Case I have covered to find address

* First check in last 10 lines : IF not found then First 10 lines
* Capture the line where we have address word
* Split the address line in the list using “,” as a delimiter
* If last element of list is our city name and having “.” Then using replace method of string to replace “.” Into space and then using strip method to trim spaces.
* Then checking each element of list and comparing it with Cities Name list
* If I have name of city in address line but words are delimited by “,” but city name is not delimited with “,” then we will use word tokenizing on each element of list and will check word by word.

Code:

for sentence in lines:#find the index of the sentence where the degree is find and then analyse that sentence

sen=" ".join([words[0].lower() for words in sentence]) #string of words in sentence

if re.search('address',sen):

if len(sen)>0:

lines1 = [(el.replace(".","")).strip() for el in sen.split(",") if len(el) > 0]

#print(lines2)

if len(lines1[1:])>0:

for new\_sen in lines1[-2:]:

# print(new\_sen)

if new\_sen in CitiesTown:

print("Address Found:"+new\_sen.upper())

else:

sen\_tokenised1= nltk.word\_tokenize(new\_sen)

for new\_sen2 in sen\_tokenised1:

if new\_sen2 in CitiesTown:

print("Address Found:"+new\_sen2.upper())

else:

print("Address Not Found:"+new\_sen2)