VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

CSE4022 – Natural Language Processing

Project Report

TITLE – RHYMER DICTIONARY

B.Tech(IT),Winter -2018-19

Submitted By:

Ritveak Dugar
(16BIT0407)

OBJECTIVE:

The objective of this project is to find a set of rhyming words with a particular word which is somehow related to an another given word.

List of Libraries used:

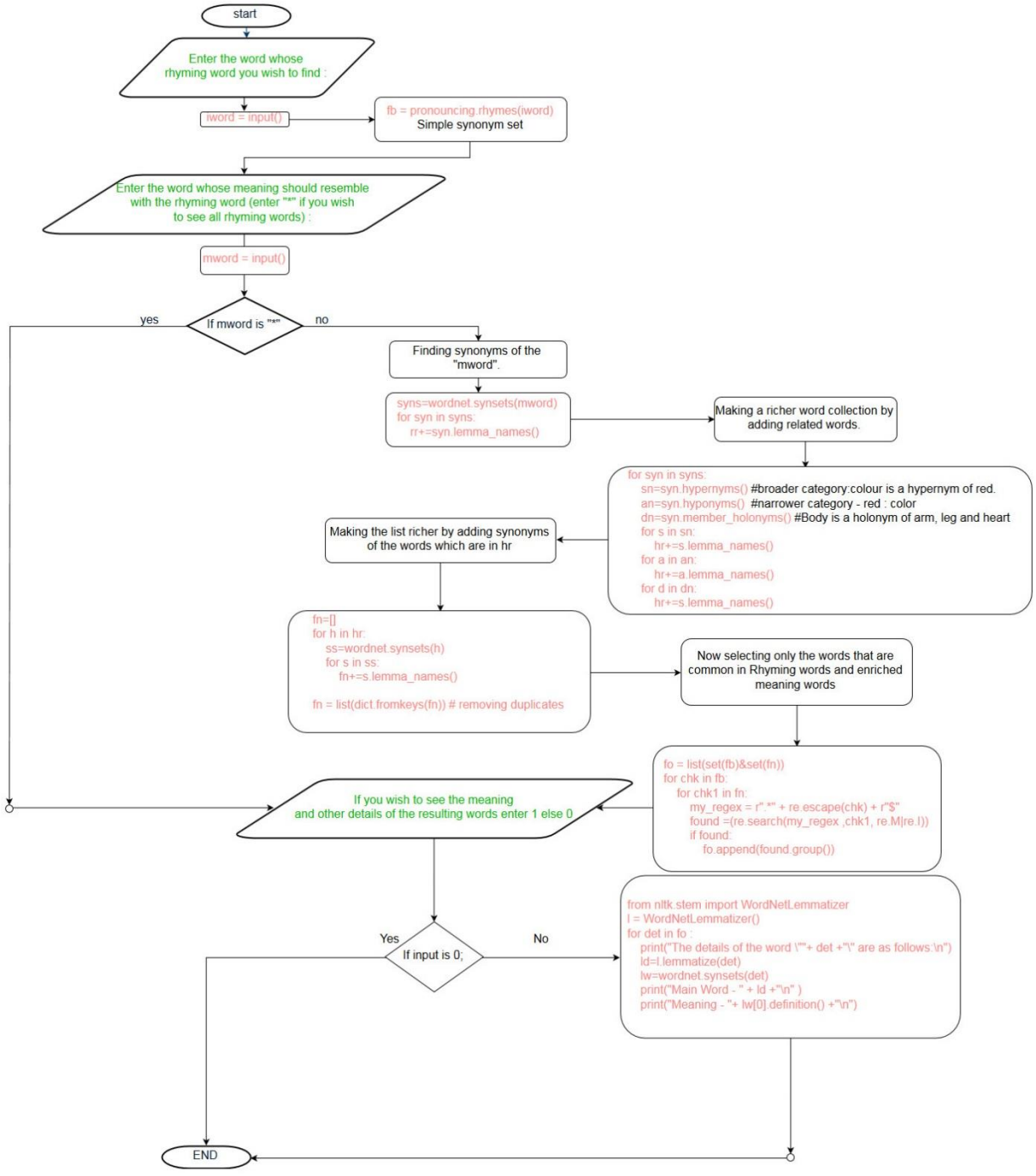
1. Pronouncing: Used for finding rhyming words, this library uses cmudict as corpus reference to find phonetically resembling words to a given word.
2. re: Required for using regular expression for string comparison.
3. sys: Required for using sys.exit(), to exit the program.
4. wordnet: Required for finding synonyms, hypernyms, hyponyms, holonyms and meaning of words.

Code:

```
import pronouncing
import re
import sys
from nltk.corpus import wordnet
import nltk
# a proper module separately made for finding rhyming words; also based on
cmudict.
print("Enter the word whose rhyming word you wish to find : ")
iword = input()

fb = pronouncing.rhymes(iword)
#Simple synonym set
rr=[]
print("Enter the word whose meaning should resemble with the rhyming word
\n(enter \"*\" if you wish to see all rhyming words) : ")
mword = input()
if(mword=="*"):
    print(fb)
    sys.exit()
syns=wordnet.synsets(mword)
for syn in syns:
    rr+=syn.lemma_names()

#Since simple synonym set is not enough, lets add more related words
#finding all related words among which rhyming words is to be found

hr=[]
syns=wordnet.synsets(mword)
for syn in syns:
```

```
    sn=syn.hypernyms()#broader category:colour is a hypernym of red.
    an=syn.hyponyms() #narrower category - red : color
    dn=syn.member_holonyms()#Body is a holonym of arm, leg and heart
    for s in sn:
        hr+=s.lemma_names()
    for a in an:
        hr+=a.lemma_names()
    for d in dn:
        hr+=s.lemma_names()

#making the list richer by adding synonyms of the words which are in hr.
fn=[]
for h in hr:
    ss=wordnet.synsets(h)
    for s in ss:
        fn+=s.lemma_names()

fn = list(dict.fromkeys(fn)) # removing duplicates
#now selecting only the words that are common in both
import re
fo = list(set(fb)&set(fn))
for chk in fb:
    for chk1 in fn:
        my_regex = r".*" + re.escape(chk) + r"$"
        found =(re.search(my_regex ,chk1, re.M|re.I))
        if found:
            fo.append(found.group())
fo=list(set(fo))

print(fo)
mean=input("If you wish to see the meaning and other details \nof the resulting
words enter 1 else 0")

if(mean=="0"):
    sys.exit()
from nltk.stem import WordNetLemmatizer
l = WordNetLemmatizer()
for det in fo :
    print("The details of the word \""+ det +"\" are as follows:\n")
    ld=l.lemmatize(det)
    lw=wordnet.synsets(det)
    print("Main Word - " + ld +"\n" )
    print("Meaning - "+ lw[0].definition() +"\n")
```

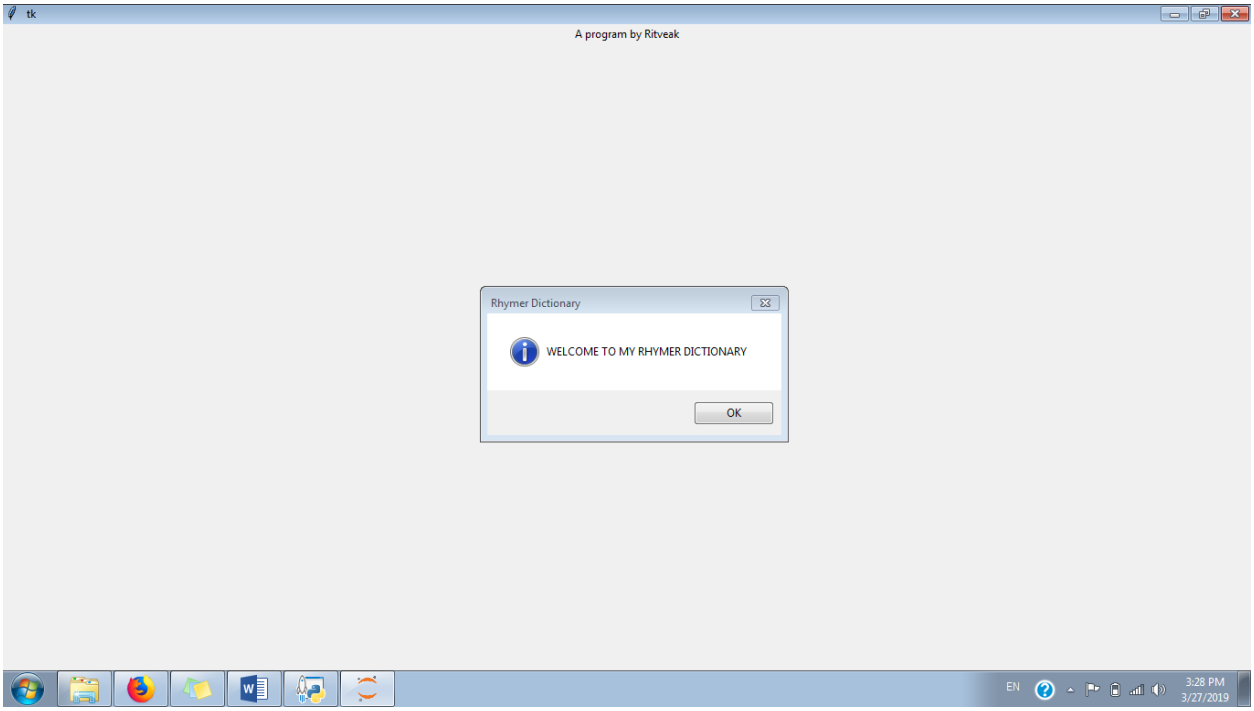
Flowchart:



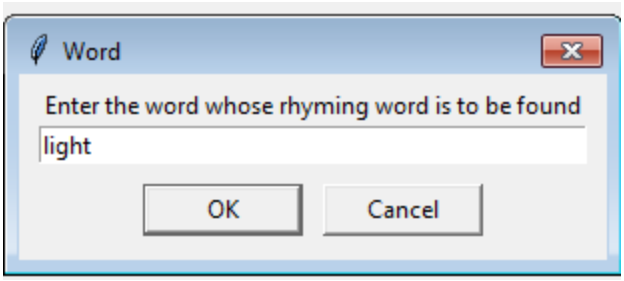
Github Link : https://github.com/ritveak/Rhymer-Dictionary

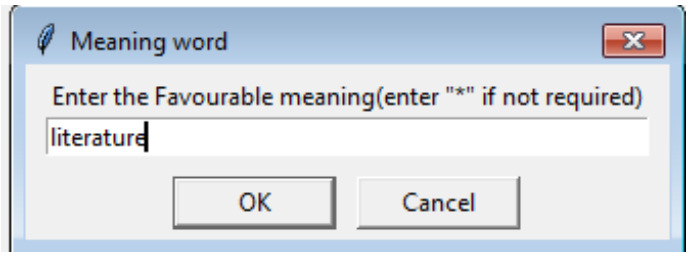
Screenshot of an example:

Welcome page:

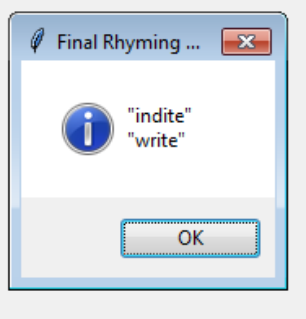


Taking Inputs:

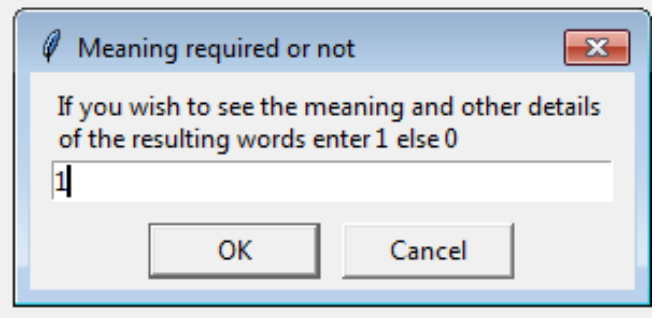




Getting list of words:



Opting for more info:



More info :