import pandas as pd

extraction = pd.read\_csv("total with label.csv",encoding='utf-8')

from soynlp.word import WordExtractor

word\_extractor = WordExtractor(

min\_frequency = 100,

min\_cohesion\_forward = 0.05,

min\_right\_branching\_entropy = 0)

word\_extractor.train(extraction.tweet)

words = word\_extractor.extract()

import math

def word\_score(score):

return (score.cohesion\_forward \* math.exp(score.right\_branching\_entropy))

wordsforLst = []

leftside\_frequency = []

cohesion\_forward = []

right\_branching\_entropy = []

for word, score in sorted(words.items(), key=lambda x:word\_score(x[1]), reverse=True)[:]:

wordsforLst.append(word)

leftside\_frequency.append(score.leftside\_frequency)

cohesion\_forward.append(score.cohesion\_forward)

right\_branching\_entropy.append(score.right\_branching\_entropy)

WordsLst= pd.DataFrame({'words':wordsforLst,'leftside\_frequncy':leftside\_frequency,'cohesion\_forward':cohesion\_forward,'right\_branching\_entropy':right\_branching\_entropy})

WordsLst.to\_csv("worfdLst\_sonlpy.csv",encoding = 'euc-kr')