

Lab-3

March 13, 2024

1) Import Libraries

```
[ ]: import pandas as pd
import re
import string
import nltk
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
```

2) Load Data

```
[ ]: data = pd.read_csv("train.csv")
data
```

```
[ ]:
      id keyword location \
0      1      NaN      NaN
1      4      NaN      NaN
2      5      NaN      NaN
3      6      NaN      NaN
4      7      NaN      NaN
...  ...  ...  ...
7608 10869      NaN      NaN
7609 10870      NaN      NaN
7610 10871      NaN      NaN
7611 10872      NaN      NaN
7612 10873      NaN      NaN

      text target
0  Our Deeds are the Reason of this #earthquake M...      1
1      Forest fire near La Ronge Sask. Canada      1
2  All residents asked to 'shelter in place' are ...      1
3  13,000 people receive #wildfires evacuation or...      1
4  Just got sent this photo from Ruby #Alaska as ...      1
...  ...  ...
7608 Two giant cranes holding a bridge collapse int...      1
7609 @aria_ahrar @TheTawniest The out of control w...      1
7610 M1.94 [01:04 UTC]?5km S of Volcano Hawaii. htt...      1
7611 Police investigating after an e-bike collided ...      1
7612 The Latest: More Homes Razed by Northern Calif...      1
```

[7613 rows x 5 columns]

```
[ ]: df = data[['text', 'target']]
df['text'] = df['text'].str.lower()

df
```

C:\Users\d-kin\AppData\Local\Temp\ipykernel_13348\1695621088.py:2:

SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['text'] = df['text'].str.lower()
```

```
[ ]:
      text  target
0  our deeds are the reason of this #earthquake m...      1
1      forest fire near la ronge sask. canada      1
2  all residents asked to 'shelter in place' are ...      1
3  13,000 people receive #wildfires evacuation or...      1
4  just got sent this photo from ruby #alaska as ...      1
...
7608 two giant cranes holding a bridge collapse int...      1
7609 @aria_ahrury @thetawniest the out of control w...      1
7610 m1.94 [01:04 utc]?5km s of volcano hawaii. htt...      1
7611 police investigating after an e-bike collided ...      1
7612 the latest: more homes razed by northern calif...      1
```

[7613 rows x 2 columns]

3) Clean Data

```
[ ]: url = "https://www.somesite.com This is othe text."

pattern = re.compile(r'https?://\S*')
pattern.findall(url)
```

```
[ ]: ['https://www.somesite.com']
```

```
[ ]: pattern.sub("$$$ ", url)
```

```
[ ]: '$$$ This is othe text.'
```

```
[ ]: df['text'] = df['text'].apply(lambda x: pattern.sub("", x))
df.head(50)
```

C:\Users\d-kin\AppData\Local\Temp\ipykernel_13348\3698504754.py:1:

SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['text'] = df['text'].apply(lambda x: pattern.sub("", x))
```

```
[ ]:                                     text  target
0  our deeds are the reason of this #earthquake m...      1
1          forest fire near la ronge sask. canada      1
2  all residents asked to 'shelter in place' are ...      1
3  13,000 people receive #wildfires evacuation or...      1
4  just got sent this photo from ruby #alaska as ...      1
5  #rockyfire update => california hwy. 20 closed...      1
6  #flood #disaster heavy rain causes flash flood...      1
7  i'm on top of the hill and i can see a fire in...      1
8  there's an emergency evacuation happening now ...      1
9  i'm afraid that the tornado is coming to our a...      1
10         three people died from the heat wave so far      1
11 haha south tampa is getting flooded hah- wait ...      1
12 #raining #flooding #florida #tampabay #tampa 1...      1
13         #flood in bago myanmar #we arrived bago      1
14 damage to school bus on 80 in multi car crash ...      1
15                               what's up man?      0
16                               i love fruits      0
17                               summer is lovely      0
18                               my car is so fast      0
19                               what a goooooooooaaaaa!!!!!!      0
20                               this is ridiculous...      0
21                               london is cool ;)      0
22                               love skiing      0
23                               what a wonderful day!      0
24                               loooooool      0
25                               no way...i can't eat that shit      0
26                               was in nyc last week!      0
27                               love my girlfriend      0
28                               coool :)      0
29                               do you like pasta?      0
30                               the end!      0
31                               @bbcmtd wholesale markets ablaze      1
32         we always try to bring the heavy. #metal #rt      0
33 #africanbaze: breaking news:nigeria flag set a...      1
34         crying out for more! set me ablaze      0
35 on plus side look at the sky last night it was...      0
36 @phdsquares #mufc they've built so much hype a...      0
37         inec office in abia set ablaze -      1
```

```

38 barbados #bridgetown jamaica ûò two cars set ... 1
39 ablaze for you lord :d 0
40 check these out: #nsfw 0
41 on the outside you're ablaze and alive\nbut yo... 0
42 had an awesome time visiting the cfc head offi... 0
43 soooo pumped for ablaze ???? @southridgelife 0
44 i wanted to set chicago ablaze with my preachi... 0
45 i gained 3 followers in the last week. you? kn... 0
46 how the west was burned: thousands of wildfire... 1
47 building the perfect tracklist to life leave t... 0
48 check these out: #nsfw 0
49 first night with retainers in. it's quite weir... 0

```

```
[ ]: string.punctuation
```

```
[ ]: '!"#$%&\'()*+,-./:;<=>?@[\\]^_`{|}~'
```

```
[ ]: text = "Hello $# world ?% find *() this <>"
pattern2 = re.compile(r'[\w\s]')
pattern2.findall(text)
```

```
[ ]: ['$', '#', '?', '%', '*', '(', ')', '<', '>']
```

```
[ ]: df['text'] = df['text'].apply(lambda x: pattern2.sub("", x))
df.head(50)
```

C:\Users\d-kin\AppData\Local\Temp\ipykernel_13348\754453846.py:1:

SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['text'] = df['text'].apply(lambda x: pattern2.sub("", x))
```

```
[ ]:
      text  target
0  our deeds are the reason of this earthquake ma... 1
1      forest fire near la ronge sask canada 1
2  all residents asked to shelter in place are be... 1
3  13000 people receive wildfires evacuation orde... 1
4  just got sent this photo from ruby alaska as s... 1
5  rockyfire update  california hwy 20 closed in ... 1
6  flood disaster heavy rain causes flash floodin... 1
7  im on top of the hill and i can see a fire in ... 1
8  theres an emergency evacuation happening now i... 1
9  im afraid that the tornado is coming to our area 1
10 three people died from the heat wave so far 1
11 haha south tampa is getting flooded hah wait a... 1

```

12	raining flooding florida tampabay tampa 18 or ...	1
13	flood in bago myanmar we arrived bago	1
14	damage to school bus on 80 in multi car crash ...	1
15	whats up man	0
16	i love fruits	0
17	summer is lovely	0
18	my car is so fast	0
19	what a goooooooooaaaaaal	0
20	this is ridiculous	0
21	london is cool	0
22	love skiing	0
23	what a wonderful day	0
24	loooooool	0
25	no wayi cant eat that shit	0
26	was in nyc last week	0
27	love my girlfriend	0
28	coool	0
29	do you like pasta	0
30	the end	0
31	bbcmttd wholesale markets ablaze	1
32	we always try to bring the heavy metal rt	0
33	africanbaze breaking newsnigeria flag set abla...	1
34	crying out for more set me ablaze	0
35	on plus side look at the sky last night it was...	0
36	phdsquares mufc theyve built so much hype arou...	0
37	inec office in abia set ablaze	1
38	barbados bridgetown jamaica ùò two cars set ab...	1
39	ablaze for you lord d	0
40	check these out nsfw	0
41	on the outside youre ablaze and alive\nbut you...	0
42	had an awesome time visiting the cfc head offi...	0
43	soooo pumped for ablaze southridgelife	0
44	i wanted to set chicago ablaze with my preachi...	0
45	i gained 3 followers in the last week you know...	0
46	how the west was burned thousands of wildfires...	1
47	building the perfect tracklist to life leave t...	0
48	check these out nsfw	0
49	first night with retainers in its quite weird ...	0

4) Stemming of the vocabulary

```
[ ]: [ _ for _ in text.split()]
```

```
[ ]: ['Hello', '$#', 'world', '?%', 'find', '*()', 'this', '<>']
```

```
[ ]: df['text'] = df['text'].apply(lambda x: [ _ for _ in x.split()])
df['text'].head(20)
```

C:\Users\d-kin\AppData\Local\Temp\ipykernel_13348\2964385030.py:1:

SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['text'] = df['text'].apply(lambda x: [_ for _ in x.split()])
```

```
[ ]: 0      [our, deeds, are, the, reason, of, this, earth...
      1      [forest, fire, near, la, ronge, sask, canada]
      2      [all, residents, asked, to, shelter, in, place...
      3      [13000, people, receive, wildfires, evacuation...
      4      [just, got, sent, this, photo, from, ruby, ala...
      5      [rockyfire, update, california, hwy, 20, close...
      6      [flood, disaster, heavy, rain, causes, flash, ...
      7      [im, on, top, of, the, hill, and, i, can, see,...
      8      [theres, an, emergency, evacuation, happening,...
      9      [im, afraid, that, the, tornado, is, coming, t...
     10      [three, people, died, from, the, heat, wave, s...
     11      [haha, south, tampa, is, getting, flooded, hah...
     12      [raining, flooding, florida, tampabay, tampa, ...
     13      [flood, in, bago, myanmar, we, arrived, bago]
     14      [damage, to, school, bus, on, 80, in, multi, c...
     15      [whats, up, man]
     16      [i, love, fruits]
     17      [summer, is, lovely]
     18      [my, car, is, so, fast]
     19      [what, a, goooooooooaaaaaal]
Name: text, dtype: object
```

```
[ ]: # nltk.download('stopwords')
eng_stop = stopwords.words('English')
eng_stop
```

```
[ ]: ['i',
      'me',
      'my',
      'myself',
      'we',
      'our',
      'ours',
      'ourselves',
      'you',
      "you're",
      "you've",
      "you'll",
      "you'd",
```

'your',
'yours',
'yourself',
'yourselves',
'he',
'him',
'his',
'himself',
'she',
"she's",
'her',
'hers',
'herself',
'it',
"it's",
'its',
'itself',
'they',
'them',
'their',
'theirs',
'themselves',
'what',
'which',
'who',
'whom',
'this',
'that',
"that'll",
'these',
'those',
'am',
'is',
'are',
'was',
'were',
'be',
'been',
'being',
'have',
'has',
'had',
'having',
'do',
'does',
'did',
'doing',

'a',
'an',
'the',
'and',
'but',
'if',
'or',
'because',
'as',
'until',
'while',
'of',
'at',
'by',
'for',
'with',
'about',
'against',
'between',
'into',
'through',
'during',
'before',
'after',
'above',
'below',
'to',
'from',
'up',
'down',
'in',
'out',
'on',
'off',
'over',
'under',
'again',
'further',
'then',
'once',
'here',
'there',
'when',
'where',
'why',
'how',
'all',

'any',
'both',
'each',
'few',
'more',
'most',
'other',
'some',
'such',
'no',
'nor',
'not',
'only',
'own',
'same',
'so',
'than',
'too',
'very',
's',
't',
'can',
'will',
'just',
'don',
"don't",
'should',
"should've",
'now',
'd',
'll',
'm',
'o',
're',
've',
'y',
'ain',
'aren',
"aren't",
'couldn',
"couldn't",
'didn',
"didn't",
'doesn',
"doesn't",
'hadn',
"hadn't",

```
'hasn',
'hasn't',
'haven',
'haven't',
'isn',
'isn't',
'ma',
'mightn',
'mightn't',
'mustn',
'mustn't',
'needn',
'needn't',
'shan',
'shan't',
'shouldn',
'shouldn't',
'wasn',
'wasn't',
'weren',
'weren't',
'won',
'won't',
'wouldn',
'wouldn't']
```

```
[ ]: [_ for _ in text.split() if _ not in eng_stop]
```

```
[ ]: ['Hello', '$#', 'world', '?%', 'find', '*()', '<>']
```

```
[ ]: df['text'] = df['text'].apply(lambda x: [_ for _ in x if _ not in eng_stop])
df['text'].head(20)
```

C:\Users\d-kin\AppData\Local\Temp\ipykernel_13348\3909481511.py:1:

SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['text'] = df['text'].apply(lambda x: [_ for _ in x if _ not in eng_stop])
```

```
[ ]: 0      [deeds, reason, earthquake, may, allah, forgiv...
      1      [forest, fire, near, la, ronge, sask, canada]
      2      [residents, asked, shelter, place, notified, o...
      3      [13000, people, receive, wildfires, evacuation...
      4      [got, sent, photo, ruby, alaska, smoke, wildfi...
      5      [rockyfire, update, california, hwy, 20, close...
```

```

6      [flood, disaster, heavy, rain, causes, flash, ...
7          [im, top, hill, see, fire, woods]
8      [theres, emergency, evacuation, happening, bui...
9          [im, afraid, tornado, coming, area]
10         [three, people, died, heat, wave, far]
11      [haha, south, tampa, getting, flooded, hah, wa...
12      [raining, flooding, florida, tampabay, tampa, ...
13         [flood, bago, myanmar, arrived, bago]
14      [damage, school, bus, 80, multi, car, crash, b...
15          [whats, man]
16          [love, fruits]
17          [summer, lovely]
18          [car, fast]
19          [goooooooooaaaaaal]
Name: text, dtype: object

```

```
[ ]: st = PorterStemmer()
```

```
[ ]: st.stem("running")
```

```
[ ]: 'run'
```

```
[ ]: [st.stem(_) for _ in text.split()]
```

```
[ ]: ['hello', '$#%', 'world', '?%', 'find', '*()', 'thi', '<>']
```

```
[ ]: df['text'] = df['text'].apply(lambda x: [st.stem(_) for _ in x])
df['text'].head(20)
```

C:\Users\d-kin\AppData\Local\Temp\ipykernel_13348\1317486952.py:1:

SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df['text'] = df['text'].apply(lambda x: [st.stem(_) for _ in x])
```

```

[ ]: 0      [deed, reason, earthquak, may, allah, forgiv, us]
1          [forest, fire, near, la, rong, sask, canada]
2      [resid, ask, shelter, place, notifi, offic, ev...
3      [13000, peopl, receiv, wildfir, evacu, order, ...
4      [got, sent, photo, rubi, alaska, smoke, wildfi...
5      [rockyfir, updat, california, hwi, 20, close, ...
6      [flood, disast, heavi, rain, caus, flash, floo...
7          [im, top, hill, see, fire, wood]
8      [there, emerg, evacu, happen, build, across, s...
9          [im, afraid, tornado, come, area]

```

```

10             [three, peopl, die, heat, wave, far]
11     [haha, south, tampa, get, flood, hah, wait, se...
12     [rain, flood, florida, tampabay, tampa, 18, 19...
13             [flood, bago, myanmar, arriv, bago]
14     [damag, school, bu, 80, multi, car, crash, break]
15             [what, man]
16             [love, fruit]
17             [summer, love]
18             [car, fast]
19             [goooooooooaaaaaal]
Name: text, dtype: object

```

5) Create Vocabulary Dictionary

```

[ ]: dictionary = []
    for rows in df['text']:
        for word in rows:
            dictionary.append(word)

dictionary = list(set(dictionary))
dictionary

```

```

[ ]: ['strang',
      'cannib',
      '662',
      'tent',
      'hair',
      'sin',
      'andrew',
      'show',
      'jitter',
      'benkin97',
      'edt',
      'barrack',
      'unwant',
      'institut',
      'chpsre',
      'salvag',
      'hashtag',
      'ng2x5',
      '7510880',
      '600',
      'shortfal',
      'panoram',
      'tragedi',
      'yyeso',
      'kendal',
      '21',

```

'nightsad',
'attila',
'ransack',
'snazzychipz',
'flare',
'morehouse64',
'dispensari',
'baduû',
'festac',
'requiem',
'0025',
'charcoal',
'real_liam_payn',
'gr',
'fleek',
'tremblayeh',
'appli',
'fallen',
'ear',
'midweek',
'ooh',
'mk',
'tmake',
'camilacabello97',
'tak',
'revel',
'182624',
'barrier',
'teeth',
'about',
'slam',
'noodl',
'ûihatchet',
'surround',
'khuzdar',
'prefer',
'envis',
'tech',
'stallion150',
'sto',
'worth',
'sirtophamhat',
'2015',
'milita',
'ganseyman',
'9km',
'emv',

'sponsor',
'fang',
'cascad',
'fullscreen',
'nuclear',
'afterhaiyan',
'meelllitt',
'cachet',
'forney',
'teamatowinn',
'randomthought',
'ted',
'stat',
'packag',
'hvnewsnetwork',
'crimin',
'wmata',
'mmda',
'indiefilm',
'helo',
'7speed',
'banquet',
'sink',
'1d',
'massag',
'cutter',
'hinata',
'jones94kyl',
'faro',
'jack',
'jersey',
'grief',
'msnbc',
'nh1new',
'troop',
'bangladesh',
'alexandrapullin',
'bieber',
'groom',
'conclus',
'squeaver',
'loui',
'omw',
'court',
'seismicresist',
'heyimbeeyt',
'chuckswindol',

'otherwisethey',
'eurocrisi',
'lolol',
'gloriavelez',
'regim',
'sheûª',
'greek',
'desper',
'ûhoax',
'ika',
'ici',
'christianvalu',
'maxi',
'stopû',
'madara',
'4tp_kotc',
'optimist',
'snowbal',
'bomber',
'competitor',
'unharm',
'7k',
'pkk',
'coleslaw',
'fuel',
'jenner',
'teeû_',
'area',
'queen',
'4193',
'teddi',
'baruch',
'gotthard',
'kiss',
'pak',
'bernardino',
'righteou',
'gmcr',
'deb117',
'book',
'njenga',
'abandonedp',
'profittothepeopl',
'blinker',
'fli',
'rd',
'kenji',

'fleet',
'rigour',
'raft',
'nasasolarsystem',
'18th',
'yyj',
'striker',
'around',
'devalu',
'la_siren',
'redeem',
'ml',
'oc',
'nake',
'vivian',
'camilla_33',
'lovefood',
'moan',
'human',
'irand',
'atleast',
'emot',
'jamienny',
'pointless',
'local',
'dust',
'partial',
'tunein',
'allen_enbot',
'7a730p',
'tangletalk',
'came',
'android',
'shantaeskyy',
'cainsuch',
'scum',
'portug',
'astroturf',
'pleasur',
'abstract',
'ethereal_7',
'asleep',
'program',
'solid',
'offers2go',
'rainbow',
'today',

'attempt',
'four',
'folieacat',
'conditionspriv',
'regc',
'gargron',
'attent',
'matt_bez',
'lownd',
'whoop',
'window',
'neal',
'ps',
'nobodi',
'blgrego',
'wooden',
'chriscesq',
'justthebottl',
'writes4coffe',
'sarumi',
'ak',
'basic',
'knight',
'brb',
'otherû_',
'oversight',
'chairman',
'tb',
'faster',
'luka',
'lich',
'i580',
'lume',
'oral',
'houston',
'halsey',
'briton',
'fox',
'149',
'x1411',
'mesick',
'icrc',
'vital',
'ordain',
'forestfir',
'robpulsenew',
'cam',

'order',
'algeria',
'soonersportstv',
'cream',
'hebrooon',
'cowboy',
'yall',
'fr',
'oklahomaok',
'languag',
'nickcannon',
'airwav',
'cabin',
'dorret',
'darkread',
'bushman',
'next',
'attach',
'imagini',
'terror',
'warû_',
'tomdean86',
'coreyash',
'pomo',
'cleveland',
'truly_st',
'natsnat',
'bounti',
'virg',
'initi',
'accus',
'hold',
'proud',
'district',
'hyatt',
'9newsmorn',
'murfreesboro',
'legendari',
'fennovoima',
'deputi',
'panti',
'pantherattack',
'wakho',
'japanes',
'via',
'fulfil',
'wasnû@t',

'spyro',
'minut',
'mili_5499',
'290',
'dyxtmrydu',
'thoroughli',
'nank',
'newyear',
'gridlock',
'adriasimon_',
'amo',
'cameronciletti',
'__',
'quiz',
'tunesblak',
'amazon',
'treescap',
'23800',
'coalit',
'tlk',
'hmu',
'deezercolombia',
'dominicana',
'okwx',
'bagsûª',
'canonusaimag',
'antiblight',
'specci',
'duo',
'ocean',
'111020',
'photographi',
'paper',
'waziristan',
'jake_adavi',
'freddi',
'betz',
's61231a',
'rrbc',
'voslooru',
'forbesasia',
'bhavana',
'incident',
'hayden',
'freeamirnow',
'doon',
'mwnhappi',

'wp',
'hitch',
'titl',
'retail',
'nypd',
'wit',
'sona',
'jose',
'magnitudeâââêml',
'60second',
'15000270364',
'renewsit',
'scofield',
'ralph',
'2pmclose',
'emili',
'lefti',
'read',
'left',
'newciti',
'tan',
'justin_l',
'unfair',
'rover',
'dalinthanelan',
'kunalkapoor',
'slosheriff',
'youngsaf',
'cliptv',
'cta',
'zar',
'captur',
'hagerstown',
'goonda',
'ensu',
'fkhanag',
'lucysforsal',
'showdown',
'necro',
'understandû',
'pageshi',
'lancast',
'savour',
'xma',
'restrict',
'medal',
'skim',

'generalnew',
'meinlcymb',
'ecofriendli',
'17jul2015',
'927',
'mo',
'nuclearweapon',
'yield',
'rcityporn',
'sx',
'recent',
'sequenc',
'ny1',
'jesuss',
'horsetalknz',
'nbcnightlynew',
'shouldnt',
'wart',
'nflexpertpick',
'unaddress',
'daborsch',
'elm',
'_charleyisqueen',
'fixingmatch',
'question',
'exacerb',
'politician',
'marijuana',
'bunt',
'sale',
'johnson',
'casualtiesû',
'newz_sacramento',
'rwrabbit',
'char',
'inst',
'humbl',
'trillion',
'startup',
'wale',
'death',
'ultimalucha',
'crazi',
'eo',
'spinningbot',
'histori',
'coif',

'oh',
'videoveranomtv',
'rub',
'space_wolverin',
'theconspiracytheori',
'bit',
'byr',
'thruuu',
'oomf',
'itsqueenbabi',
'kaboom',
'plymouth',
'24regionâ€œâ€œnear',
'raabchar_28',
'theyd',
'monarchi',
'plagu',
'uabstephenlong',
'sh',
'carmik',
'darken',
'fiascothat',
'roberthard',
'overblown',
'payment',
'hanna',
'i77',
'measur',
'battlefield',
'netlabel',
'pronounc',
'gostei',
'cabbag',
'fashion',
'subreddit',
'keithyyl',
'tune',
'snowstormdespit',
'aiiamericangiri',
'hockey',
'calgari',
'bluejay',
'stopevict',
'fredosantana300',
'ep18',
'rom',
'sweat',

'redempt',
'urin',
'attackonstil',
'upheav',
'paracord',
'saint',
'textur',
'peopl',
'slate',
'chilli',
'realjaxclon',
'8m',
'upgrad',
'kosher',
'fdx',
'whistl',
'shipsxanchor',
'clinic',
'locat',
'sergiopiaggio',
'litani',
'dragonval',
'sigh',
'marketingâ€mediocr',
'mapl',
'315',
'saison',
'prais',
'browni',
'karinagarciaxo_',
'üiwe',
'khq',
'doctor',
'shitton',
'vinustrip',
'taki',
'willienelson',
'544',
'willian',
'jackhol',
'submit',
'gooooooooo',
'fade',
'london',
'largest',
'music',
'drfriedencdc',

'humphrey',
'smile',
'florida',
'censorship',
'reportli',
'itz',
'eg',
'piperwearsthep',
'wasteland',
'holymileyray',
'diebold',
'mumbairiot9293',
'leedstraif',
'scott',
'ave',
'11yearold',
'main',
'wonderkid',
'realm',
'tix',
'brisban',
'foxdebatequest',
'xbox',
'condemn',
'indecis',
'outcri',
'crossbodi',
'fettilootch',
'remain',
'390',
'fell',
'activ',
'grate',
'angharadjames16',
'overflow',
'droid',
'southbound',
'perrybellegard',
'almusafirah_',
'mainten',
'scrape',
'carmi',
'evildead',
'gloria',
'sneak',
'tampabay',
'putin',

'sputter',
'probe',
'souda',
'retweet',
'q',
'mav',
'600û_',
'garza',
'kurdistan',
'malt',
'okla',
'peanut',
'keegan172',
'foster',
'airlift',
'3945',
'sittingû_',
'bradford',
'amcx',
'bbm',
'enhanc',
'wwa',
'browserâhijack',
'evad',
'2hi',
'slumber',
'calwildfir',
'corrupt',
'congest',
'wars_goddess',
'unknown',
'trailer',
'gonna',
'share',
'emmerdal',
'jacket',
'c130',
'197629',
'shuffl',
'usamisan',
'beacon',
'twcnewsclt',
'butter',
'lightseraph',
'cantwont',
'yup',
'lightningstrik',

'gain',
'oem',
'marley',
'racer',
'rspca',
'unplug',
'mfi',
'geoffrickli',
'shaw',
'drama',
'intactmh370partliftsoddsplaneglidednotcrashedintosea',
'dell',
'dick',
'deliv',
'balconi',
'quit',
'mlg',
'gt',
'laughtrad',
'bought',
'prodemocraci',
'search',
'mandem',
'012032',
'monsoon',
'mithitenni',
'shadowflam',
'battle_dom',
'newswatch',
'sun',
'wasp',
'tx',
'tc',
'cocki',
'score',
'descript',
'yiraneuni',
'prom',
'jean',
'fossil',
'aldrig',
'breakfaston',
'sikh',
'skyscrap',
'ny',
'wipp',
'treati',

'squirrel',
'filter',
'ramag',
'ensur',
'4422',
'uptotheminut',
'rochdal',
'nf',
'17',
'hostageamp2',
'libertarianluk',
'mac',
'1015',
'moscow',
'akcs1',
'testicl',
'rang',
'nasahurrican',
'anu',
'waimea',
'ps3',
'k_matako_bot',
'iit',
'thick',
'sham',
'jude_mugabi',
'alexeivolkov1',
'planetari',
'bethlehem',
'awn',
'2k15',
'lifethreaten',
'epilepticû_',
'crashingû_',
'candi',
'tunnel',
'hollywarnexx',
'starr',
'neg',
'hinterestin',
'midsumm',
'beaten',
'schedul',
'thrive',
'almost',
'vine',
'rapidli',

'segment',
'musik',
'3x5',
'happenû_',
'stefano',
'display',
'unsurprisingli',
'triciaoneillphoto',
'worn',
'4000',
'tepat',
'robertmeyer9',
'cartoonylik',
'method',
'downright',
'fact',
'goe',
'dix',
'seemeth',
'classic',
'cargo',
'defenc',
'forbath',
'know',
'standardis',
'averag',
'indiegogo',
'organicallyrud',
'slide',
'rorington95',
'governor',
'younoon',
'england',
'jasminehuerta24',
'homi',
'unavoid',
'comeeee',
'30û_',
'dantwitty52',
'mind',
'quicker',
'impress',
'cunt',
'hom',
'humid',
'symbol',
'horizon',

'eleph',
'guatemala',
'jen',
'fran',
'absurd',
'pine',
'unit',
'sleeper',
'warriorcord',
'setting4success',
'69',
'nitishkumar',
'savag',
'ûifdni',
'democracyûª',
'pod',
'pianohand',
'abcnews24',
'üïymcglau',
'answerû_',
'taxston',
'growingupincoloradoû',
'tracey',
'rar',
'ground',
'knox',
'gfe',
'kasabw',
'contig',
'durrellb',
'appar',
'bluetooth',
'tree',
'hornybigbadwolf',
'hungov',
'waterfur',
'landslid',
'floodzon',
'populardemand',
'citi',
'offtrackhors',
'bah',
'ps4',
'cast',
'samajp32',
'isla',
'tribal',

'sand',
'traci',
'lifelettercaf',
'countri',
'bro',
'coast2coastdj',
'sct012',
'mississauga',
'aboard',
'tyron',
'andâ€ˆchina',
'counti',
'badotweet',
'liqd',
'mediev',
'blink',
'subatom',
'imposs',
'thegam',
'tor',
'onlin',
'xd',
'rosemarytraval',
'thda',
'vitali',
'brzjc',
'prisonlik',
'novalismi',
'alba',
'carmel',
'workspac',
'crane',
'socialmediadriven',
'simply_vain',
'saniti',
'_dangerousbean',
'michele_nfpa',
'yoga',
'lhh',
'spsgsp',
'aladdin',
'110358',
'gargoyl',
'extern',
'tomato',
'diamorfiend',
'180',

'air',
'chapter',
'monday',
'collabor',
'uh',
'puls',
'indigo6',
'cstsmith',
'joshua',
'auction',
'colnharun',
'affili',
'difficultpeopl',
'sparkz',
'glaucoma',
'acid',
'tutor',
'jpanzar',
'journal',
'thatsabinegirl',
'kind',
'nennicook',
'wb',
'hazardddanger',
'disastereofici',
'faulti',
'intox',
'alifaditha',
'imagin',
'greyjoy',
'ameribag',
'bentley',
'democrat',
'satir',
'gosh',
'nasti',
'sing',
'kshllcenterpri1',
'templat',
'longll',
'adhd',
'weddinghour',
'cs6',
'jadehelm15',
'uriminzok',
'lethal',
'be4',

'jonvoyag',
'ihhen',
'switch',
'school',
'date',
'mackayim',
'influent',
'nno',
'predynast',
'dorismatsui',
'tow',
'pathfind',
'bid',
'careless',
'resolutevan',
'hashd',
'runaway',
'tr',
'savagen',
'kylewappl',
'ya_boi_luk',
'alameda',
'zaibatsunew',
'warcraft',
'sandun',
'hanna_brooksi',
'allinwithchri',
'demeanor',
'1038pm',
'pozarmi',
'taungbazar',
'lethbridg',
'fatherofthre',
'pile',
'item',
'630830',
'believei',
'dampen',
'harmoni',
'youngû_',
'purs',
'realmandyrain',
'wildlook',
'075',
'chocol',
'razak',
'heartless',


```

'privat',
'sirensicebreak',
'nigga',
'_oliviaann_',
'shooook',
'ashayo',
'muttatek',
'jeez',
'act',
'blizzard_draco',
'rdhorndal',
'1500',
'amazin',
'komatfans33',
'certifi',
'reebok',
'gotriangl',
'que',
'rescuer',
'trail',
'bicyclesuv',
'ford',
'125',
'aâemiddleag',
'bedroom',
'theash',
'smem',
'elliott',
'cochran',
'c720',
'unsign',
'burton',
'remedi',
'fat',
...]

```

```
[ ]: len(dictionary)
```

```
[ ]: 14969
```

6) Label Encode words in dictionary.

```
[ ]: word_set = {}
for i, w in enumerate(dictionary):
    word_set[w] = i
```

```
[ ]: word_set
```

```
[ ]: {'strang': 0,  
      'cannib': 1,  
      '662': 2,  
      'tent': 3,  
      'hair': 4,  
      'sin': 5,  
      'andrew': 6,  
      'show': 7,  
      'jitter': 8,  
      'benkin97': 9,  
      'edt': 10,  
      'barrack': 11,  
      'unwant': 12,  
      'institut': 13,  
      'chpsre': 14,  
      'salvag': 15,  
      'hashtag': 16,  
      'ng2x5': 17,  
      '7510880': 18,  
      '600': 19,  
      'shortfal': 20,  
      'panoram': 21,  
      'tragedi': 22,  
      'yyeso': 23,  
      'kendal': 24,  
      '21': 25,  
      'nightsad': 26,  
      'attila': 27,  
      'ransack': 28,  
      'snazzychipz': 29,  
      'flare': 30,  
      'morehouse64': 31,  
      'dispensari': 32,  
      'baduû_': 33,  
      'festac': 34,  
      'requiem': 35,  
      '0025': 36,  
      'charcoal': 37,  
      'real_liam_payn': 38,  
      'gr': 39,  
      'fleek': 40,  
      'tremblayeh': 41,  
      'appli': 42,  
      'fallen': 43,  
      'ear': 44,  
      'midweek': 45,  
      'oooh': 46,
```

'mk': 47,
'tmake': 48,
'camilacabello97': 49,
'tak': 50,
'revel': 51,
'182624': 52,
'barrier': 53,
'teeth': 54,
'about': 55,
'slam': 56,
'noodl': 57,
'üihatchet': 58,
'surround': 59,
'khuzdar': 60,
'prefer': 61,
'envis': 62,
'tech': 63,
'stallion150': 64,
'sto': 65,
'worth': 66,
'sirtophamhat': 67,
'2015': 68,
'milita': 69,
'ganseyman': 70,
'9km': 71,
'emv': 72,
'sponsor': 73,
'fang': 74,
'cascad': 75,
'fullscreen': 76,
'nuclear': 77,
'afterhaiyan': 78,
'meelllitt': 79,
'cachet': 80,
'forney': 81,
'teamatowinn': 82,
'randomthought': 83,
'ted': 84,
'stat': 85,
'packag': 86,
'hvnewsnetwork': 87,
'crimin': 88,
'wmata': 89,
'mmda': 90,
'indiefilm': 91,
'helo': 92,
'7speed': 93,

'banquet': 94,
'sink': 95,
'ld': 96,
'massag': 97,
'cutter': 98,
'hinata': 99,
'jones94kyl': 100,
'faro': 101,
'jack': 102,
'jersey': 103,
'grief': 104,
'msnbc': 105,
'nh1new': 106,
'troop': 107,
'bangladesh': 108,
'alexandrapullin': 109,
'bieber': 110,
'groom': 111,
'conclus': 112,
'squeaver': 113,
'loui': 114,
'omw': 115,
'court': 116,
'seismicresist': 117,
'heyimbeeyt': 118,
'chuckswindol': 119,
'otherwisethey': 120,
'eurocrisi': 121,
'lolol': 122,
'gloriavelez': 123,
'regim': 124,
'sheûª': 125,
'greek': 126,
'desper': 127,
'ûhoax': 128,
'ika': 129,
'ici': 130,
'christianvalu': 131,
'maxi': 132,
'stopû': 133,
'madara': 134,
'4tp_kotc': 135,
'optimist': 136,
'snowbal': 137,
'bomber': 138,
'competitor': 139,
'unharm': 140,

'7k': 141,
'pkk': 142,
'coleslaw': 143,
'fuel': 144,
'jenner': 145,
'teeû_': 146,
'area': 147,
'queen': 148,
'4193': 149,
'teddi': 150,
'baruch': 151,
'gotthard': 152,
'kiss': 153,
'pak': 154,
'bernardino': 155,
'righteou': 156,
'gmcr': 157,
'deb117': 158,
'book': 159,
'njenga': 160,
'abandonedp': 161,
'profittothepeopl': 162,
'blinker': 163,
'fli': 164,
'rd': 165,
'kenji': 166,
'fleet': 167,
'rigour': 168,
'raft': 169,
'nasasolarsystem': 170,
'18th': 171,
'yyj': 172,
'striker': 173,
'around': 174,
'devalu': 175,
'la_siren': 176,
'redeem': 177,
'ml': 178,
'oc': 179,
'nake': 180,
'vivian': 181,
'camilla_33': 182,
'lovefood': 183,
'moan': 184,
'human': 185,
'irand': 186,
'atleast': 187,

'emot': 188,
'jamienny': 189,
'pointless': 190,
'local': 191,
'dust': 192,
'partial': 193,
'tunein': 194,
'allen_enbot': 195,
'7a730p': 196,
'tangletalk': 197,
'came': 198,
'android': 199,
'shantaesky': 200,
'cainsuch': 201,
'scum': 202,
'portug': 203,
'astroturf': 204,
'pleasur': 205,
'abstract': 206,
'ethereal_7': 207,
'asleep': 208,
'program': 209,
'solid': 210,
'offers2go': 211,
'rainbow': 212,
'today': 213,
'attempt': 214,
'four': 215,
'folieacat': 216,
'conditionspriv': 217,
'regc': 218,
'gargron': 219,
'attent': 220,
'matt_bez': 221,
'lownd': 222,
'whoop': 223,
'window': 224,
'neal': 225,
'ps': 226,
'nobodi': 227,
'bigregio': 228,
'wooden': 229,
'chriscesq': 230,
'justthebottl': 231,
'writes4coffe': 232,
'sarumi': 233,
'ak': 234,

'basic': 235,
'knight': 236,
'brb': 237,
'otherû_': 238,
'oversight': 239,
'chairman': 240,
'tb': 241,
'faster': 242,
'luka': 243,
'lich': 244,
'i580': 245,
'lume': 246,
'oral': 247,
'houston': 248,
'halsey': 249,
'briton': 250,
'fox': 251,
'149': 252,
'x1411': 253,
'mesick': 254,
'icrc': 255,
'vital': 256,
'ordain': 257,
'forestfir': 258,
'robpulsenew': 259,
'cam': 260,
'order': 261,
'algeria': 262,
'soonersportstv': 263,
'cream': 264,
'hebrooon': 265,
'cowboy': 266,
'yall': 267,
'fr': 268,
'oklahomaok': 269,
'languag': 270,
'nickcannon': 271,
'airwav': 272,
'cabin': 273,
'dorret': 274,
'darkread': 275,
'bushman': 276,
'next': 277,
'attach': 278,
'imagini': 279,
'terror': 280,
'warû_': 281,

'tomdean86': 282,
'coreyash': 283,
'pomo': 284,
'cleveland': 285,
'truly_st': 286,
'natsnat': 287,
'bounti': 288,
'virg': 289,
'initi': 290,
'accus': 291,
'hold': 292,
'proud': 293,
'district': 294,
'hyatt': 295,
'9newsmorn': 296,
'murfreesboro': 297,
'legendari': 298,
'fennovoima': 299,
'deputi': 300,
'panti': 301,
'pantherattack': 302,
'wakho': 303,
'japanes': 304,
'via': 305,
'fulfil': 306,
'wasnû^at': 307,
'spyro': 308,
'minut': 309,
'mili_5499': 310,
'290': 311,
'dyxtmrydu': 312,
'thoroughli': 313,
'nank': 314,
'newyear': 315,
'gridlock': 316,
'adriasimon_': 317,
'amo': 318,
'cameronciletti': 319,
'__': 320,
'quiz': 321,
'tunesblak': 322,
'amazon': 323,
'treescap': 324,
'23800': 325,
'coalit': 326,
'tlk': 327,
'hmu': 328,

'deezercolumbia': 329,
'dominicana': 330,
'okwx': 331,
'bagsû': 332,
'canonusaimag': 333,
'antiblight': 334,
'specci': 335,
'duo': 336,
'ocean': 337,
'111020': 338,
'photographi': 339,
'paper': 340,
'waziristan': 341,
'jake_adavi': 342,
'freddi': 343,
'betz': 344,
's61231a': 345,
'rrbc': 346,
'voslooru': 347,
'forbesasia': 348,
'bhavana': 349,
'incident': 350,
'hayden': 351,
'freeamirnow': 352,
'doon': 353,
'mwnhappi': 354,
'wp': 355,
'hitch': 356,
'titl': 357,
'retail': 358,
'nypd': 359,
'wit': 360,
'sona': 361,
'jose': 362,
'magnitudeâââêml': 363,
'60second': 364,
'15000270364': 365,
'renewsit': 366,
'scofield': 367,
'ralph': 368,
'2pmclose': 369,
'emili': 370,
'lefti': 371,
'read': 372,
'left': 373,
'newciti': 374,
'tan': 375,

'justin_l': 376,
'unfair': 377,
'rover': 378,
'dalinthanelan': 379,
'kunalkapoor': 380,
'slosheriff': 381,
'youngsaf': 382,
'cliptv': 383,
'cta': 384,
'zar': 385,
'captur': 386,
'hagerstown': 387,
'goonda': 388,
'ensu': 389,
'fkhanag': 390,
'lucysforsal': 391,
'showdown': 392,
'necro': 393,
'understandû': 394,
'pageshi': 395,
'lancast': 396,
'savour': 397,
'xma': 398,
'restrict': 399,
'medal': 400,
'skim': 401,
'generalnew': 402,
'meinlcymb': 403,
'ecofriendli': 404,
'17jul2015': 405,
'927': 406,
'mo': 407,
'nuclearweapon': 408,
'yield': 409,
'rcityporn': 410,
'sx': 411,
'recent': 412,
'sequenc': 413,
'ny1': 414,
'jesuss': 415,
'horsetalknz': 416,
'nbcnightlynew': 417,
'shouldnt': 418,
'wart': 419,
'nflexpertpick': 420,
'unaddress': 421,
'daborsch': 422,

'elm': 423,
'_charleyisqueen': 424,
'fixingmatch': 425,
'question': 426,
'exacerb': 427,
'politician': 428,
'marijuana': 429,
'bunt': 430,
'sale': 431,
'johnson': 432,
'casualties': 433,
'newz_sacramento': 434,
'rwrabbit': 435,
'char': 436,
'inst': 437,
'humbl': 438,
'trillion': 439,
'startup': 440,
'wale': 441,
'death': 442,
'ultimalucha': 443,
'crazi': 444,
'eo': 445,
'spinningbot': 446,
'histori': 447,
'coif': 448,
'oh': 449,
'videoveranomtv': 450,
'rub': 451,
'space_wolverin': 452,
'theconspiracytheori': 453,
'bit': 454,
'byr': 455,
'thruuu': 456,
'oomf': 457,
'itsqueenbabi': 458,
'kaboom': 459,
'plymouth': 460,
'24regionââânear': 461,
'raabchar_28': 462,
'theyd': 463,
'monarchi': 464,
'plagu': 465,
'uabstephenlong': 466,
'sh': 467,
'carmik': 468,
'darken': 469,

'fiascothat': 470,
'roberthard': 471,
'overblown': 472,
'payment': 473,
'hanna': 474,
'i77': 475,
'measur': 476,
'battlefield': 477,
'netlabel': 478,
'pronounc': 479,
'gostei': 480,
'cabbag': 481,
'fashion': 482,
'subreddit': 483,
'keithyyl': 484,
'tune': 485,
'snowstormdespit': 486,
'aiiamericangiri': 487,
'hockey': 488,
'calgari': 489,
'bluejay': 490,
'stopevict': 491,
'fredosantana300': 492,
'ep18': 493,
'rom': 494,
'sweat': 495,
'redempt': 496,
'urin': 497,
'attackonstil': 498,
'upheav': 499,
'paracord': 500,
'saint': 501,
'textur': 502,
'peopl': 503,
'slate': 504,
'chilli': 505,
'realjaxclon': 506,
'8m': 507,
'upgrad': 508,
'kosher': 509,
'fdx': 510,
'whistl': 511,
'shipsxanchor': 512,
'clinic': 513,
'locat': 514,
'sergiopiaggio': 515,
'litani': 516,

'dragonval': 517,
'sigh': 518,
'marketingãmediocr': 519,
'mapl': 520,
'315': 521,
'saison': 522,
'prais': 523,
'browni': 524,
'karinagarciaxo_': 525,
'ûiwe': 526,
'khq': 527,
'doctor': 528,
'shitton': 529,
'vinustrip': 530,
'taki': 531,
'willienelson': 532,
'544': 533,
'willian': 534,
'jackhol': 535,
'submit': 536,
'gooooooooo': 537,
'fade': 538,
'london': 539,
'largest': 540,
'music': 541,
'drfriedencdc': 542,
'humphrey': 543,
'smile': 544,
'florida': 545,
'censorship': 546,
'reportli': 547,
'itz': 548,
'eg': 549,
'piperwearsthep': 550,
'wasteland': 551,
'holymileyray': 552,
'diebold': 553,
'mumbairiot9293': 554,
'leedstraif': 555,
'scott': 556,
'ave': 557,
'11yearold': 558,
'main': 559,
'wonderkid': 560,
'realm': 561,
'tix': 562,
'brisban': 563,

'foxdebatequest': 564,
'xbox': 565,
'condemn': 566,
'indecis': 567,
'outcri': 568,
'crossbodi': 569,
'fettilootch': 570,
'remain': 571,
'390': 572,
'fell': 573,
'activ': 574,
'grate': 575,
'angharadjames16': 576,
'overflow': 577,
'droid': 578,
'southbound': 579,
'perrybellegard': 580,
'almusafirah_': 581,
'mainten': 582,
'scrape': 583,
'carmi': 584,
'evildead': 585,
'gloria': 586,
'sneak': 587,
'tampabay': 588,
'putin': 589,
'sputter': 590,
'probe': 591,
'souda': 592,
'retweet': 593,
'q': 594,
'mav': 595,
'600û_': 596,
'garza': 597,
'kurdistan': 598,
'malt': 599,
'okla': 600,
'peanut': 601,
'keegan172': 602,
'foster': 603,
'airlift': 604,
'3945': 605,
'sittingû_': 606,
'bradford': 607,
'amcx': 608,
'bbm': 609,
'enhanc': 610,

'wwa': 611,
'browserâ€”hijack': 612,
'evad': 613,
'2hi': 614,
'slumber': 615,
'calwildfir': 616,
'corrupt': 617,
'congest': 618,
'wars_goddess': 619,
'unknown': 620,
'trailer': 621,
'gonna': 622,
'share': 623,
'emmerdal': 624,
'jacket': 625,
'c130': 626,
'197629': 627,
'shuffl': 628,
'usamisan': 629,
'beacon': 630,
'twcnewsclt': 631,
'butter': 632,
'lightseraph': 633,
'cantwont': 634,
'yup': 635,
'lightningstrik': 636,
'gain': 637,
'oem': 638,
'marley': 639,
'racer': 640,
'rspca': 641,
'unplug': 642,
'mfi': 643,
'geoffrickli': 644,
'shaw': 645,
'drama': 646,
'intactmh370partliftsoddsplaneglidednotcrashedintosea': 647,
'dell': 648,
'dick': 649,
'deliv': 650,
'balconi': 651,
'quit': 652,
'mlg': 653,
'gt': 654,
'laughtrad': 655,
'bought': 656,
'prodemocraci': 657,

'search': 658,
'mandem': 659,
'012032': 660,
'monsoon': 661,
'mithitenni': 662,
'shadowflam': 663,
'battle_dom': 664,
'newswatch': 665,
'sun': 666,
'wasp': 667,
'tx': 668,
'tc': 669,
'cocki': 670,
'score': 671,
'descript': 672,
'yiraneuni': 673,
'prom': 674,
'jean': 675,
'fossil': 676,
'alldridg': 677,
'breakfaston': 678,
'sikh': 679,
'skyscrap': 680,
'ny': 681,
'wipp': 682,
'treati': 683,
'squirrel': 684,
'filter': 685,
'ramag': 686,
'ensur': 687,
'4422': 688,
'uptotheminut': 689,
'rochdal': 690,
'nf': 691,
'17': 692,
'hostageamp2': 693,
'libertarianluk': 694,
'mac': 695,
'1015': 696,
'moscow': 697,
'akcsl': 698,
'testicl': 699,
'rang': 700,
'nasahurricane': 701,
'anu': 702,
'waimea': 703,
'ps3': 704,

'k_matako_bot': 705,
'iit': 706,
'thick': 707,
'sham': 708,
'jude_mugabi': 709,
'alexeivolkov1': 710,
'planetari': 711,
'bethlehem': 712,
'awn': 713,
'2k15': 714,
'lifethreaten': 715,
'epilepticû_': 716,
'crashingû_': 717,
'candi': 718,
'tunnel': 719,
'hollywarnexx': 720,
'starr': 721,
'neg': 722,
'hinterestin': 723,
'midsumm': 724,
'beaten': 725,
'schedul': 726,
'thrive': 727,
'almost': 728,
'vine': 729,
'rapidli': 730,
'segment': 731,
'musik': 732,
'3x5': 733,
'happenû_': 734,
'stefano': 735,
'display': 736,
'unsurprisingli': 737,
'triciaoneillphoto': 738,
'worn': 739,
'4000': 740,
'tepat': 741,
'robertmeyer9': 742,
'cartoonylik': 743,
'method': 744,
'downright': 745,
'fact': 746,
'goe': 747,
'dix': 748,
'seemeth': 749,
'classic': 750,
'cargo': 751,

'defenc': 752,
'forbath': 753,
'know': 754,
'standardis': 755,
'averag': 756,
'indiegogo': 757,
'organicallyrud': 758,
'slide': 759,
'rorington95': 760,
'governor': 761,
'younoon': 762,
'england': 763,
'jasminehuerta24': 764,
'homi': 765,
'unavoid': 766,
'comeeee': 767,
'30û_': 768,
'dantwitty52': 769,
'mind': 770,
'quicker': 771,
'impress': 772,
'cunt': 773,
'hom': 774,
'humid': 775,
'symbol': 776,
'horizon': 777,
'eleph': 778,
'guatemala': 779,
'jen': 780,
'fran': 781,
'absurd': 782,
'pine': 783,
'unit': 784,
'sleeper': 785,
'warriorcord': 786,
'setting4success': 787,
'69': 788,
'nitishkumar': 789,
'savag': 790,
'ûifdni': 791,
'democracyûª': 792,
'pod': 793,
'pianohand': 794,
'abcnews24': 795,
'ûiymcglaun': 796,
'answerû_': 797,
'taxston': 798,

'growingupincolorado': 799,
'tracey': 800,
'rar': 801,
'ground': 802,
'knox': 803,
'gfe': 804,
'kasabw': 805,
'contig': 806,
'durrellb': 807,
'appar': 808,
'bluetooth': 809,
'tree': 810,
'hornybigbadwolf': 811,
'hungov': 812,
'waterfur': 813,
'landslid': 814,
'floodzon': 815,
'populardemand': 816,
'citi': 817,
'offtrackhors': 818,
'bah': 819,
'ps4': 820,
'cast': 821,
'samajp32': 822,
'isla': 823,
'tribal': 824,
'sand': 825,
'traci': 826,
'lifelettercaf': 827,
'countri': 828,
'bro': 829,
'coast2coastdj': 830,
'sct012': 831,
'mississauga': 832,
'aboard': 833,
'tyron': 834,
'andâchina': 835,
'counti': 836,
'badotweet': 837,
'liqd': 838,
'mediev': 839,
'blink': 840,
'subatom': 841,
'imposs': 842,
'thegam': 843,
'tor': 844,
'onlin': 845,

'xd': 846,
'rosemarytraval': 847,
'thda': 848,
'vitali': 849,
'brzjc': 850,
'prisonlik': 851,
'novalismi': 852,
'alba': 853,
'carmel': 854,
'workspac': 855,
'crane': 856,
'socialmediadriven': 857,
'simply_vain': 858,
'saniti': 859,
'_dangerousbean': 860,
'michele_nfpa': 861,
'yoga': 862,
'lhh': 863,
'spsgsp': 864,
'aladdin': 865,
'110358': 866,
'gargoyl': 867,
'extern': 868,
'tomato': 869,
'diamorfiend': 870,
'180': 871,
'air': 872,
'chapter': 873,
'monday': 874,
'collabor': 875,
'uh': 876,
'puls': 877,
'indigo6': 878,
'cstsmith': 879,
'joshua': 880,
'auction': 881,
'colnharun': 882,
'affili': 883,
'difficultpeopl': 884,
'sparkz': 885,
'glaucoma': 886,
'acid': 887,
'tutor': 888,
'jpanzar': 889,
'journal': 890,
'thatsabinegirl': 891,
'kind': 892,

'nennicook': 893,
'wb': 894,
'hazarddanger': 895,
'disastereofici': 896,
'faulti': 897,
'intox': 898,
'alifaditha': 899,
'imagin': 900,
'greyjoy': 901,
'ameribag': 902,
'bentley': 903,
'democrat': 904,
'satir': 905,
'gosh': 906,
'nasti': 907,
'sing': 908,
'kshllcenterprii1': 909,
'templat': 910,
'longll': 911,
'adhd': 912,
'weddinghour': 913,
'cs6': 914,
'jadehelm15': 915,
'uriminzok': 916,
'lethal': 917,
'be4': 918,
'jonvoyag': 919,
'ihhen': 920,
'switch': 921,
'school': 922,
'date': 923,
'mackayim': 924,
'influenc': 925,
'nno': 926,
'predynast': 927,
'dorismatsui': 928,
'tow': 929,
'pathfind': 930,
'bid': 931,
'careless': 932,
'resolutevan': 933,
'hashd': 934,
'runaway': 935,
'tr': 936,
'savagen': 937,
'kylewappl': 938,
'ya_boi_luk': 939,

'alameda': 940,
'zaibatsunew': 941,
'warcraft': 942,
'sandun': 943,
'hanna_brooksi': 944,
'allinwithchri': 945,
'demeanor': 946,
'1038pm': 947,
'pozarmi': 948,
'taungbazar': 949,
'lethbridg': 950,
'fatherofthre': 951,
'pile': 952,
'item': 953,
'630830': 954,
'believei': 955,
'dampen': 956,
'harmoni': 957,
'youngû_': 958,
'purs': 959,
'realmandyrain': 960,
'wildlook': 961,
'075': 962,
'chocol': 963,
'razak': 964,
'heartless': 965,
'privat': 966,
'sirensicebreak': 967,
'nigga': 968,
'_oliviaann_': 969,
'shook': 970,
'ashayo': 971,
'muttatek': 972,
'jeez': 973,
'act': 974,
'blizzard_draco': 975,
'rdhorndal': 976,
'1500': 977,
'amazon': 978,
'kombatfans33': 979,
'certifi': 980,
'reebok': 981,
'gottriangl': 982,
'que': 983,
'rescuer': 984,
'trail': 985,
'bicyclesuv': 986,

```
'ford': 987,  
'125': 988,  
'aa@middleag': 989,  
'bedroom': 990,  
'theash': 991,  
'smem': 992,  
'elliott': 993,  
'cochran': 994,  
'c720': 995,  
'unsign': 996,  
'burton': 997,  
'remedi': 998,  
'fat': 999,  
...}
```