

Course – Machine Learning for Trading

Project - 2014Fall7646 Project 1 A

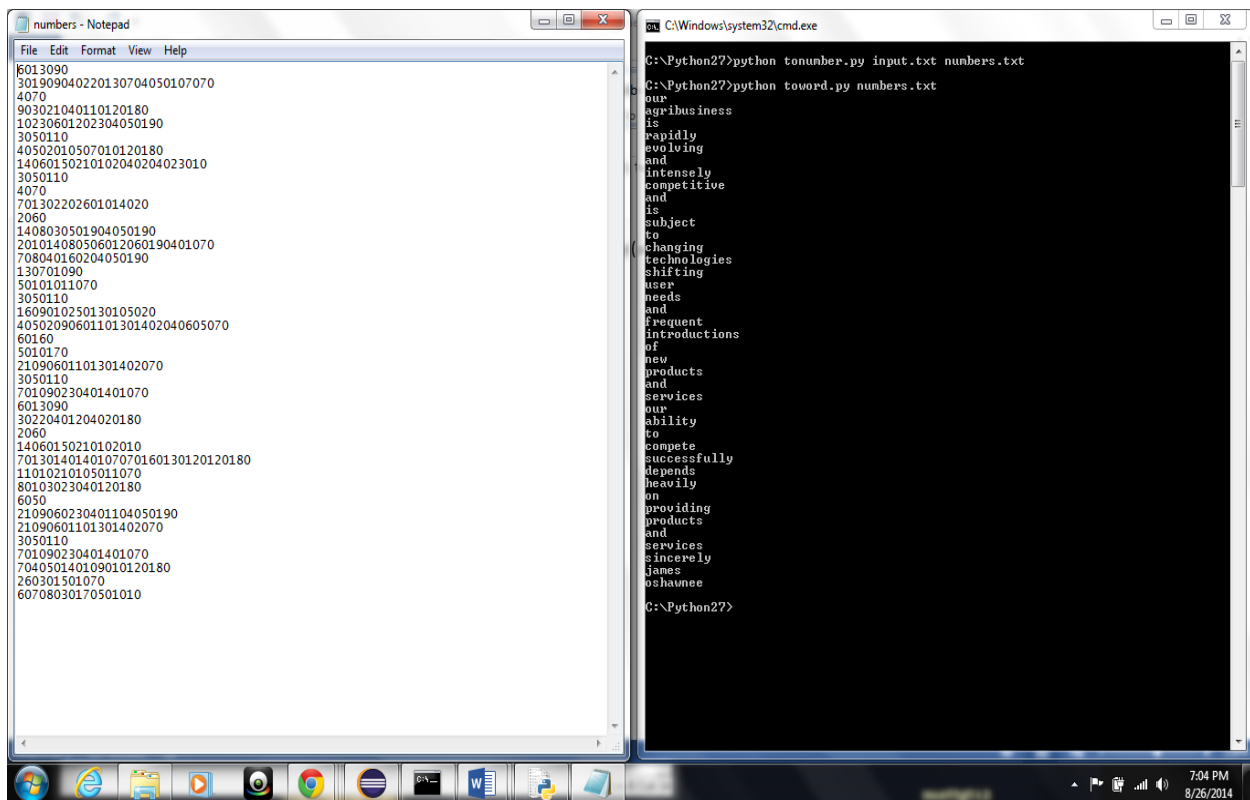
Student Name – Aishvarya Krishnan

In this assignment, I am using number '0' as delimiter and a static dictionary with numbers (not containing 0) assigned for all alphabets.

The encoding scheme in tonumber.py - For every word in the input.txt, I convert the alphabet to a number as mapped in the dictionary, add delimiter '0' after every number and concatenate them to represent the word as a unique number.

The decoding scheme in toword.py – For every number, I split the number at '0', find the corresponding alphabet in dictionary and concatenate the alphabets to retrieve the word.

Output of tonumber.py (numbers.txt) and toword.py (printed on cmd prompt)–



The screenshot displays a Windows desktop environment. On the left, a Notepad window titled 'numbers - Notepad' contains a list of numbers generated by the tonumber.py script. These numbers are unique encodings of words, created by concatenating the numeric values of each letter in a word, separated by '0'. For example, 'our' is encoded as '6013090'. The list includes words like 'our', 'agribusiness', 'rapidly', 'evolving', 'and', 'intensely', 'competitive', 'subject', 'to', 'changing', 'technologies', 'shifting', 'user', 'needs', 'and', 'frequent', 'introductions', 'of', 'new', 'products', 'and', 'services', 'our', 'ability', 'to', 'compete', 'successfully', 'depends', 'heavily', 'on', 'providing', 'products', 'and', 'services', 'sincerely', 'james', 'oshaunee'. On the right, a Command Prompt window titled 'C:\Windows\system32\cmd.exe' shows the execution of two Python scripts. The first command is 'C:\Python27>python tonumber.py input.txt numbers.txt', which successfully generates the numbers.txt file. The second command is 'C:\Python27>python toword.py numbers.txt', which prints the decoded words to the command prompt: 'our', 'agribusiness', 'rapidly', 'evolving', 'and', 'intensely', 'competitive', 'subject', 'to', 'changing', 'technologies', 'shifting', 'user', 'needs', 'and', 'frequent', 'introductions', 'of', 'new', 'products', 'and', 'services', 'our', 'ability', 'to', 'compete', 'successfully', 'depends', 'heavily', 'on', 'providing', 'products', 'and', 'services', 'sincerely', 'james', 'oshaunee'. The system clock in the bottom right corner indicates the time is 7:04 PM on 8/26/2014.

```
File Edit Format View Help
6013090
301909040220130704050107070
4070
903021040110120180
10230601202304050190
3050110
40502010507010120180
14060150210102040204023010
3050110
4070
701302202601014020
2060
1408030501904050190
201014080506012060190401070
708040160204050190
130701090
50101011070
3050110
1609010250130105020
40502090601101301402040605070
60160
5010170
21090601101301402070
3050110
701090230401401070
6013090
30220401204020180
2060
14060150210102010
70130140140107070160130120180
11010210105011070
80103023040120180
6050
2109060230401104050190
21090601101301402070
3050110
701090230401401070
704050140109010120180
260301501070
60708030170501010

C:\Windows\system32\cmd.exe
C:\Python27>python tonumber.py input.txt numbers.txt
C:\Python27>python toword.py numbers.txt
our
agribusiness
is
rapidly
evolving
and
intensely
competitive
and
is
subject
to
changing
technologies
shifting
user
needs
and
frequent
introductions
of
new
products
and
services
our
ability
to
compete
successfully
depends
heavily
on
providing
products
and
services
sincerely
james
oshaunee
C:\Python27>
```