

CSE 5320/7320

Assignment 11

Due: Sunday Apr 28

The NP grammar rule we explored in class

"NP: {<DT>*<JJ>*<NN.*>}"

Will not match the following noun phrases (show with their tags)

- earlier/JJR stages/NNS
- His/PRP\$ Mansion/NNP House/NNP speech/NN
- the/DT fastest/JJS developing/VBG trends/NNS

Come up with a grammar rule that will identify all the above as Noun Phrases (NP)

Use the following code to jump start your effort.

Note: this code was run in PyCharm using the Anaconda distribution.

```
import nltk
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')

text = "The fantastic Rollo went to the expensive store to buy the candy."

tokens = nltk.word_tokenize(text)
print (tokens)

tags = nltk.pos_tag(tokens)
print (tags)

grammar = "NP: {<DT>*<JJ>*<NN.*>}"
cp = nltk.RegexpParser(grammar)

result = cp.parse(tags)
print (result)
```

Submit:

- a PDF of your code
- the output showing the NP