return words

File 🡨 open(filename)

File\_content 🡨 read（File）

Close File

Words 🡨 File\_content split

Extract\_words\_from\_file(filename)

Find\_element(alist, item)

loop

the No. int(len(alist)/2) word = item

yes

No

the No. int(len(alist)/2) word < item

Alist 🡨 the second half of Alist

yes

No

the No. int(len(alist)/2) word > item

Alist 🡨 the first half of Alist

yes

yes

No

len(alist) <=1

No

yes

return words

Text\_to\_words(text\_list)

Words 🡨 []

loop

i 🡨 the i\_th element of text\_list

My\_substitutions 🡨 i.maketrans(“ABCD…”\\”abcd…”

Cleaned\_text 🡨 i.translate(my\_substtutions)

Wds 🡨 split cleaned\_text

Words 🡨 words + wds

all the words in the list are checked

No

yes

return word\_list

return alist

Uniquelize(alist)

Join a list with space into a string

Split the string, in order to remove the possible space in the list of words

Turn the list into a set, and turn the set back to a list, in order to remove the possible duplicate words in the list

Word\_not\_in\_count(file, vocab)

Extract\_words\_from\_file，to obtain a text\_list from the book file

Text\_to\_words，to obtain a word\_list

Uniquelize, to remove the duplicate words in word\_list

Extract\_words\_from\_file，obtain a vocabulary\_list from the vocabulary file

loop

i 🡨 the i\_th element of words

find\_element(vocab, i)

yes

No

append i to words\_not\_in

All the words are checked

No

yes

return len(words\_not\_in)