAZ|C’ is what a gazette feature looks like.
2. RegexNER
   1. It is one of the basic annotators in Stanford core nlp.
   2. The format of regexner file is

Bachelor of Arts DEGREE

Bachelor of (Arts|Laws|Science|Engineering|Divinity) DEGREE

* 1. Command ‘ java -cp "\*" -Xmx2g edu.stanford.nlp.pipeline.StanfordCoreNLP -annotators tokenize,ssplit,pos,lemma,regexner -file input.txt -regexner.mappping jgregexner.txt’ can be an overlay of StanfordNER and the only recognized entities are the one in the pattern file – jgregexner.txt
  2. Command ‘ java -cp "\*" -Xmx2g edu.stanford.nlp.pipeline.StanfordCoreNLP -annotators tokenize,ssplit,pos,lemma,regexner,ner -file input.txt -regexner.mappping jgregexner.txt’ can use the results of regexner to relabel results of ner. Only we should to do is declare the entities could be relabeled in the pattern file. The format is (LOCATION is the correct label and PERSON is the entity could be relabeled.)

Lalor LOCATION PERSON

* 1. When using the relabeled method provided by Stanford core nlp, I can not find the way to replace ner model with user-defined ner model. But we can use regexner as a post-processing method and add the relabeled function by self.