Telecom Software lab EEP 773

Assignment 9

Learning Python and git

Submitted by:
Nitin K. Lohar
2012EEZ8074

September 24, 2014

Contents

Chapter 1

Problem statement

Write python program to find emotions of users based on the different smileys occurring in the text and looking up emotion of that smiley from a given dictionary. Also, output percentage of each emotion occurring in the text.

Chapter 2

Solution

We have written python script for the problem. Script written is given in later sections.

2.1 Assumptions

We assume that smiley can occur only after a white space or a full stop. Smiley occurring just after a word (without a space) is not considered as a smiley. Although, there can be a word just after the smiley. Similarly, we assume that ;);) represents sarcastic mood. Any other smiley repeating without space or dot is ignored.
While deciding mood of a user, we count number of occurrence for each mood and then whichever mood has highest count, we assume that mood. If two mood have same max occurrence, we choose randomly.

2.2 Logic

We have created a dictionary first using the file "mood_dict.txt". After this, we read line by line contents of "contents.txt" file and search for regular expression of smileys. Then we add those occurrences to the user database. Finally, we sort the user database according to the occurrence of mood and write the max occurred mood to the file. Then we write occurrence of each mood in the file in percentage.

Chapter 3

Results

The code files and output files are attached. Also, same have been pushed to the github account.

3.1 Python code for the task

```
#!/usr/bin/python

import re

# create a dictionary from the file "filename"
def create_dict(filename):
    f = open(filename, mode='rU')
    lst = []
    for line in f:
        lst.append(line)
    f.close()
    i = 0
    dictionary = {}
    emotions = []
    for string in lst:
        words = string.split()
        if len(words) < 1: continue
        if i == 0:
            for word in words:
                emotions.append(word)
        else:
            j = 0
            for word in words:
                dictionary[word] = emotions[j]
                j+=1
        i+=1
    if ':P' in dictionary.keys():
        dictionary[':p']=dictionary[':P']
    dictionary[';);)']='Crook'
    return (emotions, dictionary)


def read_file(dict_file,content_file):
    total = 0
    (moods, dictionary) = create_dict(dict_file)
    f = open(content_file, mode='rU')
```

```
37      dict2 = {}  # Dictionary to keep track of users' mood
        mood_count={}   # keep track of number of times a mood occurred
39      for mood in moods:  # Initialize a dictionary to keep count of
        each mood
            mood_count[mood]=0
41      for line in f:
            user = line[0]
43          #print user, (user in dict2.keys())
            if not(user in dict2.keys()):
45              dict2[user]={}
                for mm in moods:
47                  dict2[user][mm]=0
            found = re.findall(r'[\s\.](:[\)\(PpD]|:\-[o/]|:\'\(|B\-\)|
        x\-\(|;\);\)|;\)|O_O|=_=|>_<)',line);
49          for entry in found:
                mood = dictionary[entry]
51              dict2[user][mood] = dict2[user][mood]+1
                mood_count[mood]+=1
53              total += 1
        f.close()
55      f1 = open("output.txt",'w')
#       print mood_count.items()
57      for entry in dict2.keys():
            res = sorted(dict2[entry],key=dict2[entry].get,reverse=True
        ) # If two moods have equal max count, we choose one randomly
59          f1.write(entry+" : "+res[0]+"\n")
        f1.write("\n"+"-"*20+"\n")
61      for entry in mood_count:
            f1.write(entry+" : "+repr(mood_count[entry]/float(total)
        *100)+" %\n")
63  #           print dict2[entry]
        f1.close()
65
67
    def main():
69      read_file("mood_dict.txt","content.txt")
#       (moods, dictionary) = create_dict("mood_dict.txt")
71  #       print dictionary.items()
    #       print moods
73
    if __name__ == '__main__':
75      main()
```

assgn9.py