**Work2vec**

**1.訓練模型**

import warnings

warnings.filterwarnings(action = 'ignore', category = UserWarning, module = 'gensim')

from gensim.models import word2vec

# 主要透過 gensim 訓練成 model 並供使用

class Train(object):

def \_\_init\_\_(self):

pass

# 可參考 https://radimrehurek.com/gensim/models/word2vec.html 更多運用

def train(self):

print("訓練中...(喝個咖啡吧^0^)")

# Load file

sentence = word2vec.Text8Corpus("word1")

# Setting degree and Produce Model(Train)

model = word2vec.Word2Vec(sentence, size = 50, window = 10, min\_count = 5, workers = 4, sg = 1)

# Save model

model.wv.save\_word2vec\_format(u"makeup.model.bin", binary = True)

print("model 已儲存完畢")

if \_\_name\_\_ == "\_\_main\_\_":

t = Train()

# 訓練(shallow semantic space)

t.train()

**2.搜尋詞向量**

import warnings

warnings.filterwarnings(action = 'ignore', category = UserWarning, module = 'gensim')

from gensim.models.keyedvectors import KeyedVectors

# 載入 model 並去運用

def main():

# 可參考 https://radimrehurek.com/gensim/models/word2vec.html 更多運用

# How to use bin(model)?

word\_vectors = KeyedVectors.load\_word2vec\_format("makeup.model.bin", binary = True)

# print("'歐美'前10名相似:")

a = input('關鍵字:')

res = word\_vectors.wv.most\_similar(a, topn = 20)

for item in res:

print(item[0] + "," + str(item[1]))

if \_\_name\_\_ == "\_\_main\_\_":

main()