In [1]:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from subprocess import check_output
%matplotlib inline
import plotly.offline as py
py.init_notebook_mode(connected=True)
import plotly.graph_objs as go
import plotly.tools as tls
import os
import gc

import re
from nltk.corpus import stopwords
import distance
from nltk.stem import PorterStemmer
from bs4 import BeautifulSoup
```

In [2]:

```
df = pd.read_csv('Quora_train.csv')
print("Number of Data point:",df.shape[0])
```

```
Number of Data point: 404290
```

In [3]:

```
df.head()
```

Out[3]:

	id	qid1	qid2	question1	question2	is_duplicate
0	0	1	2	What is the step by step guide to invest in sh...	What is the step by step guide to invest in sh...	0
1	1	3	4	What is the story of Kohinoor (Koh-i-Noor) Dia...	What would happen if the Indian government sto...	0
2	2	5	6	How can I increase the speed of my internet co...	How can Internet speed be increased by hacking...	0
3	3	7	8	Why am I mentally very lonely? How can I solve...	Find the remainder when [math]23^{24}[/math] i...	0
4	4	9	10	Which one dissolve in water quikly sugar, salt...	Which fish would survive in salt water?	0

In [4]:

```
df.info()
```

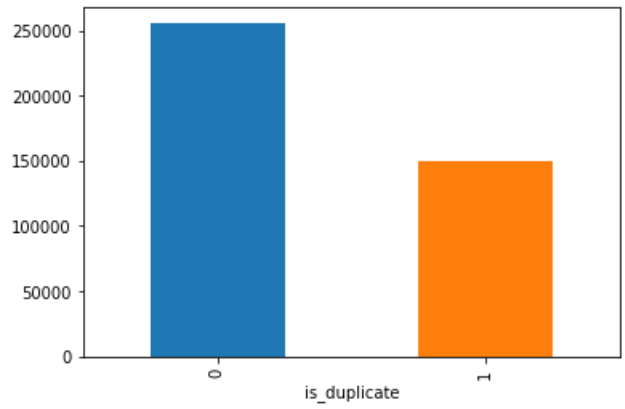
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 404290 entries, 0 to 404289
Data columns (total 6 columns):
id              404290 non-null int64
qid1            404290 non-null int64
qid2            404290 non-null int64
question1       404289 non-null object
question2       404288 non-null object
is_duplicate    404290 non-null int64
dtypes: int64(4), object(2)
memory usage: 18.5+ MB
```

In [5]:

```
df.groupby("is_duplicate")['id'].count().plot.bar()
```

Out[5]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x226881d80f0>
```



In [6]:

```
print("Total No. of QUESTION for pairing: ",len(df))
```

```
Total No. of QUESTION for pairing:  404290
```

In [7]:

```
print("QUESTION PAIRS ARE NOT SIMILAR",100-round(df['is_duplicate'].mean()*100,2))
print("QUESTION PAIRS ARE SIMILAR",round(df['is_duplicate'].mean()*100,2))
```

```
QUESTION PAIRS ARE NOT SIMILAR 63.08
QUESTION PAIRS ARE SIMILAR 36.92
```

In [8]:

```
qids = pd.Series(df['qid1'].tolist()+df['qid2'].tolist())
```

In [9]:

```
len(qids)
```

Out[9]:

```
808580
```

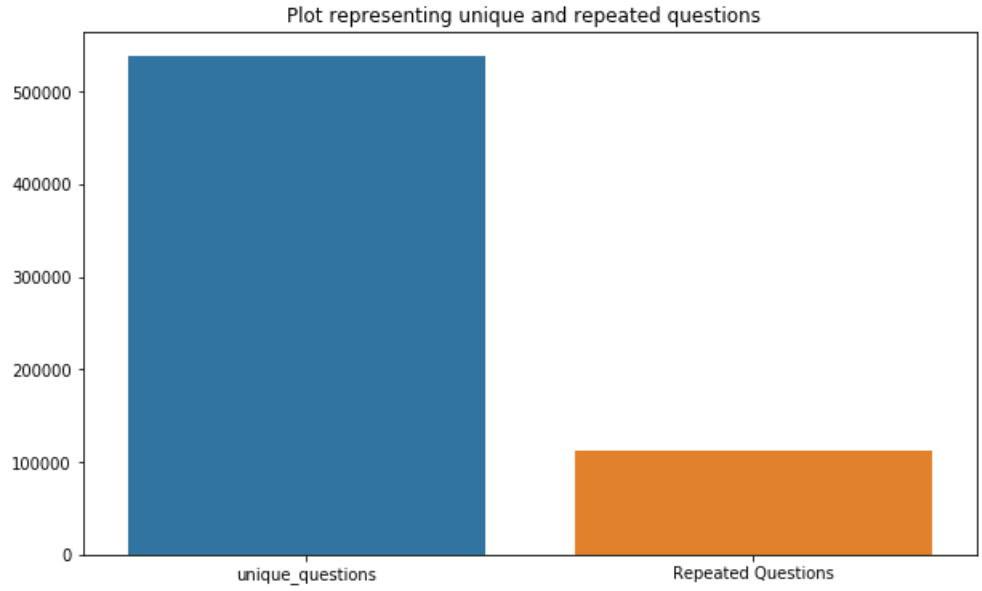
In [10]:

```
unique_qids = len(np.unique(qids))
qs_morethan_onetime =  np.sum(qids.value_counts()>1)
print("Total number of Unique Questions are",unique_qids)
print ('Number of unique questions that appear more than one time: {}
({}%)\n'.format(qs_morethan_onetime,qs_morethan_onetime/unique_qids*100))

print ('Max number of times a single question is repeated: {}\n'.format(max(qids.value_counts())))


q_vals=qids.value_counts()

q_vals=q_vals.values
```

```
Total number of Unique Questions are 537933
Number of unique questions that appear more than one time: 111780 (20.77953945937505%)

Max number of times a single question is repeated: 157
```

In [11]:

```
x = ["unique_questions" , "Repeated Questions"]
y =  [unique_qids , qs_morethan_onetime]

plt.figure(figsize=(10, 6))
plt.title ("Plot representing unique and repeated questions  ")
sns.barplot(x,y)
plt.show()
```



In [12]:

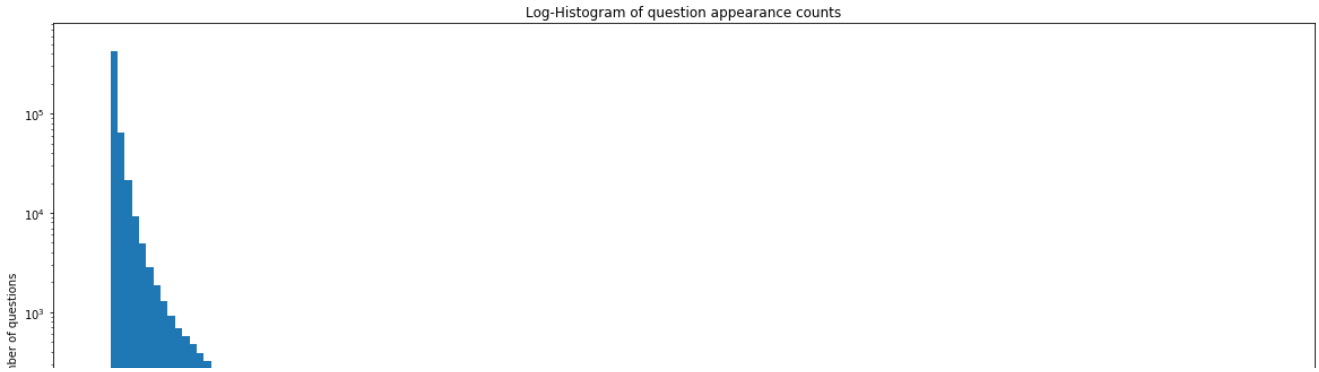
```
pair_duplicates =
df[['qid1','qid2','is_duplicate']].groupby(['qid1','qid2']).count().reset_index()
print ("Number of duplicate questions",(pair_duplicates).shape[0] - df.shape[0])
```

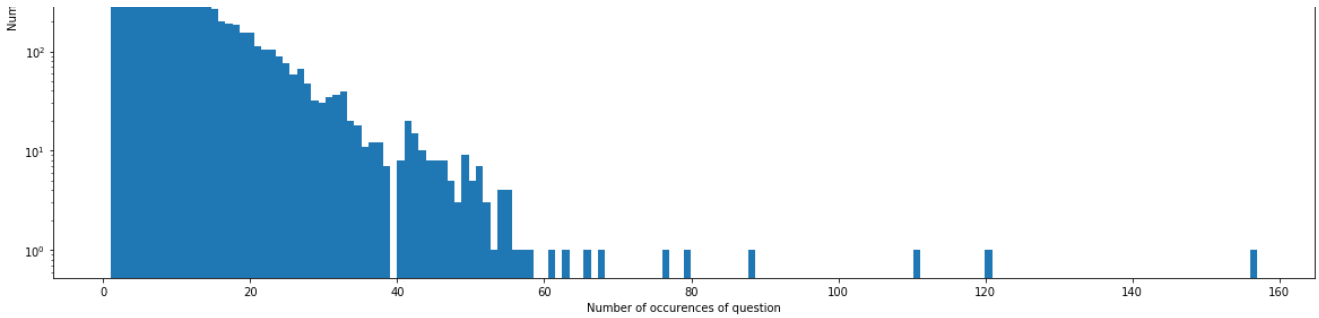
```
Number of duplicate questions 0
```

In [13]:

```
plt.figure(figsize=(20,10))
plt.hist(qids.value_counts(),bins = 160)
plt.yscale('log', nonposy='clip')

plt.title('Log-Histogram of question appearance counts')

plt.xlabel('Number of occurences of question')

plt.ylabel('Number of questions')

print ('Maximum number of times a single question is repeated: {}\n'.format(max(qids.value_counts(
)))
```

```
Maximum number of times a single question is repeated: 157
```





In [14]:

```
nan_rows = df[df.isnull().any(1)]
print(nan_rows)
```

```
            id    qid1    qid2                          question1  \
105780  105780  174363  174364    How can I develop android app?
201841  201841  303951  174364  How can I create an Android app?
363362  363362  493340  493341                               NaN

                                             question2  is_duplicate
105780                                             NaN             0
201841                                             NaN             0
363362  My Chinese name is Haichao Yu. What English na...             0
```

In [15]:

```
df = df.fillna('')
nan_rows = df[df.isnull().any(1)]
print(nan_rows)
```

```
Empty DataFrame
Columns: [id, qid1, qid2, question1, question2, is_duplicate]
Index: []
```

Let us now construct a few features like:

- **freq_qid1** = Frequency of qid1's
- **freq_qid2** = Frequency of qid2's
- **q1len** = Length of q1
- **q2len** = Length of q2
- **q1_n_words** = Number of words in Question 1
- **q2_n_words** = Number of words in Question 2
- **word_Common** = (Number of common unique words in Question 1 and Question 2)
- **word_Total** =(Total num of words in Question 1 + Total num of words in Question 2)
- **word_share** = (word_common)/(word_Total)
- **freq_q1+freq_q2** = sum total of frequency of qid1 and qid2
- **freq_q1-freq_q2** = absolute difference of frequency of qid1 and qid2

In [16]:

```
if os.path.isfile('df_fe_without_preprocessing_train.csv'):
    df = pd.read_csv("df_fe_without_preprocessing_train.csv",encoding='latin-1')
else:
    df['freq_qid1'] = df.groupby('qid1')['qid1'].transform('count')
    df['freq_qid2'] = df.groupby('qid2')['qid2'].transform('count')
    df['q1len'] = df['question1'].str.len()
    df['q2len'] = df['question2'].str.len()
    df['q1_n_words'] = df['question1'].apply(lambda row: len(row.split(" ")))
    df['q2_n_words'] = df['question2'].apply(lambda row: len(row.split(" ")))

    def normalized_word_Common(row):
        w1 = set(map(lambda word: word.lower().strip(), row['question1'].split(" ")))
        w2 = set(map(lambda word: word.lower().strip(), row['question2'].split(" ")))
        return 1.0 * len(w1 & w2)
    df['word_Common'] = df.apply(normalized_word_Common, axis=1)

    def normalized_word_Total(row):
```

```
        w1 = set(map(lambda word: word.lower().strip(), row['question1'].split(" ")))
        w2 = set(map(lambda word: word.lower().strip(), row['question2'].split(" ")))
        return 1.0 * (len(w1) + len(w2))
    df['word_Total'] = df.apply(normalized_word_Total, axis=1)

    def normalized_word_share(row):
        w1 = set(map(lambda word: word.lower().strip(), row['question1'].split(" ")))
        w2 = set(map(lambda word: word.lower().strip(), row['question2'].split(" ")))
        return 1.0 * len(w1 & w2)/(len(w1) + len(w2))
    df['word_share'] = df.apply(normalized_word_share, axis=1)

    df['freq_q1+q2'] = df['freq_qid1']+df['freq_qid2']
    df['freq_q1-q2'] = abs(df['freq_qid1']-df['freq_qid2'])

    df.to_csv("df_fe_without_preprocessing_train.csv", index=False)

df.head()
```

Out[16]:

	id	qid1	qid2	question1	question2	is_duplicate	freq_id1	freq_id2	q1len	q2len	q1_n_words	q2_n_words	Word_c
0	0	1	2	What is the step by step guide to invest in sh...	What is the step by step guide to invest in sh...	0	1	1	66	57	14	12	10.0
1	1	3	4	What is the story of Kohinoor (Koh-i-Noor) Dia...	What would happen if the Indian government sto...	0	4	1	51	88	8	13	4.0
2	2	5	6	How can I increase the speed of my internet co...	How can Internet speed be increased by hacking...	0	1	1	73	59	14	10	4.0
3	3	7	8	Why am I mentally very lonely? How can I solve...	Find the remainder when [math]23^{24}[/math] i...	0	1	1	50	65	11	9	0.0
4	4	9	10	Which one dissolve in water quikly sugar, salt...	Which fish would survive in salt water?	0	3	1	76	39	13	7	2.0

In [20]:

```
import os
```

In [17]:

```
print ("Minimum length of the questions in question1 : " , min(df['q1_n_words']))

print ("Minimum length of the questions in question2 : " , min(df['q2_n_words']))

print ("Number of Questions with minimum length [question1] :", df[df['q1_n_words']== 1].shape[0])
```

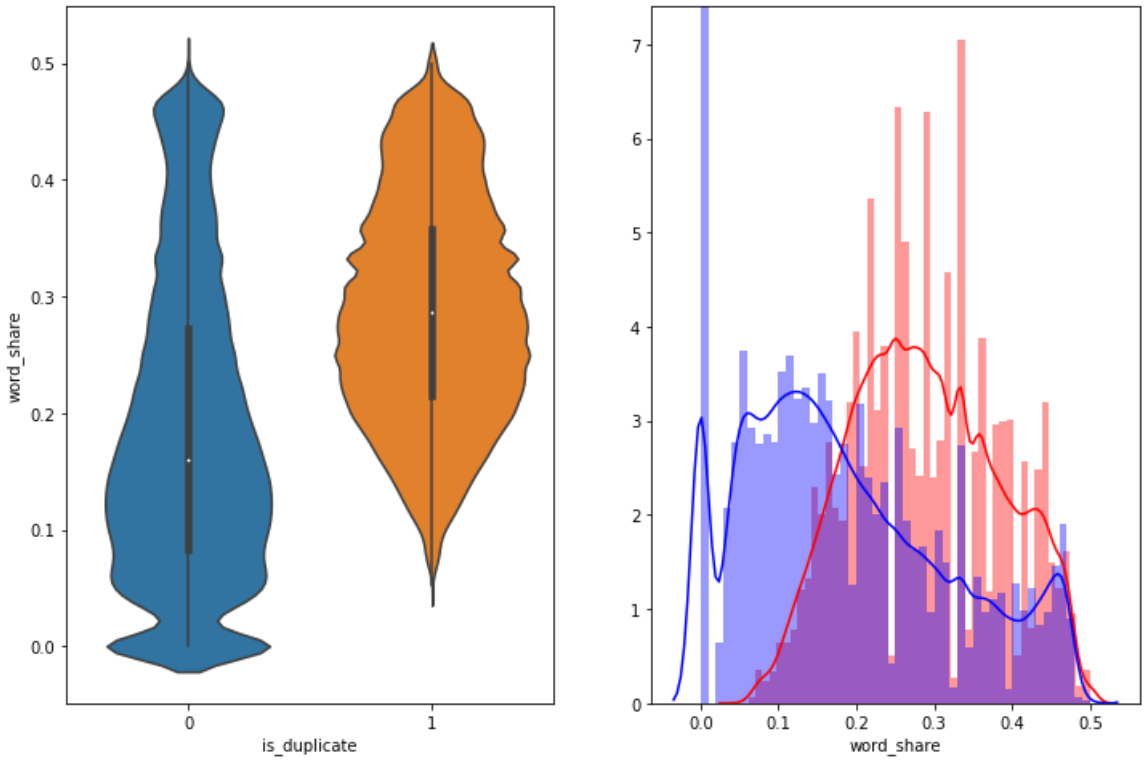
```
print ("Number of Questions with minimum length [question2] :", df[df['q2_n_words']== 1].shape[0])
```

```
Minimum length of the questions in question1 :  1
Minimum length of the questions in question2 :  1
Number of Questions with minimum length [question1] : 67
Number of Questions with minimum length [question2] : 24
```

In [18]:

```
plt.figure(figsize=(12, 8))

plt.subplot(1,2,1)
sns.violinplot(x = 'is_duplicate', y = 'word_share', data = df[0:])

plt.subplot(1,2,2)
sns.distplot(df[df['is_duplicate'] == 1.0]['word_share'][0:] , label = "1", color = 'red')
sns.distplot(df[df['is_duplicate'] == 0.0]['word_share'][0:] , label = "0" , color = 'blue' )
plt.show()
```

```
C:\Users\LENOVO\Anaconda3\lib\site-packages\matplotlib\axes\_axes.py:6462: UserWarning:

The 'normed' kwarg is deprecated, and has been replaced by the 'density' kwarg.

C:\Users\LENOVO\Anaconda3\lib\site-packages\matplotlib\axes\_axes.py:6462: UserWarning:

The 'normed' kwarg is deprecated, and has been replaced by the 'density' kwarg.
```

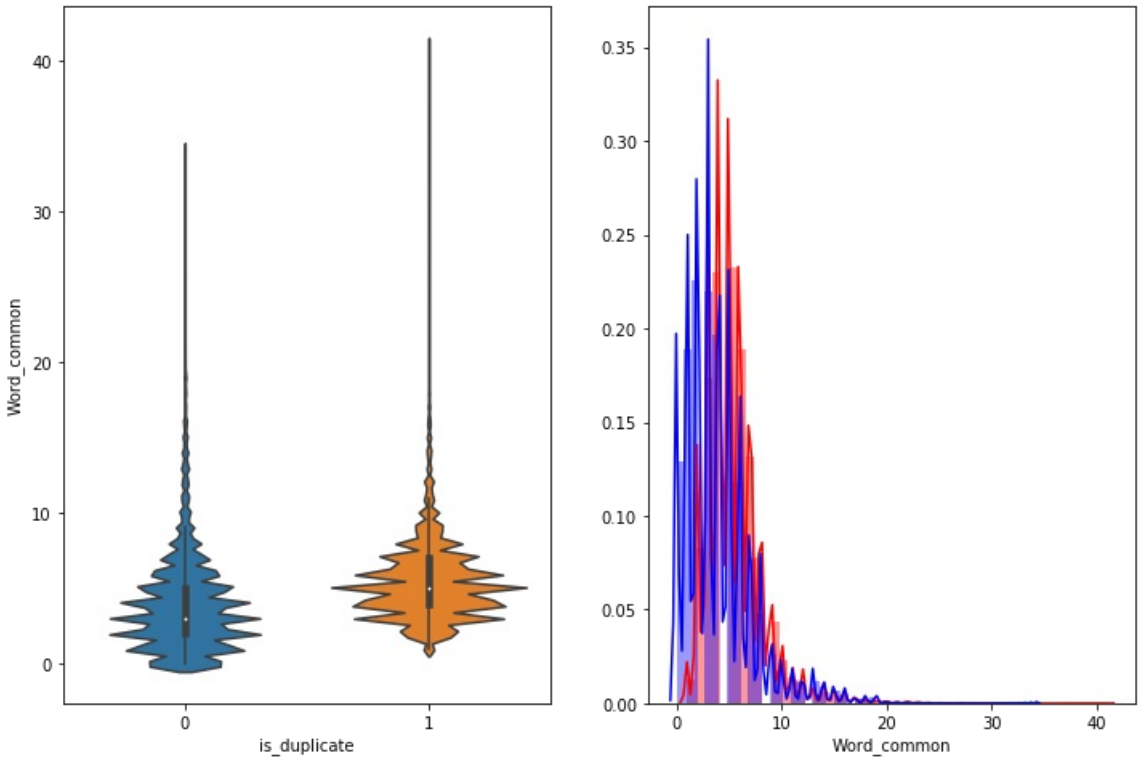


In [19]:

```
plt.figure(figsize=(12, 8))

plt.subplot(1,2,1)
sns.violinplot(x = 'is_duplicate', y = 'Word_common', data = df[0:])

plt.subplot(1,2,2)
sns.distplot(df[df['is_duplicate'] == 1.0]['Word_common'][0:] , label = "1", color = 'red')
sns.distplot(df[df['is_duplicate'] == 0.0]['Word_common'][0:] , label = "0" , color = 'blue' )
plt.show()
```

```
C:\Users\LENOVO\Anaconda3\lib\site-packages\matplotlib\axes\_axes.py:6462: UserWarning:

The 'normed' kwarg is deprecated, and has been replaced by the 'density' kwarg.
```

```
C:\Users\LENOVO\Anaconda3\lib\site-packages\matplotlib\axes\_axes.py:6462: UserWarning:

The 'normed' kwarg is deprecated, and has been replaced by the 'density' kwarg.
```



In [20]:

```
from fuzzywuzzy import fuzz
from sklearn.manifold import TSNE
from wordcloud import WordCloud , STOPWORDS
from os import path
from PIL import Image
```

```
C:\Users\LENOVO\Anaconda3\lib\site-packages\fuzzywuzzy\fuzz.py:11: UserWarning:

Using slow pure-python SequenceMatcher. Install python-Levenshtein to remove this warning
```

In [25]:

```
!pip install fuzzywuzzy
```

```
Requirement already satisfied: fuzzywuzzy in c:\users\lenovo\anaconda3\lib\site-packages (0.17.0)

distributed 1.21.8 requires msgpack, which is not installed.
You are using pip version 10.0.1, however version 18.1 is available.
You should consider upgrading via the 'python -m pip install --upgrade pip' command.
```

In [21]:

```
if os.path.isfile('df_fe_without_preprocessing_train.csv'):
    df = pd.read_csv('df_fe_without_preprocessing_train.csv')
    df.fillna('')
    df.head()
else:
    print('get df_fe_without_preprocessing_train.csv  or run the previous notebook')
```

In [22]:

```
df.head(2)
```

Out[22]:

	id	qid1	qid2	question1	question2	is_duplicate	freq_id1	freq_id2	q1len	q2len	q1_n_words	q2_n_words	Word_con
0	0	1	2	What is the step by step guide to invest in sh...	What is the step by step guide to invest in sh...	0	1	1	66	57	14	12	10.0
1	1	3	4	What is the story of Kohinoor (Koh-i-Noor) Dia...	What would happen if the Indian government sto...	0	4	1	51	88	8	13	4.0

3.4 Preprocessing of Text

- Preprocessing:
  - Removing html tags
  - Removing Punctuations
  - Performing stemming
  - Removing Stopwords
  - Expanding contractions etc.

In [23]:

```
SAFE_DIV= 0.0001
STOP_WORDS = stopwords.words("english")
def preprocessing(x):
    x = str(x).lower()
    x=x.replace(",000,000", "m").replace(",000", "k").replace("’", "'").replace("‘", "'")\
                           .replace("won't", "will not").replace("cannot", "can not").replace("can'
", "can not")\
                           .replace("n't", " not").replace("what's", "what is").replace("it's", "it
is")\
                           .replace("'ve", " have").replace("i'm", "i am").replace("'re", " are")\
                           .replace("he's", "he is").replace("she's", "she is").replace("'s", " own
)\
                           .replace("%", " percent ").replace("₹", " rupee ").replace("$", " dollar
")\
                           .replace("€", " euro ").replace("'ll", " will")
    x = re.sub(r"([0-9]+)000000",r"\1m",x)
    x=re.sub(r"([0-9]+)000",r"\1k",x)
    porter = PorterStemmer()
    pattern = re.compile('\w')
    if type(x)==type(''):
        x=re.sub(pattern,'',x)
    if type(x) == type(''):
        x= porter.stem(x)
        example1 = BeautifulSoup(x)
        x= example1.get_text()
    return x
```

In [24]:

```
from nltk.corpus import stopwords
import re
from nltk.stem import PorterStemmer
from bs4 import BeautifulSoup
```

Definition:

- **Token**: You get a token by splitting sentence a space
- **Stop_Word** : stop words as per NLTK.
- **Word** : A token that is not a stop_word

Features:

- **cwc_min** : Ratio of common_word_count to min lenghth of word count of Q1 and Q2
  cwc_min = common_word_count / (min(len(q1_words), len(q2_words))

- **cwc_max** : Ratio of common_word_count to max lenghth of word count of Q1 and Q2
  cwc_max = common_word_count / (max(len(q1_words), len(q2_words))

- **csc_min** : Ratio of common_stop_count to min lenghth of stop count of Q1 and Q2
  csc_min = common_stop_count / (min(len(q1_stops), len(q2_stops))

- **csc_max** : Ratio of common_stop_count to max lenghth of stop count of Q1 and Q2
  csc_max = common_stop_count / (max(len(q1_stops), len(q2_stops))

- **ctc_min** : Ratio of common_token_count to min lenghth of token count of Q1 and Q2
  ctc_min = common_token_count / (min(len(q1_tokens), len(q2_tokens))

- **ctc_max** : Ratio of common_token_count to max lenghth of token count of Q1 and Q2
  ctc_max = common_token_count / (max(len(q1_tokens), len(q2_tokens))

- **last_word_eq** : Check if First word of both questions is equal or not
  last_word_eq = int(q1_tokens[-1] == q2_tokens[-1])

- **first_word_eq** : Check if First word of both questions is equal or not
  first_word_eq = int(q1_tokens[0] == q2_tokens[0])

- **abs_len_diff** : Abs. length difference
  abs_len_diff = abs(len(q1_tokens) - len(q2_tokens))

- **mean_len** : Average Token Length of both Questions
  mean_len = (len(q1_tokens) + len(q2_tokens))/2

- **fuzz_ratio** : https://github.com/seatgeek/fuzzywuzzy#usage http://chairnerd.seatgeek.com/fuzzywuzzy-fuzzy-string-matching-in-python/

- **fuzz_partial_ratio** : https://github.com/seatgeek/fuzzywuzzy#usage http://chairnerd.seatgeek.com/fuzzywuzzy-fuzzy-string-matching-in-python/

- **token_sort_ratio** : https://github.com/seatgeek/fuzzywuzzy#usage http://chairnerd.seatgeek.com/fuzzywuzzy-fuzzy-string-matching-in-python/

- **token_set_ratio** : https://github.com/seatgeek/fuzzywuzzy#usage http://chairnerd.seatgeek.com/fuzzywuzzy-fuzzy-string-matching-in-python/

- **longest_substr_ratio** : Ratio of length longest common substring to min lenghth of token count of Q1 and Q2
  longest_substr_ratio = len(longest common substring) / (min(len(q1_tokens), len(q2_tokens))

In [25]:

```
def get_token_features(q1, q2):
    token_features = [0.0]*10

    # Converting the Sentence into Tokens:
    q1_tokens = q1.split()
    q2_tokens = q2.split()

    if len(q1_tokens) == 0 or len(q2_tokens) == 0:
        return token_features
    # Get the non-stopwords in Questions
```

```
    q1_words = set([word for word in q1_tokens if word not in STOP_WORDS])
    q2_words = set([word for word in q2_tokens if word not in STOP_WORDS])

    #Get the stopwords in Questions
    q1_stops = set([word for word in q1_tokens if word in STOP_WORDS])
    q2_stops = set([word for word in q2_tokens if word in STOP_WORDS])

    # Get the common non-stopwords from Question pair
    common_word_count = len(q1_words.intersection(q2_words))

    # Get the common stopwords from Question pair
    common_stop_count = len(q1_stops.intersection(q2_stops))

    # Get the common Tokens from Question pair
    common_token_count = len(set(q1_tokens).intersection(set(q2_tokens)))


    token_features[0] = common_word_count / (min(len(q1_words), len(q2_words)) + SAFE_DIV)
    token_features[1] = common_word_count / (max(len(q1_words), len(q2_words)) + SAFE_DIV)
    token_features[2] =common_stop_count / (min(len(q1_stops), len(q2_stops)) + SAFE_DIV)
    token_features[3] = common_stop_count / (max(len(q1_stops), len(q2_stops)) + SAFE_DIV)
    token_features[4] = common_token_count / (min(len(q1_tokens), len(q2_tokens)) + SAFE_DIV)
    token_features[5] = common_token_count / (max(len(q1_tokens), len(q2_tokens)) + SAFE_DIV)

    # Last word of both question is same or not
    token_features[6] = int(q1_tokens[-1] == q2_tokens[-1])

    # First word of both question is same or not
    token_features[7] = int(q1_tokens[0] == q2_tokens[0])

    token_features[8] = abs(len(q1_tokens) - len(q2_tokens))

    #Average Token Length of both Questions
    token_features[9] = (len(q1_tokens) + len(q2_tokens))/2
    return token_features

# get the Longest Common sub string

def get_longest_substr_ratio(a, b):
    strs = list(distance.lcsubstrings(a, b))
    if len(strs) == 0:
        return 0
    else:
        return len(strs[0]) / (min(len(a), len(b)) + 1)

def extract_features(df):
    # preprocessing each question
    df["question1"] = df["question1"].fillna("").apply(preprocessing)
    df["question2"] = df["question2"].fillna("").apply(preprocessing)

    print("token features...")

    # Merging Features with dataset

    token_features = df.apply(lambda x: get_token_features(x["question1"], x["question2"]), axis=1)

    df["cwc_min"]       = list(map(lambda x: x[0], token_features))
    df["cwc_max"]       = list(map(lambda x: x[1], token_features))
    df["csc_min"]       = list(map(lambda x: x[2], token_features))
    df["csc_max"]       = list(map(lambda x: x[3], token_features))
    df["ctc_min"]       = list(map(lambda x: x[4], token_features))
    df["ctc_max"]       = list(map(lambda x: x[5], token_features))
    df["last_word_eq"]  = list(map(lambda x: x[6], token_features))
    df["first_word_eq"] = list(map(lambda x: x[7], token_features))
    df["abs_len_diff"]  = list(map(lambda x: x[8], token_features))
    df["mean_len"]      = list(map(lambda x: x[9], token_features))

    #Computing Fuzzy Features and Merging with Dataset

    # do read this blog: http://chairnerd.seatgeek.com/fuzzywuzzy-fuzzy-string-matching-in-python/
    # https://stackoverflow.com/questions/31806695/when-to-use-which-fuzz-function-to-compare-2-st
rings
    # https://github.com/seatgeek/fuzzywuzzy
    print("fuzzy features..")
    df["token_set_ratio"]       = df.apply(lambda x: fuzz.token_set_ratio(x["question1"],
x["question2"]), axis=1)
    # The token sort approach involves tokenizing the string in question, sorting the tokens alpha
```

```
betically, and
    # then joining them back into a string We then compare the transformed strings with a simple r
atio().
    df["token_sort_ratio"]      = df.apply(lambda x: fuzz.token_sort_ratio(x["question1"],
x["question2"]), axis=1)
    df["fuzz_ratio"]            = df.apply(lambda x: fuzz.QRatio(x["question1"], x["question2"]), a
is=1)
    df["fuzz_partial_ratio"]    = df.apply(lambda x: fuzz.partial_ratio(x["question1"],
x["question2"]), axis=1)
    df["longest_substr_ratio"]  = df.apply(lambda x: get_longest_substr_ratio(x["question1"], x["qu
estion2"]), axis=1)
    return df
```

In [26]:

```
if os.path.isfile('nlp_features_train.csv'):
    df = pd.read_csv("nlp_features_train.csv",encoding='latin-1')
    df.fillna('')
else:
    print("Extracting features for train:")
    df = pd.read_csv("Quora_train.csv")
    df = extract_features(df)
    df.to_csv("nlp_features_train.csv", index=False)
df.head(2)
```

Out[26]:

	id	qid1	qid2	question1	question2	is_duplicate	cwc_min	cwc_max	csc_min	csc_max	...	ctc_max	last_word_eq	f
0	0	1	2	what is the step by step guide to invest in sh...	what is the step by step guide to invest in sh...	0	0.999980	0.833319	0.999983	0.999983	...	0.785709	0.0	1
1	1	3	4	what is the story of kohinoor koh i noor dia...	what would happen if the indian government sto...	0	0.799984	0.399996	0.749981	0.599988	...	0.466664	0.0	1

2 rows × 21 columns

3.5.1 Analysis of extracted features

In [27]:

```
df_duplicate = df[df['is_duplicate'] == 1]
dfp_nonduplicate = df[df['is_duplicate'] == 0]

# Converting 2d array of q1 and q2 and flatten the array: like {{1,2},{3,4}} to {1,2,3,4}
p = np.dstack([df_duplicate["question1"], df_duplicate["question2"]]).flatten()
n = np.dstack([dfp_nonduplicate["question1"], dfp_nonduplicate["question2"]]).flatten()

print ("Number of data points in class 1 (duplicate pairs) :",len(p))
print ("Number of data points in class 0 (non duplicate pairs) :",len(n))

#Saving the np array into a text file
#np.savetxt('train_p.txt', p, delimiter=' ', fmt='%s')
#np.savetxt('train_n.txt', n, delimiter=' ', fmt='%s')
```

```
Number of data points in class 1 (duplicate pairs) : 298526
Number of data points in class 0 (non duplicate pairs) : 510054
```

In [28]:

```
import pickle
def savetofile(obj,filename):
    pickle.dump(obj,open(filename+".txt","wb"))
def openfromfile(filename):
    temp = pickle.load(open(filename+".txt","rb"))
    return temp
```

In [29]:

```
d = path.dirname('.')
textp_w= open(path.join(d,'train_p.txt')).read()
textn_w = open(path.join(d,'train_n.txt')).read()
stopwords = set(STOPWORDS)
stopwords.add("said")
stopwords.add("br")
stopwords.add("")
stopwords.remove("not")
stopwords.remove("no")
stopwords.remove("like")
print ("Total number of words in duplicate pair questions :",len(textp_w))
print ("Total number of words in non duplicate pair questions :",len(textn_w))
```

```
Total number of words in duplicate pair questions : 16109886
Total number of words in non duplicate pair questions : 3335825
```

In [30]:

```
wc = WordCloud(background_color = 'white',max_words= len(textp_w),stopwords = stopwords)
wc.generate(textp_w)
print("word cloud for duplicate Question pairs")
plt.imshow(wc, interpolation='bilinear')
plt.axis("off")
plt.show()
```

```
word cloud for duplicate Question pairs
```

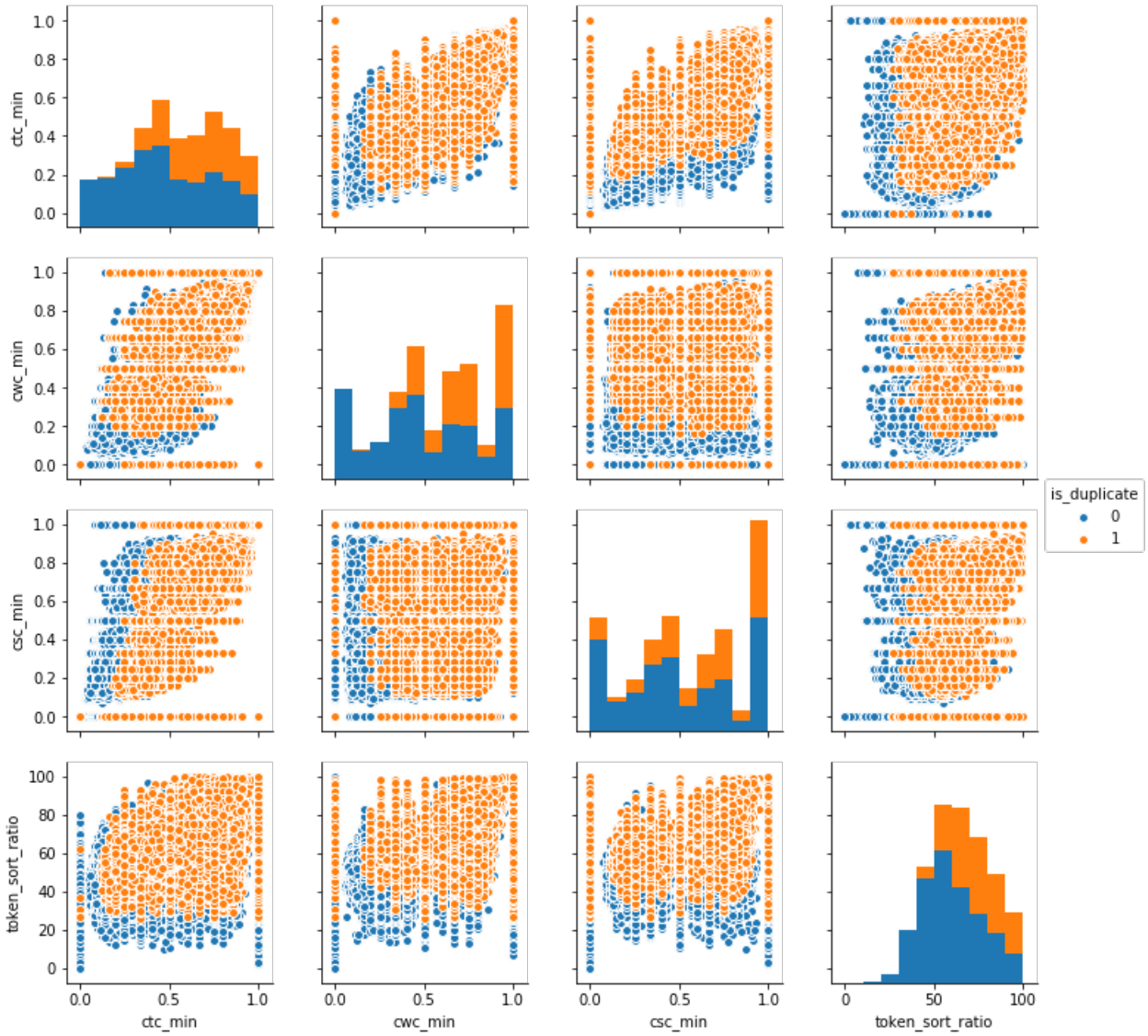
In [31]:

```
wc = WordCloud(background_color="white", max_words=len(textn_w),stopwords=stopwords)
# generate word cloud
wc.generate(textn_w)
print ("Word Cloud for non-Duplicate Question pairs:")
plt.imshow(wc, interpolation='bilinear')
plt.axis("off")
plt.show()
```

```
Word Cloud for non-Duplicate Question pairs:
```

In [32]:

```
n = df.shape[0]
sns.pairplot(df[['ctc_min','cwc_min','csc_min','token_sort_ratio','is_duplicate']][0:n],hue =
'is_duplicate',vars=['ctc_min', 'cwc_min', 'csc_min', 'token_sort_ratio'])
plt.show()
```

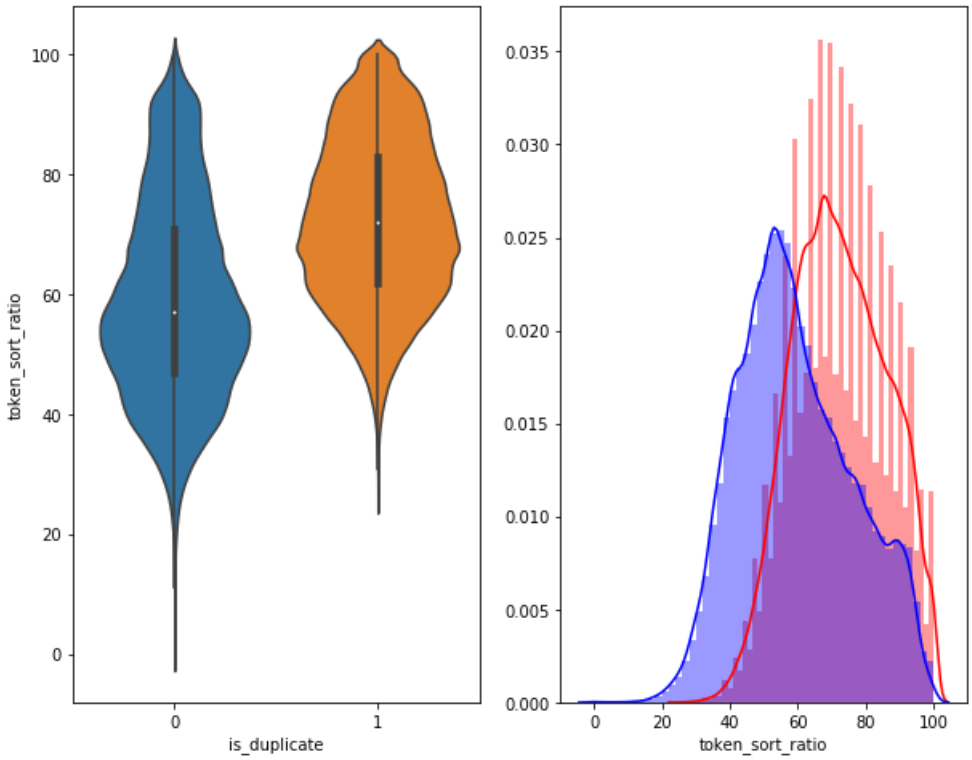


In [33]:

```
# Distribution of the token_sort_ratio
plt.figure(figsize=(10, 8))

plt.subplot(1,2,1)
sns.violinplot(x = 'is_duplicate', y = 'token_sort_ratio', data = df[0:] , )

plt.subplot(1,2,2)
sns.distplot(df[df['is_duplicate'] == 1.0]['token_sort_ratio'][0:] , label = "1", color = 'red')
sns.distplot(df[df['is_duplicate'] == 0.0]['token_sort_ratio'][0:] , label = "0" , color = 'blue' )
plt.show()
```

```
C:\Users\LENOVO\Anaconda3\lib\site-packages\matplotlib\axes\_axes.py:6462: UserWarning:

The 'normed' kwarg is deprecated, and has been replaced by the 'density' kwarg.

C:\Users\LENOVO\Anaconda3\lib\site-packages\matplotlib\axes\_axes.py:6462: UserWarning:
```

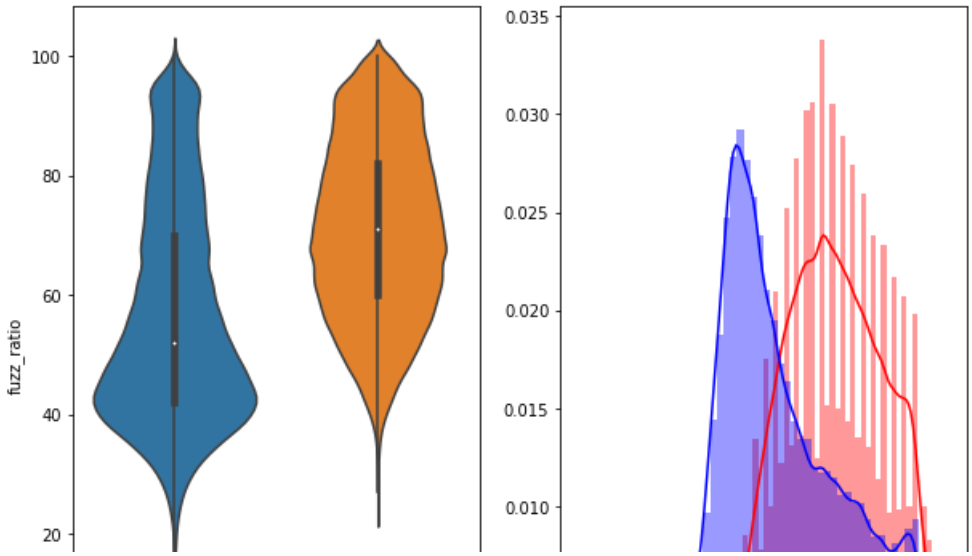
```
The 'normed' kwarg is deprecated, and has been replaced by the 'density' kwarg.
```

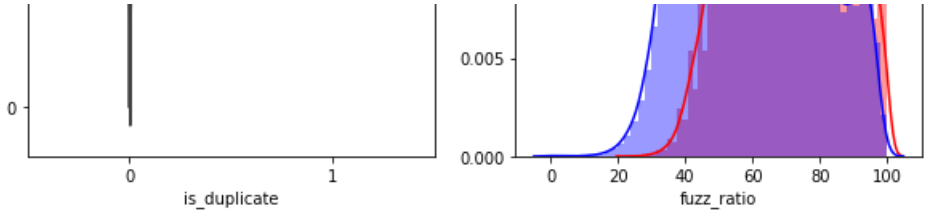


In [34]:

```
plt.figure(figsize=(10, 8))

plt.subplot(1,2,1)
sns.violinplot(x = 'is_duplicate', y = 'fuzz_ratio', data = df[0:] , )

plt.subplot(1,2,2)
sns.distplot(df[df['is_duplicate'] == 1.0]['fuzz_ratio'][0:] , label = "1", color = 'red')
sns.distplot(df[df['is_duplicate'] == 0.0]['fuzz_ratio'][0:] , label = "0" , color = 'blue' )
plt.show()
```

```
C:\Users\LENOVO\Anaconda3\lib\site-packages\matplotlib\axes\_axes.py:6462: UserWarning:

The 'normed' kwarg is deprecated, and has been replaced by the 'density' kwarg.

C:\Users\LENOVO\Anaconda3\lib\site-packages\matplotlib\axes\_axes.py:6462: UserWarning:

The 'normed' kwarg is deprecated, and has been replaced by the 'density' kwarg.
```





TSNE VISUALIZATION

In [58]:

```
from sklearn.preprocessing import MinMaxScaler

dfp_subsampled = df[0:5000]
X = MinMaxScaler().fit_transform(dfp_subsampled[['cwc_min', 'cwc_max', 'csc_min', 'csc_max' ,
'ctc_min' , 'ctc_max' , 'last_word_eq', 'first_word_eq' , 'abs_len_diff' , 'mean_len' , 'token_set_
ratio' , 'token_sort_ratio' ,  'fuzz_ratio' , 'fuzz_partial_ratio' , 'longest_substr_ratio']])
y = dfp_subsampled['is_duplicate'].values
```

In [59]:

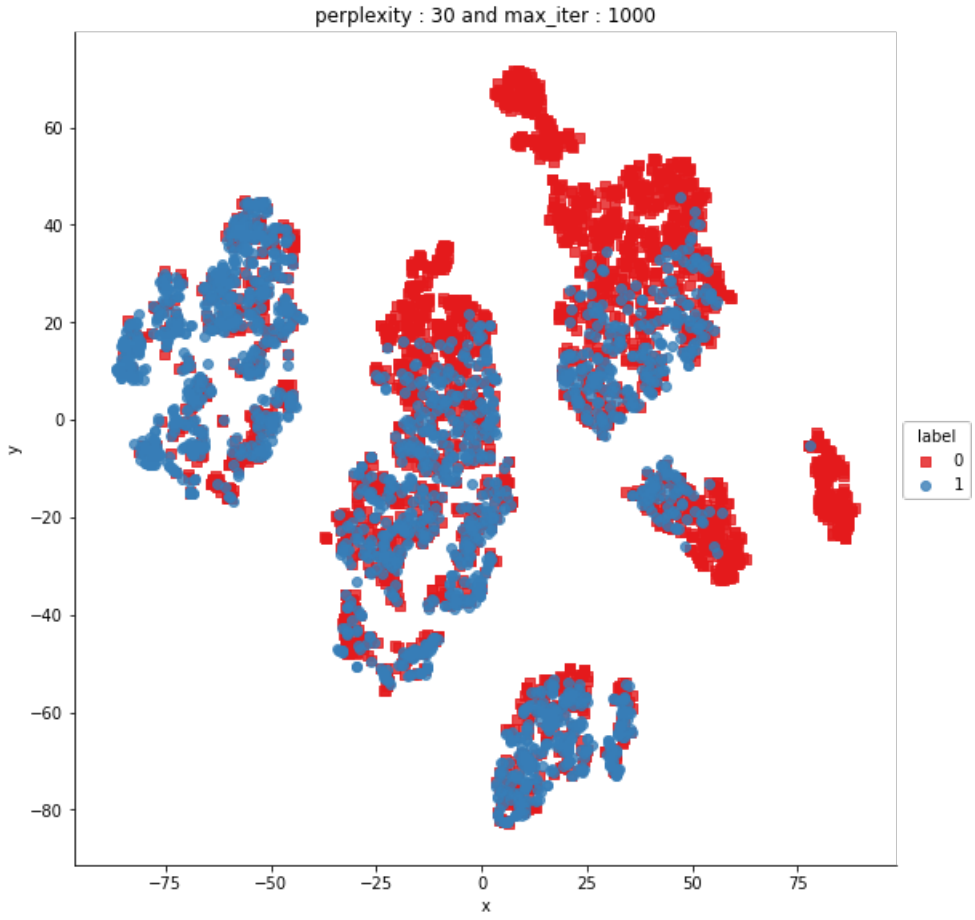
```
tsne2d = TSNE(
    n_components=2,
    init='random', # pca
    random_state=101,
    method='barnes_hut',
    n_iter=1000,
    verbose=2,
    angle=0.5
).fit_transform(X)
```

```
[t-SNE] Computing 91 nearest neighbors...
[t-SNE] Indexed 5000 samples in 0.134s...
[t-SNE] Computed neighbors for 5000 samples in 0.543s...
[t-SNE] Computed conditional probabilities for sample 1000 / 5000
[t-SNE] Computed conditional probabilities for sample 2000 / 5000
[t-SNE] Computed conditional probabilities for sample 3000 / 5000
[t-SNE] Computed conditional probabilities for sample 4000 / 5000
[t-SNE] Computed conditional probabilities for sample 5000 / 5000
[t-SNE] Mean sigma: 0.116557
[t-SNE] Computed conditional probabilities in 0.470s
[t-SNE] Iteration 50: error = 80.8968964, gradient norm = 0.0430571 (50 iterations in 8.061s)
[t-SNE] Iteration 100: error = 70.3833160, gradient norm = 0.0099593 (50 iterations in 6.123s)
[t-SNE] Iteration 150: error = 68.6159134, gradient norm = 0.0056708 (50 iterations in 6.126s)
[t-SNE] Iteration 200: error = 67.7694321, gradient norm = 0.0040581 (50 iterations in 6.194s)
[t-SNE] Iteration 250: error = 67.2746048, gradient norm = 0.0033067 (50 iterations in 6.152s)
[t-SNE] KL divergence after 250 iterations with early exaggeration: 67.274605
[t-SNE] Iteration 300: error = 1.7729300, gradient norm = 0.0011900 (50 iterations in 6.198s)
[t-SNE] Iteration 350: error = 1.3714967, gradient norm = 0.0004818 (50 iterations in 6.026s)
[t-SNE] Iteration 400: error = 1.2036748, gradient norm = 0.0002779 (50 iterations in 5.979s)
[t-SNE] Iteration 450: error = 1.1132656, gradient norm = 0.0001889 (50 iterations in 6.061s)
[t-SNE] Iteration 500: error = 1.0582460, gradient norm = 0.0001434 (50 iterations in 5.998s)
[t-SNE] Iteration 550: error = 1.0222589, gradient norm = 0.0001180 (50 iterations in 6.045s)
[t-SNE] Iteration 600: error = 0.9984865, gradient norm = 0.0001015 (50 iterations in 6.044s)
[t-SNE] Iteration 650: error = 0.9830498, gradient norm = 0.0000958 (50 iterations in 6.106s)
[t-SNE] Iteration 700: error = 0.9726909, gradient norm = 0.0000877 (50 iterations in 6.138s)
[t-SNE] Iteration 750: error = 0.9647216, gradient norm = 0.0000823 (50 iterations in 6.088s)
[t-SNE] Iteration 800: error = 0.9582971, gradient norm = 0.0000755 (50 iterations in 6.073s)
[t-SNE] Iteration 850: error = 0.9531373, gradient norm = 0.0000697 (50 iterations in 6.070s)
[t-SNE] Iteration 900: error = 0.9484153, gradient norm = 0.0000696 (50 iterations in 6.095s)
[t-SNE] Iteration 950: error = 0.9445393, gradient norm = 0.0000659 (50 iterations in 6.061s)
[t-SNE] Iteration 1000: error = 0.9412127, gradient norm = 0.0000674 (50 iterations in 6.810s)
[t-SNE] Error after 1000 iterations: 0.941213
```

In [60]:

```
df = pd.DataFrame({'x':tsne2d[:,0], 'y':tsne2d[:,1] ,'label':y})

# draw the plot in appropriate place in the grid
sns.lmplot(data=df, x='x', y='y', hue='label', fit_reg=False, size=8,palette="Set1",markers=['s','o
```

```
'])
plt.title("perplexity : {} and max_iter : {}".format(30, 1000))
plt.show()
```



In [ ]:

```
trace1 = go.Scatter3d(
    x=tsne3d[:,0],
    y=tsne3d[:,1],
    z=tsne3d[:,2],
    mode='markers',
    marker=dict(
        sizemode='diameter',
        color = y,
        colorscale = 'Portland',
        colorbar = dict(title = 'duplicate'),
        line=dict(color='rgb(255, 255, 255)'),
        opacity=0.75
    )
)

data=[trace1]
layout=dict(height=800, width=800, title='3d embedding with engineered features')
fig=dict(data=data, layout=layout)
py.iplot(fig, filename='3DBubble')
```

In [35]:

```
import pandas as pd
import matplotlib.pyplot as plt
import re
import time
import warnings
import numpy as np
from nltk.corpus import stopwords
from sklearn.preprocessing import normalize
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfVectorizer
warnings.filterwarnings("ignore")
import sys
```

```
import sys
import os
import pandas as pd
import numpy as np
from tqdm import tqdm

# exctract word2vec vectors
# https://github.com/explosion/spaCy/issues/1721
# http://landinghub.visualstudio.com/visual-cpp-build-tools
```

In [36]:

```
# avoid decoding problems
df = pd.read_csv("Quora_train.csv")

# encode questions to unicode
# https://stackoverflow.com/a/6812069
# ----------------- python 2 ---------------------
# df['question1'] = df['question1'].apply(lambda x: unicode(str(x),"utf-8"))
# df['question2'] = df['question2'].apply(lambda x: unicode(str(x),"utf-8"))
# ----------------- python 3 ---------------------
df['question1'] = df['question1'].apply(lambda x: str(x))
df['question2'] = df['question2'].apply(lambda x: str(x))
```

In [37]:

```
df.head()
```

Out[37]:

	id	qid1	qid2	question1	question2	is_duplicate
0	0	1	2	What is the step by step guide to invest in sh...	What is the step by step guide to invest in sh...	0
1	1	3	4	What is the story of Kohinoor (Koh-i-Noor) Dia...	What would happen if the Indian government sto...	0
2	2	5	6	How can I increase the speed of my internet co...	How can Internet speed be increased by hacking...	0
3	3	7	8	Why am I mentally very lonely? How can I solve...	Find the remainder when [math]23^{24}[/math] i...	0
4	4	9	10	Which one dissolve in water quikly sugar, salt...	Which fish would survive in salt water?	0

In [38]:

```
from sklearn.feature_extraction.text import TfidfVectorizer
```

In [39]:

```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.feature_extraction.text import CountVectorizer
questions = list(df['question1']) + list(df['question2'])
tfidf = TfidfVectorizer(lowercase=False,)
tfidf.fit_transform(questions)
word2tfidf = dict(zip(tfidf.get_feature_names(), tfidf.idf_))
```

In [40]:

```
word2tfidf
```

Out[40]:

```
{'00': 10.04915924690928,
 '000': 7.304591037001674,
 '0000': 13.909888957949875,
 '000000': 13.909888957949875,
 '00000000': 13.909888957949875,
 '0000000000': 13.50442384984171,
 '0000001': 13.909888957949875,
 '00000074': 13.909888957949875
```

```
'0000021210': 13.909888957949875,
'00001': 13.909888957949875,
'0001': 13.909888957949875,
'0002': 13.909888957949875,
'000INR': 13.909888957949875,
'000K': 13.909888957949875,
'000Rs': 12.811276669281765,
'000USD': 13.909888957949875,
'000WebHost': 13.909888957949875,
'000ft': 13.909888957949875,
'000rs': 13.909888957949875,
'000s': 13.909888957949875,
'000webhost': 13.909888957949875,
'001': 12.811276669281765,
'0015': 13.909888957949875,
'0019': 12.99359822607572,
'002': 13.21674177738993,
'0021': 13.909888957949875,
'0035': 13.909888957949875,
'003SC': 13.909888957949875,
'005': 13.909888957949875,
'0051': 13.50442384984171,
'007': 13.50442384984171,
'008': 13.50442384984171,
'008801703772104': 13.909888957949875,
'0088882578743': 13.909888957949875,
'00AM': 13.50442384984171,
'00GHz': 13.909888957949875,
'00V': 13.909888957949875,
'00am': 12.11812948872182,
'00g': 13.909888957949875,
'00pm': 13.50442384984171,
'00s': 13.909888957949875,
'01': 10.59570295327735,
'0113': 13.909888957949875,
'012': 13.21674177738993,
'013J': 13.909888957949875,
'014': 13.50442384984171,
'018v4': 13.909888957949875,
'01GBP': 13.909888957949875,
'01st': 13.909888957949875,
'02': 11.963978808894563,
'020': 13.909888957949875,
'02038': 13.909888957949875,
'022': 13.50442384984171,
'025': 13.909888957949875,
'026': 13.909888957949875,
'03': 11.769822794453605,
'032': 13.50442384984171,
'0320': 13.909888957949875,
'034': 13.50442384984171,
'035': 13.909888957949875,
'035tx': 13.50442384984171,
'036': 13.909888957949875,
'038': 13.909888957949875,
'03g': 13.909888957949875,
'04': 9.920904911385602,
'04009': 13.50442384984171,
'041x': 13.909888957949875,
'042': 13.909888957949875,
'045': 13.909888957949875,
'04LTS': 13.909888957949875,
'05': 11.65859715934338,
'050': 13.21674177738993,
'052': 13.909888957949875,
'0524': 13.909888957949875,
'05246': 13.909888957949875,
'054': 13.909888957949875,
'055': 13.909888957949875,
'0569': 13.909888957949875,
'058': 13.50442384984171,
'05_mar': 13.909888957949875,
'06': 11.270831628334617,
'062': 13.50442384984171,
'065g': 13.909888957949875,
'066': 13.50442384984171,
```

```
'0665745511651039': 13.909888957949875,
'067972477X': 13.909888957949875,
'068': 13.50442384984171,
'06x08': 13.909888957949875,
'07': 11.963978808894563,
'0706': 13.909888957949875,
'072': 13.909888957949875,
'0745': 13.909888957949875,
'075': 13.909888957949875,
'07904378172': 13.909888957949875,
'079tx': 13.909888957949875,
'07ish': 13.909888957949875,
'08': 11.894985937407611,
'080': 13.909888957949875,
'083': 13.909888957949875,
'086': 13.21674177738993,
'088': 13.50442384984171,
'08888888888': 13.909888957949875,
'088989914321': 13.909888957949875,
'08GB': 13.909888957949875,
'09': 11.137300235710095,
'090': 13.50442384984171,
'099': 13.50442384984171,
'09987510597': 13.909888957949875,
'0A': 13.50442384984171,
'0GB': 13.909888957949875,
'0GHz': 13.21674177738993,
'0L': 13.909888957949875,
'0MQ': 13.909888957949875,
'0V': 12.99359822607572,
'0XC00000E9': 13.909888957949875,
'0and': 13.909888957949875,
'0are': 13.909888957949875,
'0dev': 13.909888957949875,
'0f': 13.909888957949875,
'0g': 13.50442384984171,
'0kg': 13.909888957949875,
'0l': 13.909888957949875,
'0m': 13.21674177738993,
'0n7': 13.909888957949875,
'0s': 12.811276669281765,
'0w': 13.21674177738993,
'0x': 13.50442384984171,
'0x0000605': 13.909888957949875,
'0x0020': 13.50442384984171,
'0x0c0007': 13.909888957949875,
'0x80004004': 13.909888957949875,
'0x80070422': 13.50442384984171,
'0x80070424': 13.909888957949875,
'0x80070426': 12.99359822607572,
'0x80072efd': 13.909888957949875,
'0x80072f78': 13.909888957949875,
'0x80073CF6': 13.909888957949875,
'0x80073cf0': 13.909888957949875,
'0x800f081f': 13.50442384984171,
'0xc000007b': 13.50442384984171,
'0xc0000142': 13.50442384984171,
'0xc000014c': 13.909888957949875,
'0xc0000225': 13.909888957949875,
'0xc00007b': 13.909888957949875,
'0xdde224f0': 13.50442384984171,
'10': 6.198339978320729,
'100': 7.244205240167467,
'1000': 5.742820779608637,
'10000': 9.319832409771832,
'100000': 11.60730386495583,
'1000000': 12.657125989454507,
'10000000': 13.21674177738993,
'100000000': 12.99359822607572,
'10000000000': 13.909888957949875,
'10000INR': 12.657125989454507,
'10000mAH': 13.909888957949875,
'10000mAh': 13.909888957949875,
'10000rs': 13.50442384984171,
'10000²': 13.909888957949875,
'1000MB': 13.21674177738993,
'1000Rs': 11.46754192258067,
'1000cc': 13.909888957949875,
```

```
'1000cc': 13.909888957949875,
'1000k': 13.909888957949875,
'1000kb': 13.50442384984171,
'1000km': 13.909888957949875,
'1000mAh': 13.909888957949875,
'1000mg': 13.50442384984171,
'1000note': 13.909888957949875,
'1000p': 13.909888957949875,
'1000rpm': 13.909888957949875,
'1000rs': 11.047688077020407,
'1000s': 12.405811561173602,
'1000sq': 13.909888957949875,
'1000tKB': 13.50442384984171,
'1000th': 13.909888957949875,
'1000v': 13.909888957949875,
'1000views': 13.50442384984171,
'1000w': 13.50442384984171,
'1001': 13.50442384984171,
'100100': 13.50442384984171,
'10041': 13.909888957949875,
'10050': 13.909888957949875,
'1006': 13.50442384984171,
'1008': 13.909888957949875,
'1009': 13.909888957949875,
'100AUD': 13.909888957949875,
'100Ah': 13.50442384984171,
'100C': 13.50442384984171,
'100CC': 13.909888957949875,
'100CE': 13.909888957949875,
'100K': 11.10652857704334,
'100KW': 13.909888957949875,
'100M': 12.811276669281765,
'100MB': 13.50442384984171,
'100MW': 13.909888957949875,
'100Mbps': 12.657125989454507,
'100P2': 13.909888957949875,
'100US': 13.909888957949875,
'100V': 13.50442384984171,
'100W': 13.50442384984171,
'100ahr': 13.50442384984171,
'100bestbuy': 13.909888957949875,
'100bhp': 13.909888957949875,
'100cc': 12.811276669281765,
'100degrees': 13.909888957949875,
'100ft': 13.50442384984171,
'100g': 13.21674177738993,
'100gm': 13.50442384984171,
'100k': 10.196316891245568,
'100kb': 13.50442384984171,
'100kg': 13.909888957949875,
'100km': 13.909888957949875,
'100kw': 13.21674177738993,
'100lbs': 13.909888957949875,
'100m': 11.60730386495583,
'100mA': 13.50442384984171,
'100mAh': 13.909888957949875,
'100marks': 13.909888957949875,
'100mb': 13.50442384984171,
'100mcg': 13.909888957949875,
'100mg': 12.99359822607572,
'100mm': 13.50442384984171,
'100mph': 13.909888957949875,
'100ng': 13.909888957949875,
'100rs': 13.909888957949875,
'100s': 13.50442384984171,
'100sft': 13.909888957949875,
'100square': 13.909888957949875,
'100th': 13.21674177738993,
'100usd': 13.909888957949875,
'100w': 12.99359822607572,
'101': 11.10652857704334,
'1010KB': 13.50442384984171,
'1011': 13.909888957949875,
'10129': 13.909888957949875,
'1013': 13.909888957949875,
'1015': 13.909888957949875,
'1016': 13.21674177738993,
'10170': 13.909888957949875,
```

```
'10170': 13.90988895794987 5,
'101yr': 13.909888957949875,
'102': 11.65859715934338,
'1020': 13.21674177738993,
'102000': 13.909888957949875,
'1022': 13.50442384984171,
'1023': 13.909888957949875,
'1024': 11.511993685151504,
'1024MB': 13.21674177738993,
'1027': 13.909888957949875,
'1028': 13.50442384984171,
'103': 12.523594596829986,
'1030': 13.50442384984171,
'1031': 12.99359822607572,
'103lbs': 13.909888957949875,
'104': 12.20514086571145,
'1040': 13.50442384984171,
'10400': 13.21674177738993,
'1040mph': 13.909888957949875,
'1042s': 13.909888957949875,
'1048576': 13.50442384984171,
'105': 11.270831628334617,
'1050': 12.11812948872182,
'10500': 13.909888957949875,
'1053': 13.909888957949875,
'105306812205': 13.909888957949875,
'1055': 13.50442384984171,
'1056': 13.909888957949875,
'1057': 13.909888957949875,
'105hp': 13.909888957949875,
'105k': 13.909888957949875,
'105mm': 13.909888957949875,
'106': 12.300451045515775,
'1060': 11.46754192258067,
'1065': 13.21674177738993,
'1066': 13.909888957949875,
'1067': 13.909888957949875,
'1068': 13.50442384984171,
'107': 11.38416031364162,
'1070': 11.60730386495583,
'1070s': 13.909888957949875,
'1075': 13.50442384984171,
'10770': 13.909888957949875,
'1079rank': 13.909888957949875,
'108': 11.65859715934338,
'1080': 11.769822794453605,
'1080P': 13.50442384984171,
'1080i': 12.300451045515775,
'1080p': 10.208586983837382,
'1080ti': 13.909888957949875,
'1080x2560': 13.909888957949875,
'108K': 13.909888957949875,
'108y': 13.909888957949875,
'109': 12.20514086571145,
'1090': 13.909888957949875,
'1091': 13.909888957949875,
'1097': 13.909888957949875,
'1099': 11.558513700786397,
'1099s': 13.50442384984171,
'109cc': 13.909888957949875,
'109km': 13.909888957949875,
'10A': 13.21674177738993,
'10B': 13.909888957949875,
'10C': 13.909888957949875,
'10CGPA': 12.99359822607572,
'10D': 13.50442384984171,
'10F': 13.909888957949875,
'10G': 13.909888957949875,
'10GB': 13.909888957949875,
'10Gb': 13.909888957949875,
'10H': 13.909888957949875,
'10K': 11.10652857704334,
'10Kg': 13.909888957949875,
'10Kgs': 13.909888957949875,
'10L': 13.50442384984171,
'10M': 13.50442384984171,
'10MM': 13.909888957949875,
```

```
'10Mbps': 13.909888957949875,
'10P': 13.909888957949875,
'10PM': 13.909888957949875,
'10Q': 13.909888957949875,
'10RNC': 13.909888957949875,
'10US': 13.909888957949875,
'10V': 13.909888957949875,
'10W': 13.50442384984171,
'10X': 13.21674177738993,
'10Y': 13.909888957949875,
'10a': 13.909888957949875,
'10am': 13.21674177738993,
'10cgpa': 13.21674177738993,
'10cl': 13.909888957949875,
'10cm': 13.21674177738993,
'10d': 13.909888957949875,
'10e3': 13.50442384984171,
'10ft': 13.50442384984171,
'10g': 13.21674177738993,
'10i': 13.909888957949875,
'10k': 9.314769107815286,
'10kg': 11.344939600488338,
'10kgs': 12.99359822607572,
'10km': 13.909888957949875,
'10lac': 13.909888957949875,
'10lacs': 13.909888957949875,
'10lbs': 13.50442384984171,
'10m': 11.65859715934338,
'10m3': 13.909888957949875,
'10mL': 13.909888957949875,
'10mb': 13.909888957949875,
'10mbps': 13.909888957949875,
'10mg': 12.038086781048284,
'10mins': 13.909888957949875,
'10mm': 12.99359822607572,
'10months': 13.50442384984171,
'10ms': 13.909888957949875,
'10n': 13.909888957949875,
'10oz': 13.909888957949875,
'10pm': 13.21674177738993,
'10points': 13.909888957949875,
'10rs': 13.50442384984171,
'10s': 12.99359822607572,
'10sec': 13.50442384984171,
'10sqrt': 13.909888957949875,
'10th': 8.606584049890799,
'10times': 13.909888957949875,
'10v': 13.909888957949875,
'10w': 13.21674177738993,
'10x': 12.300451045515775,
'10years': 13.21674177738993,
'10yrs': 13.50442384984171,
'10z': 13.909888957949875,
'10π': 13.909888957949875,
'11': 7.688305245791258,
'110': 10.354540896460461,
'1100': 11.83044741627004,
'11000': 12.99359822607572,
'11000INR': 13.21674177738993,
'11001': 13.909888957949875,
'110017': 13.909888957949875,
'1100D': 13.50442384984171,
'1101': 13.909888957949875,
'1101101': 13.909888957949875,
'11023': 13.909888957949875,
'1108': 13.909888957949875,
'110V': 11.344939600488338,
'110k': 13.21674177738993,
'110kgs': 13.909888957949875,
'110kw': 13.50442384984171,
'110lbs': 13.909888957949875,
'110m': 13.50442384984171,
'110mm': 13.909888957949875,
'110v': 12.300451045515775,
'111': 11.963978808894563,
'1111': 13.50442384984171,
'11110': 13.909888957949875,
```

```
'1112': 12.038086781048284,
'111221': 13.909888957949875,
'112': 11.60730386495583,
'1120': 13.909888957949875,
'11200': 13.909888957949875,
'11217': 13.909888957949875,
'1125': 13.909888957949875,
'113': 12.523594596829986,
'11301': 12.99359822607572,
'113y': 13.909888957949875,
'114': 12.038086781048284,
'1140': 13.909888957949875,
'1143': 13.909888957949875,
'1146': 13.909888957949875,
'115': 11.30719927250549,
'1150': 13.50442384984171,
'1151': 13.909888957949875,
'11516': 13.909888957949875,
'115197': 13.909888957949875,
'115B': 13.909888957949875,
'115K': 13.909888957949875,
'115L': 13.909888957949875,
'115v': 13.909888957949875,
'116': 12.657125989454507,
'116bpm': 13.50442384984171,
'116tx': 13.909888957949875,
'117': 11.894985937407611,
'1170': 12.99359822607572,
'118': 12.99359822607572,
'1180': 13.909888957949875,
'11808': 13.909888957949875,
'119': 12.811276669281765,
'1190': 13.50442384984171,
'11900': 13.909888957949875,
'11969': 13.909888957949875,
'119GB': 13.50442384984171,
'11AM': 13.909888957949875,
'11B': 13.909888957949875,
'11C': 13.909888957949875,
'11G': 13.909888957949875,
'11K': 13.909888957949875,
'11KV': 13.21674177738993,
'11PM': 13.50442384984171,
'11ac': 13.50442384984171,
'11am': 13.909888957949875,
'11and12': 13.909888957949875,
'11b': 13.909888957949875,
'11class': 13.909888957949875,
'11days': 12.657125989454507,
'11g': 12.657125989454507,
'11gb': 13.909888957949875,
'11give': 13.909888957949875,
'11kV': 13.909888957949875,
'11kg': 13.909888957949875,
'11kv': 13.50442384984171,
'11n': 13.21674177738993,
'11pm': 13.21674177738993,
'11s': 12.811276669281765,
'11th': 9.143450624365661,
'11x': 13.50442384984171,
'11x11': 13.909888957949875,
'11y': 13.50442384984171,
'11years': 13.21674177738993,
'11yrs': 13.909888957949875,
'11µF': 13.50442384984171,
'12': 7.348858292053302,
'120': 9.949075788352298,
'1200': 11.235740308523347,
'12000': 10.711215840399195,
'120000': 13.50442384984171,
'1200D': 12.038086781048284,
'1200cc': 13.50442384984171,
'1200d': 13.21674177738993,
'1200hp': 13.909888957949875,
'1200m': 13.909888957949875,
'1200watts': 13.909888957949875,
'1202': 13.50442384984171,
```

```
'12031': 13.50442384984171,
'1204': 13.50442384984171,
'12071': 13.909888957949875,
'120Hz': 13.909888957949875,
'120K': 12.811276669281765,
'120V': 12.523594596829986,
'120gb': 13.909888957949875,
'120k': 12.405811561173602,
'120kg': 13.909888957949875,
'120km': 13.909888957949875,
'120kms': 13.909888957949875,
'120lbs': 13.909888957949875,
'120m': 13.50442384984171,
'120mm': 13.21674177738993,
'120v': 13.21674177738993,
'121': 12.405811561173602,
'1210': 13.909888957949875,
'1211': 13.50442384984171,
'1212': 13.909888957949875,
'12124': 13.50442384984171,
'1213': 13.909888957949875,
'121392': 13.909888957949875,
'12163': 13.50442384984171,
'121832': 13.909888957949875,
'122': 11.894985937407611,
'12203': 13.909888957949875,
'1221': 13.909888957949875,
'1225': 13.909888957949875,
'122600': 13.909888957949875,
'12295': 13.909888957949875,
'122nd': 13.909888957949875,
'123': 11.894985937407611,
'12300': 13.909888957949875,
'1234': 13.21674177738993,
'12345': 13.21674177738993,
'123456': 13.50442384984171,
'1234567': 13.50442384984171,
'12345678': 13.909888957949875,
'123RF': 13.909888957949875,
'123exchanges': 13.50442384984171,
'123greetings': 13.909888957949875,
'123movies': 13.909888957949875,
'123rf': 13.50442384984171,
'124': 12.11812948872182,
'1240': 12.99359822607572,
'1240nm': 13.909888957949875,
'1245': 13.909888957949875,
'125': 10.47590175346473,
'1250': 13.909888957949875,
'12500': 13.909888957949875,
'125000': 13.909888957949875,
'125126': 13.909888957949875,
'125353fshsgsjhs2527bwv2u888': 13.909888957949875,
'1257': 13.909888957949875,
'125700': 13.909888957949875,
'12577': 13.50442384984171,
'1258': 13.21674177738993,
'12587': 13.909888957949875,
'12593': 13.50442384984171,
'125AMP': 13.909888957949875,
'125CC': 13.50442384984171,
'125K': 13.909888957949875,
'125V': 13.909888957949875,
'125cc': 12.523594596829986,
'125k': 13.50442384984171,
'125miles': 13.909888957949875,
'125mph': 13.909888957949875,
'126': 11.83044741627004,
'1260': 13.909888957949875,
'1267': 13.909888957949875,
'127': 11.769822794453605,
'1270': 13.50442384984171,
'127000': 13.909888957949875,
'1279': 13.909888957949875,
'127lbs': 13.909888957949875,
'128': 10.889464071805513,
'1280': 13.909888957949875,
```

```
'12800s': 13.909888957949875,
'12801': 13.909888957949875,
'12802': 13.909888957949875,
'1280x720': 13.909888957949875,
'1285': 13.909888957949875,
'12890': 13.909888957949875,
'12893': 13.909888957949875,
'128GB': 12.20514086571145,
'128bit': 13.909888957949875,
'128gb': 12.657125989454507,
'128mb': 13.909888957949875,
'129': 12.811276669281765,
'1290': 13.21674177738993,
'1293LZ': 13.909888957949875,
'12975': 13.909888957949875,
'1298': 13.909888957949875,
'12AM': 13.50442384984171,
'12BB': 13.50442384984171,
'12GB': 13.21674177738993,
'12Gb': 13.909888957949875,
'12H143': 13.909888957949875,
'12IU': 13.909888957949875,
'12K': 13.21674177738993,
'12Lac': 13.909888957949875,
'12MP': 13.909888957949875,
'12PM': 13.909888957949875,
'12V': 11.83044741627004,
'12XX': 13.909888957949875,
'12am': 12.99359822607572,
'12c': 12.405811561173602,
'12class': 13.909888957949875,
'12cm': 13.909888957949875,
'12cubic': 13.909888957949875,
'12g': 13.50442384984171,
'12gb': 13.909888957949875,
'12inch': 13.909888957949875,
'12k': 10.865366520226452,
'12km': 13.909888957949875,
'12mbps': 13.909888957949875,
'12mm': 12.811276669281765,
'12oz': 13.909888957949875,
'12pm': 12.20514086571145,
'12std': 13.21674177738993,
'12th': 7.963160304683285,
'12thclass': 13.909888957949875,
'12thee': 13.50442384984171,
'12ths': 13.909888957949875,
'12v': 11.46754192258067,
'12volt': 13.909888957949875,
'12x': 13.21674177738993,
'12yo': 13.909888957949875,
'13': 7.8732120680203455,
'130': 10.525498694604101,
'1300': 12.11812948872182,
'13000': 11.424982308161875,
'13000mAh': 13.21674177738993,
'1300D': 11.46754192258067,
'1300USD': 13.909888957949875,
'1300d': 12.657125989454507,
'1300mA': 12.99359822607572,
'13014': 13.909888957949875,
'13060': 13.909888957949875,
'130K': 12.99359822607572,
'130cc': 13.50442384984171,
'130ft': 13.909888957949875,
'130k': 12.811276669281765,
'130lb': 13.909888957949875,
'130lbs': 13.909888957949875,
'130s': 13.909888957949875,
'131': 12.99359822607572,
'1310': 13.21674177738993,
'13101988': 13.909888957949875,
'13112221': 13.909888957949875,
'1313': 13.909888957949875,
'1314': 13.909888957949875,
'132': 12.405811561173602,
'1320': 13.909888957949875,
```

```
'132Kv': 13.909888957949875,
'133': 13.21674177738993,
'1331': 13.909888957949875,
'1333mhz': 13.909888957949875,
'1334': 13.909888957949875,
'1337': 12.405811561173602,
'133MB': 13.909888957949875,
'134': 12.523594596829986,
'1344': 13.909888957949875,
'1345': 13.909888957949875,
'1348': 13.50442384984171,
'134bp': 13.909888957949875,
'135': 11.235740308523347,
'13500': 13.50442384984171,
'13538': 13.909888957949875,
'1354': 13.909888957949875,
'135k': 13.909888957949875,
'135kg': 13.909888957949875,
'135mm': 13.21674177738993,
'136': 12.811276669281765,
'1361': 11.424982308161875,
'13615': 13.909888957949875,
'1366': 13.50442384984171,
'1366x768': 13.50442384984171,
'137': 12.405811561173602,
'1370': 13.909888957949875,
'1373': 13.909888957949875,
'138': 12.405811561173602,
'1380': 13.909888957949875,
'13800': 13.909888957949875,
'138th': 13.909888957949875,
'1394': 13.909888957949875,
'13B': 13.909888957949875,
'13ISK': 13.909888957949875,
'13K': 13.21674177738993,
'13L': 13.909888957949875,
'13MP': 13.909888957949875,
'13b': 13.50442384984171,
'13cm': 13.50442384984171,
'13h': 13.909888957949875,
'13inch': 13.909888957949875,
'13k': 12.523594596829986,
'13m': 13.909888957949875,
'13mbs': 13.909888957949875,
'13oz': 13.909888957949875,
'13th': 10.939474492380175,
'13x': 13.50442384984171,
'13year': 13.909888957949875,
'13years': 13.909888957949875,
'13yo': 13.909888957949875,
'14': 7.929738170742167,
'140': 10.340356261468505,
'1400': 12.20514086571145,
'14000': 12.300451045515775,
'14001': 13.909888957949875,
'14015': 13.909888957949875,
'14060M': 13.909888957949875,
'14097': 13.909888957949875,
'140K': 13.21674177738993,
'140M': 13.909888957949875,
'140MM': 13.909888957949875,
'140k': 12.99359822607572,
'140kg': 13.909888957949875,
'140mm': 12.300451045515775,
'140s': 13.909888957949875,
'141': 12.811276669281765,
'1412': 13.909888957949875,
'14159265359': 13.909888957949875,
'1416': 13.909888957949875,
'14196': 13.21674177738993,
'141V': 13.909888957949875,
'141v': 13.909888957949875,
'142': 12.405811561173602,
'14200': 13.909888957949875,
'142V': 13.50442384984171,
'143': 12.657125989454507,
'1430': 13.909888957949875,
```

```
'1432': 13.909888957949875,
'1434': 13.50442384984171,
'1437': 13.909888957949875,
'143b': 13.50442384984171,
'143movies': 13.50442384984171,
'144': 11.30719927250549,
'1440': 13.50442384984171,
'144002': 13.909888957949875,
'1440x2560': 13.909888957949875,
'1444': 13.50442384984171,
'14443': 13.909888957949875,
'144Hz': 13.50442384984171,
'144hz': 13.21674177738993,
'144k': 12.99359822607572,
'144v': 13.909888957949875,
'145': 11.963978808894563,
'1450': 13.909888957949875,
'1453': 13.909888957949875,
'145k': 13.909888957949875,
'146': 13.21674177738993,
'14601': 13.909888957949875,
'147': 11.894985937407611,
'1470': 13.909888957949875,
'1475fps': 13.21674177738993,
'147cm': 13.909888957949875,
'148': 11.963978808894563,
'1480': 13.50442384984171,
'1485': 13.909888957949875,
'149': 12.20514086571145,
'1490': 13.50442384984171,
'1491': 13.50442384984171,
'1492': 13.21674177738993,
'1495': 13.909888957949875,
'149Q': 13.909888957949875,
'149V': 13.909888957949875,
'14D': 13.909888957949875,
'14L': 13.909888957949875,
'14R': 13.909888957949875,
'14cm': 13.50442384984171,
'14days': 11.558513700786397,
'14e': 13.909888957949875,
'14g': 13.909888957949875,
'14k': 12.99359822607572,
'14nm': 13.909888957949875,
'14posts': 13.909888957949875,
'14th': 10.542593127963402,
'14year': 13.909888957949875,
'15': 7.337606415255868,
'150': 9.23706012348797,
'1500': 10.428648868614184,
'15000': 8.4065920107153,
'150000': 13.21674177738993,
'15000rs': 12.811276669281765,
'1500MHz': 13.909888957949875,
'1500cc': 13.50442384984171,
'1500km': 13.909888957949875,
'1500mAh': 13.909888957949875,
'1502': 13.909888957949875,
'1505M': 13.909888957949875,
'150AH': 13.909888957949875,
'150CC': 13.50442384984171,
'150K': 13.21674177738993,
'150R': 12.405811561173602,
'150cc': 11.83044741627004,
'150cm': 13.909888957949875,
'150ft': 13.909888957949875,
'150gms': 13.909888957949875,
'150ibs': 13.909888957949875,
'150k': 12.20514086571145,
'150kg': 13.909888957949875,
'150km': 13.909888957949875,
'150lbs': 13.21674177738993,
'150m': 13.50442384984171,
'150mbps': 13.50442384984171,
'150mm': 13.50442384984171,
'150pm': 13.909888957949875,
'150r': 13.50442384984171,
```

```
'151': 11.712664380613656,
'1510': 13.909888957949875,
'15100': 13.909888957949875,
'151012': 13.21674177738993,
'151271': 13.50442384984171,
'1513': 13.909888957949875,
'1515': 13.909888957949875,
'1515HD': 13.909888957949875,
'1516': 13.909888957949875,
'152': 12.20514086571145,
'1520': 12.523594596829986,
'1525': 12.405811561173602,
'152k': 13.909888957949875,
'153': 11.894985937407611,
'1536': 13.21674177738993,
'153V': 13.909888957949875,
'153cm': 13.50442384984171,
'154': 12.405811561173602,
'1547': 13.909888957949875,
'1549': 12.99359822607572,
'154V': 13.909888957949875,
'155': 10.965449978783434,
'1550': 13.909888957949875,
'15500': 13.909888957949875,
'1550nm': 13.50442384984171,
'1557': 13.50442384984171,
'155cc': 13.50442384984171,
'155mbs': 13.909888957949875,
'156': 12.300451045515775,
'15600': 12.99359822607572,
'15625': 13.909888957949875,
'1564': 13.909888957949875,
'15666': 13.909888957949875,
'156cm': 13.909888957949875,
'157': 12.20514086571145,
'1570': 13.909888957949875,
'1571': 13.909888957949875,
'1575M': 13.909888957949875,
'1577': 13.909888957949875,
'1579': 13.909888957949875,
'157cm': 13.909888957949875,
'158': 12.523594596829986,
'15800': 13.909888957949875,
'1588': 13.909888957949875,
'159': 12.300451045515775,
'15924': 13.909888957949875,
'159cm': 13.909888957949875,
'15G': 13.21674177738993,
'15GB': 13.909888957949875,
'15ISK': 13.909888957949875,
'15J': 13.909888957949875,
'15K': 10.577684447774672,
'15L': 13.909888957949875,
'15LPA': 13.909888957949875,
'15MM': 13.909888957949875,
'15R': 13.21674177738993,
'15SE': 13.909888957949875,
'15T': 13.909888957949875,
'15ab219': 13.909888957949875,
'15au101nx': 13.909888957949875,
'15bn': 13.909888957949875,
'15cc': 13.909888957949875,
'15cm': 13.50442384984171,
'15dayes': 13.909888957949875,
'15days': 12.99359822607572,
'15ft': 13.909888957949875,
'15hours': 13.909888957949875,
'15iBD': 13.909888957949875,
'15k': 8.803943484049295,
'15kVA': 13.909888957949875,
'15kg': 12.99359822607572,
'15kgs': 12.300451045515775,
'15km': 13.21674177738993,
'15m': 12.811276669281765,
'15mg': 12.038086781048284,
'15mm': 13.909888957949875,
'15s': 13.50442384984171,
```

```
'15th': 11.047688077020407,
'15x': 13.50442384984171,
'15y': 13.909888957949875,
'15years': 13.21674177738993,
'15yo': 13.21674177738993,
'15yr': 13.909888957949875,
'15yrs': 13.909888957949875,
'16': 7.622030397788091,
'160': 10.01806865983925,
'1600': 11.511993685151504,
'16000': 12.523594596829986,
'1600m': 13.50442384984171,
'1600mhz': 13.909888957949875,
'1600s': 12.99359822607572,
'1604': 13.50442384984171,
'1607': 13.21674177738993,
'160GB': 13.909888957949875,
'160Q': 13.21674177738993,
'160R': 12.038086781048284,
'160V': 13.909888957949875,
'160cm': 13.909888957949875,
'160k': 13.50442384984171,
'160lbs': 13.21674177738993,
'160m': 13.50442384984171,
'160mg': 13.909888957949875,
'160q': 13.909888957949875,
'160r': 12.657125989454507,
'161': 11.894985937407611,
'1611': 11.46754192258067,
'1616': 13.909888957949875,
'161998': 13.50442384984171,
'161Q': 13.909888957949875,
'162': 12.20514086571145,
'16201': 13.909888957949875,
'163': 12.300451045515775,
'16315': 13.909888957949875,
'164': 12.657125989454507,
'1640': 13.909888957949875,
'16400': 13.21674177738993,
'1644': 13.909888957949875,
'164595': 13.21674177738993,
'165': 11.30719927250549,
'1650': 13.50442384984171,
'165k': 13.909888957949875,
'166': 12.657125989454507,
'167': 12.657125989454507,
'167Q': 13.50442384984171,
'167x': 13.909888957949875,
'168': 11.30719927250549,
'1680': 13.909888957949875,
'16800mAh': 13.909888957949875,
'1680x1050': 13.909888957949875,
'1687': 11.424982308161875,
'1688': 13.21674177738993,
'1689': 13.909888957949875,
'168Q': 13.909888957949875,
'169': 12.657125989454507,
'16949': 13.909888957949875,
'1696': 13.909888957949875,
'16A': 13.50442384984171,
'16GB': 11.047688077020407,
'16K': 13.50442384984171,
'16MP': 13.909888957949875,
'16TB': 13.909888957949875,
'16g': 13.50442384984171,
'16gb': 12.11812948872182,
'16k': 12.405811561173602,
'16kgs': 13.21674177738993,
'16lpa': 13.50442384984171,
'16min': 13.909888957949875,
'16mm': 13.909888957949875,
'16mol': 13.909888957949875,
'16oz': 13.909888957949875,
'16t': 13.909888957949875,
'16tg': 13.909888957949875,
'16th': 11.019517200053711,
'16volts': 13.909888957949875,
```

```
 '16x': 13.909888957949875,
 '16x7': 13.909888957949875,
 '16yo': 13.909888957949875,
 '16yrs': 13.21674177738993,
 '17': 7.960549337142565,
 '170': 11.46754192258067,
 '1700': 11.019517200053711,
 '17000': 13.909888957949875,
 '1700s': 13.50442384984171,
 '1700th': 13.909888957949875,
 '17026': 13.50442384984171,
 '1705': 13.909888957949875,
 '1706': 13.909888957949875,
 '170cm': 13.909888957949875,
 '170k': 13.909888957949875,
 '170lbs': 13.909888957949875,
 '171': 13.50442384984171,
 '1713': 13.909888957949875,
 '172': 12.20514086571145,
 '1729': 13.50442384984171,
 '173': 12.300451045515775,
 '17305': 13.909888957949875,
 '173cm': 12.99359822607572,
 '174': 13.50442384984171,
 '1745': 13.50442384984171,
 '1746': 11.46754192258067,
 '174cm': 13.50442384984171,
 '175': 11.558513700786397,
 '1750': 12.99359822607572,
 '17500': 13.909888957949875,
 '1751': 11.424982308161875,
 '1752': 13.909888957949875,
 '17535': 13.50442384984171,
 '1755': 11.46754192258067,
 '17580364': 13.909888957949875,
 '175cm': 13.50442384984171,
 '175k': 12.99359822607572,
 '175lbs': 13.909888957949875,
 '176': 12.523594596829986,
 '1761': 13.909888957949875,
 '1763': 13.21674177738993,
 '1764': 13.909888957949875,
 '176cm': 13.909888957949875,
 '177': 12.657125989454507,
 '1772': 13.909888957949875,
 '1775': 13.909888957949875,
 '1776': 12.811276669281765,
 '1777': 13.909888957949875,
 '1778': 13.50442384984171,
 '178': 12.405811561173602,
 '1780': 13.909888957949875,
 '17800': 13.909888957949875,
 '1783': 13.909888957949875,
 '1785': 13.50442384984171,
 '1787': 11.424982308161875,
 '1789': 12.657125989454507,
 '178cm': 13.21674177738993,
 '179': 12.657125989454507,
 '1793': 13.50442384984171,
 '1799': 13.909888957949875,
 '179cm': 13.909888957949875,
 '17K': 13.909888957949875,
 '17L': 13.909888957949875,
 '17M': 13.909888957949875,
 '17PS': 13.909888957949875,
 '17cm': 13.909888957949875,
 '17episodes': 13.909888957949875,
 '17feet': 13.909888957949875,
 '17k': 12.811276669281765,
 ...}
```

In [ ]:

```
nlp = spacy.load('en_core_web_sm')

vecs1 = []
```

```
# https://github.com/noamraph/tqdm
# tqdm is used to print the progress bar
for qu1 in tqdm(list(df['question1'])):
    doc1 = nlp(qu1)
    # 384 is the number of dimensions of vectors
    mean_vec1 = np.zeros([len(doc1), 384])
    for word1 in doc1:
        # word2vec
        vec1 = word1.vector
        # fetch df score
        try:
            idf = word2tfidf[str(word1)]
        except:
            idf = 0
        # compute final vec
        mean_vec1 += vec1 * idf
    mean_vec1 = mean_vec1.mean(axis=0)
    vecs1.append(mean_vec1)
df['q1_feats_m'] = list(vecs1)
```

In [ ]:

```
vecs2 = []
for qu2 in tqdm(list(df['question2'])):
    doc2 = nlp(qu2)
    mean_vec2 = np.zeros([len(doc2), 384])
    for word2 in doc2:
        # word2vec
        vec2 = word2.vector
        # fetch df score
        try:
            idf = word2tfidf[str(word2)]
        except:
            #print word
            idf = 0
        # compute final vec
        mean_vec2 += vec2 * idf
    mean_vec2 = mean_vec2.mean(axis=0)
    vecs2.append(mean_vec2)
df['q2_feats_m'] = list(vecs2)
```

In [46]:

```
#prepro_features_train.csv (Simple Preprocessing Feartures)
#nlp_features_train.csv (NLP Features)
if os.path.isfile('nlp_features_train.csv'):
    dfnlp = pd.read_csv("nlp_features_train.csv",encoding='latin-1')
else:
    print("download nlp_features_train.csv from drive or run previous notebook")

if os.path.isfile('df_fe_without_preprocessing_train.csv'):
    dfppro = pd.read_csv("df_fe_without_preprocessing_train.csv",encoding='latin-1')
else:
    print("download df_fe_without_preprocessing_train.csv from drive or run previous notebook")
```

In [69]:

```
dfnlp.head()
```

Out[69]:

	id	qid1	qid2	question1	question2	is_duplicate	cwc_min	cwc_max	csc_min	csc_max	...	ctc_max	last_word_eq	f
0	0	1	2	what is the step by step guide to invest in sh...	what is the step by step guide to invest in sh...	0	0.999980	0.833319	0.999983	0.999983	...	0.785709	0.0	1
				what is the story	what would									

	id	qid1	qid2	question1	question2	is_duplicate	cwc_min	cwc_max	csc_min	csc_max	...	ctc_max	last_word_eq	f
1	1	3	4	kohinoor koh i noor dia...	happen if the indian government sto...	0	0.799984	0.399996	0.749981	0.599988	...	0.466664	0.0	1
2	2	5	6	how can i increase the speed of my internet co...	how can internet speed be increased by hacking...	0	0.399992	0.333328	0.399992	0.249997	...	0.285712	0.0	1
3	3	7	8	why am i mentally very lonely how can i solve...	find the remainder when math 23 24 math i...	0	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.0	0
4	4	9	10	which one dissolve in water quikly sugar salt...	which fish would survive in salt water	0	0.399992	0.199998	0.999950	0.666644	...	0.307690	0.0	1

5 rows × 21 columns

In [71]:

```
df1 = dfnlp.drop(['qid1','qid2','question1','question2'],axis=1)
df2  = dfppro.drop(['qid1','qid2','question1','question2','is_duplicate'],axis=1)
df3 = df.drop(['qid1','qid2','question1','question2','is_duplicate'] , axis=1)
#df3_q1=pd.DataFrame(df3.q1_feats_values.tolist(),index=df3.index)
#df3_q2=pd.DataFrame(df3.q2_feats_values.tolist(),index=df3.index)
```

In [72]:

```
df1.head()
```

Out[72]:

	id	is_duplicate	cwc_min	cwc_max	csc_min	csc_max	ctc_min	ctc_max	last_word_eq	first_word_eq	abs_len_diff	l
0	0	0	0.999980	0.833319	0.999983	0.999983	0.916659	0.785709	0.0	1.0	2.0	
1	1	0	0.799984	0.399996	0.749981	0.599988	0.699993	0.466664	0.0	1.0	5.0	
2	2	0	0.399992	0.333328	0.399992	0.249997	0.399996	0.285712	0.0	1.0	4.0	
3	3	0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.0	2.0	
4	4	0	0.399992	0.199998	0.999950	0.666644	0.571420	0.307690	0.0	1.0	6.0	

In [73]:

```
df2.head()
```

Out[73]:

	id	freq_id1	freq_id2	q1len	q2len	q1_n_words	q2_n_words	Word_common	word_Total	word_share	freq_qid1	freq_
0	0	1	1	66	57	14	12	10.0	23.0	0.434783	1	1
1	1	4	1	51	88	8	13	4.0	20.0	0.200000	4	1
2	2	1	1	73	59	14	10	4.0	24.0	0.166667	1	1
3	3	1	1	50	65	11	9	0.0	19.0	0.000000	1	1
4	4	3	1	76	39	13	7	2.0	20.0	0.100000	3	1

In [ ]:

```
df3_q1.head()
```

In [ ]:

```
df3_q2.head()
```

In [ ]:

```
print("Number of features in nlp dataframe :", df1.shape[1])
print("Number of features in preprocessed dataframe :", df2.shape[1])
print("Number of features in question1 w2v  dataframe :", df3_q1.shape[1])
print("Number of features in question2 w2v  dataframe :", df3_q2.shape[1])
print("Number of features in final dataframe  :", df1.shape[1]+df2.shape[1]+df3_q1.shape[1]+df3_q2.
shape[1])
```

In [133]:

```
if not os.path.isfile('final_features.csv'):
    df3_q1['id'] = df1['id']
    df3_q2['id'] = df1['id']
    df1=df1.merge(df2,on='id',how='left')
    df2  = df3_q1.merge(df3_q2, on='id',how='left')
    result  = df1.merge(df2, on='id',how='left')
    result.to_csv('final_features.csv')
```

In [47]:

```
import pandas as pd
import matplotlib.pyplot as plt
import re
import time
import warnings
import sqlite3
from sqlalchemy import create_engine # database connection
import csv
import os
warnings.filterwarnings("ignore")
import datetime as dt
import numpy as np
from nltk.corpus import stopwords
from sklearn.decomposition import TruncatedSVD
from sklearn.preprocessing import normalize
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.manifold import TSNE
import seaborn as sns
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix
from sklearn.metrics.classification import accuracy_score, log_loss
from sklearn.feature_extraction.text import TfidfVectorizer
from collections import Counter
from scipy.sparse import hstack
from sklearn.multiclass import OneVsRestClassifier
from sklearn.svm import SVC
from sklearn.cross_validation import StratifiedKFold
from collections import Counter, defaultdict
from sklearn.calibration import CalibratedClassifierCV
from sklearn.naive_bayes import MultinomialNB
from sklearn.naive_bayes import GaussianNB
from sklearn.model_selection import train_test_split
from sklearn.model_selection import GridSearchCV
import math
from sklearn.metrics import normalized_mutual_info_score
from sklearn.ensemble import RandomForestClassifier



from sklearn.model_selection import cross_val_score
from sklearn.linear_model import SGDClassifier
```

```
from sklearn.linear_model import SGDClassifier
from mlxtend.classifier import StackingClassifier

from sklearn import model_selection
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import precision_recall_curve, auc, roc_curve
```

In [2]:

```
!pip install mlxtend
```

```
Collecting mlxtend
  Downloading
https://files.pythonhosted.org/packages/44/d1/1b9e85e991f836e9aaea18367ff628a6324af1005971dc9f57e51
5a4/mlxtend-0.14.0-py2.py3-none-any.whl (1.3MB)
Requirement already satisfied: scikit-learn>=0.18 in c:\users\lenovo\anaconda3\lib\site-packages
(from mlxtend) (0.19.1)
Requirement already satisfied: matplotlib>=1.5.1 in c:\users\lenovo\anaconda3\lib\site-packages
(from mlxtend) (2.2.2)
Requirement already satisfied: pandas>=0.17.1 in c:\users\lenovo\anaconda3\lib\site-packages (from
mlxtend) (0.23.0)
Requirement already satisfied: setuptools in c:\users\lenovo\anaconda3\lib\site-packages (from
mlxtend) (39.1.0)
Requirement already satisfied: scipy>=0.17 in c:\users\lenovo\anaconda3\lib\site-packages (from
mlxtend) (1.1.0)
Requirement already satisfied: numpy>=1.10.4 in c:\users\lenovo\anaconda3\lib\site-packages (from
mlxtend) (1.14.3)
Requirement already satisfied: cycler>=0.10 in c:\users\lenovo\anaconda3\lib\site-packages (from
matplotlib>=1.5.1->mlxtend) (0.10.0)
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.1 in
c:\users\lenovo\anaconda3\lib\site-packages (from matplotlib>=1.5.1->mlxtend) (2.2.0)
Requirement already satisfied: python-dateutil>=2.1 in c:\users\lenovo\anaconda3\lib\site-packages
(from matplotlib>=1.5.1->mlxtend) (2.7.3)
Requirement already satisfied: pytz in c:\users\lenovo\anaconda3\lib\site-packages (from
matplotlib>=1.5.1->mlxtend) (2018.4)
Requirement already satisfied: six>=1.10 in c:\users\lenovo\anaconda3\lib\site-packages (from
matplotlib>=1.5.1->mlxtend) (1.11.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\lenovo\anaconda3\lib\site-packages
(from matplotlib>=1.5.1->mlxtend) (1.0.1)
Installing collected packages: mlxtend
Successfully installed mlxtend-0.14.0
```

```
distributed 1.21.8 requires msgpack, which is not installed.
You are using pip version 10.0.1, however version 18.1 is available.
You should consider upgrading via the 'python -m pip install --upgrade pip' command.
```

In [ ]:

```
#Creating db file from csv
if not os.path.isfile('train.db'):
    disk_engine = create_engine('sqlite:///train.db')
    start = dt.datetime.now()
    chunksize = 180000
    j = 0
    index_start = 1
    for df in pd.read_csv('final_features.csv', names=['Unnamed: 0','id','is_duplicate','cwc_min','
cwc_max','csc_min','csc_max','ctc_min','ctc_max','last_word_eq','first_word_eq','abs_len_diff','me
an_len','token_set_ratio','token_sort_ratio','fuzz_ratio','fuzz_partial_ratio','longest_substr_rati
o','freq_qid1','freq_qid2','q1len','q2len','q1_n_words','q2_n_words','word_Common','word_Total','w
ord_share','freq_q1+q2','freq_q1-
q2','0_x','1_x','2_x','3_x','4_x','5_x','6_x','7_x','8_x','9_x','10_x','11_x','12_x','13_x','14_x',
'15_x','16_x','17_x','18_x','19_x','20_x','21_x','22_x','23_x','24_x','25_x','26_x','27_x','28_x','
29_x','30_x','31_x','32_x','33_x','34_x','35_x','36_x','37_x','38_x','39_x','40_x','41_x','42_x','4
3_x','44_x','45_x','46_x','47_x','48_x','49_x','50_x','51_x','52_x','53_x','54_x','55_x','56_x','57
_x','58_x','59_x','60_x','61_x','62_x','63_x','64_x','65_x','66_x','67_x','68_x','69_x','70_x','71_
x','72_x','73_x','74_x','75_x','76_x','77_x','78_x','79_x','80_x','81_x','82_x','83_x','84_x','85_x
','86_x','87_x','88_x','89_x','90_x','91_x','92_x','93_x','94_x','95_x','96_x','97_x','98_x','99_x'
,'100_x','101_x','102_x','103_x','104_x','105_x','106_x','107_x','108_x','109_x','110_x','111_x','
112_x','113_x','114_x','115_x','116_x','117_x','118_x','119_x','120_x','121_x','122_x','123_x','12
4_x','125_x','126_x','127_x','128_x','129_x','130_x','131_x','132_x','133_x','134_x','135_x','136_
x','137_x','138_x','139_x','140_x','141_x','142_x','143_x','144_x','145_x','146_x','147_x','148_x'
,'149_x','150_x','151_x','152_x','153_x','154_x','155_x','156_x','157_x','158_x','159_x','160_x','
161_x','162_x','163_x','164_x','165_x','166_x','167_x','168_x','169_x','170_x','171_x','172_x','17
3_x','174_x','175_x','176_x','177_x','178_x','179_x','180_x','181_x','182_x','183_x','184_x','185_
```

```
x','186_x','187_x','188_x','189_x','190_x','191_x','192_x','193_x','194_x','195_x','196_x','197_x'
,'198_x','199_x','200_x','201_x','202_x','203_x','204_x','205_x','206_x','207_x','208_x','209_x','
210_x','211_x','212_x','213_x','214_x','215_x','216_x','217_x','218_x','219_x','220_x','221_x','22
2_x','223_x','224_x','225_x','226_x','227_x','228_x','229_x','230_x','231_x','232_x','233_x','234_
x','235_x','236_x','237_x','238_x','239_x','240_x','241_x','242_x','243_x','244_x','245_x','246_x'
,'247_x','248_x','249_x','250_x','251_x','252_x','253_x','254_x','255_x','256_x','257_x','258_x','
259_x','260_x','261_x','262_x','263_x','264_x','265_x','266_x','267_x','268_x','269_x','270_x','27
1_x','272_x','273_x','274_x','275_x','276_x','277_x','278_x','279_x','280_x','281_x','282_x','283_
x','284_x','285_x','286_x','287_x','288_x','289_x','290_x','291_x','292_x','293_x','294_x','295_x'
,'296_x','297_x','298_x','299_x','300_x','301_x','302_x','303_x','304_x','305_x','306_x','307_x','
308_x','309_x','310_x','311_x','312_x','313_x','314_x','315_x','316_x','317_x','318_x','319_x','32
0_x','321_x','322_x','323_x','324_x','325_x','326_x','327_x','328_x','329_x','330_x','331_x','332_
x','333_x','334_x','335_x','336_x','337_x','338_x','339_x','340_x','341_x','342_x','343_x','344_x'
,'345_x','346_x','347_x','348_x','349_x','350_x','351_x','352_x','353_x','354_x','355_x','356_x','
357_x','358_x','359_x','360_x','361_x','362_x','363_x','364_x','365_x','366_x','367_x','368_x','36
9_x','370_x','371_x','372_x','373_x','374_x','375_x','376_x','377_x','378_x','379_x','380_x','381_
x','382_x','383_x','0_y','1_y','2_y','3_y','4_y','5_y','6_y','7_y','8_y','9_y','10_y','11_y','12_y'
,'13_y','14_y','15_y','16_y','17_y','18_y','19_y','20_y','21_y','22_y','23_y','24_y','25_y','26_y',
'27_y','28_y','29_y','30_y','31_y','32_y','33_y','34_y','35_y','36_y','37_y','38_y','39_y','40_y','
41_y','42_y','43_y','44_y','45_y','46_y','47_y','48_y','49_y','50_y','51_y','52_y','53_y','54_y','5
5_y','56_y','57_y','58_y','59_y','60_y','61_y','62_y','63_y','64_y','65_y','66_y','67_y','68_y','69
_y','70_y','71_y','72_y','73_y','74_y','75_y','76_y','77_y','78_y','79_y','80_y','81_y','82_y','83_
y','84_y','85_y','86_y','87_y','88_y','89_y','90_y','91_y','92_y','93_y','94_y','95_y','96_y','97_y
','98_y','99_y','100_y','101_y','102_y','103_y','104_y','105_y','106_y','107_y','108_y','109_y','11
0_y','111_y','112_y','113_y','114_y','115_y','116_y','117_y','118_y','119_y','120_y','121_y','122_
y','123_y','124_y','125_y','126_y','127_y','128_y','129_y','130_y','131_y','132_y','133_y','134_y'
,'135_y','136_y','137_y','138_y','139_y','140_y','141_y','142_y','143_y','144_y','145_y','146_y','
147_y','148_y','149_y','150_y','151_y','152_y','153_y','154_y','155_y','156_y','157_y','158_y','15
9_y','160_y','161_y','162_y','163_y','164_y','165_y','166_y','167_y','168_y','169_y','170_y','171_
y','172_y','173_y','174_y','175_y','176_y','177_y','178_y','179_y','180_y','181_y','182_y','183_y'
,'184_y','185_y','186_y','187_y','188_y','189_y','190_y','191_y','192_y','193_y','194_y','195_y','
196_y','197_y','198_y','199_y','200_y','201_y','202_y','203_y','204_y','205_y','206_y','207_y','20
8_y','209_y','210_y','211_y','212_y','213_y','214_y','215_y','216_y','217_y','218_y','219_y','220_
y','221_y','222_y','223_y','224_y','225_y','226_y','227_y','228_y','229_y','230_y','231_y','232_y'
,'233_y','234_y','235_y','236_y','237_y','238_y','239_y','240_y','241_y','242_y','243_y','244_y','
245_y','246_y','247_y','248_y','249_y','250_y','251_y','252_y','253_y','254_y','255_y','256_y','25
7_y','258_y','259_y','260_y','261_y','262_y','263_y','264_y','265_y','266_y','267_y','268_y','269_
y','270_y','271_y','272_y','273_y','274_y','275_y','276_y','277_y','278_y','279_y','280_y','281_y'
,'282_y','283_y','284_y','285_y','286_y','287_y','288_y','289_y','290_y','291_y','292_y','293_y','
294_y','295_y','296_y','297_y','298_y','299_y','300_y','301_y','302_y','303_y','304_y','305_y','30
6_y','307_y','308_y','309_y','310_y','311_y','312_y','313_y','314_y','315_y','316_y','317_y','318_
y','319_y','320_y','321_y','322_y','323_y','324_y','325_y','326_y','327_y','328_y','329_y','330_y'
,'331_y','332_y','333_y','334_y','335_y','336_y','337_y','338_y','339_y','340_y','341_y','342_y','
343_y','344_y','345_y','346_y','347_y','348_y','349_y','350_y','351_y','352_y','353_y','354_y','35
5_y','356_y','357_y','358_y','359_y','360_y','361_y','362_y','363_y','364_y','365_y','366_y','367_
y','368_y','369_y','370_y','371_y','372_y','373_y','374_y','375_y','376_y','377_y','378_y','379_y'
,'380_y','381_y','382_y','383_y'], chunksize=chunksize, iterator=True, encoding='utf-8', ):
        df.index += index_start
        j+=1
        print('{} rows'.format(j*chunksize))
        df.to_sql('data', disk_engine, if_exists='append')
        index_start = df.index[-1] + 1
```

```
180000 rows
```

In [48]:

```
def create_connection(db_file):
    try:

        conn = sqlite3.connect(db_file)
        return conn
    except Error as e:
        print(e)
    return None
def checkTableExists(dbcon):
    cursr = dbcon.cursor()
    str = "select name from sqlite_master where type='table'"
    table_names = dbcon.execute(str)
    print("Tables in the databse:")
    tables =table_names.fetchall()
    print(tables[0][0])
    return (len(tables))
```

In [49]:

```
read_db = 'train.db'
conn_r = create_connection(read_db)
checkTableExists(conn_r)
conn_r.close()
```

```
Tables in the databse:
data
```

In [ ]:

```
if os.path.isfile(read_db):
    conn_r = create_connection(read_db)
    if conn_r is not None:
        # for selecting first 1M rows
        # data = pd.read_sql_query("""SELECT * FROM data LIMIT 100001;""", conn_r)

        # for selecting random points
        data = pd.read_sql_query("SELECT * From data ORDER BY RANDOM() LIMIT 100001;", conn_r)
        conn_r.commit()
        conn_r.close()
```

In [ ]:

```
data.drop(data.index[0],inplace=True)
y_true = data['is_duplicate']
data.drop(['Unnamed:0','id','index','is_duplicate'],axis=1,inplace=True)
```

In [ ]:

```
data.head()
```

In [ ]:

```
cols = list(data.columns)
for i in cols:
    data[i]=data[i].apply(pd.to_numeric)
    print(i)
```

In [ ]:

```
y_true = list(map(int,y_true.values))
```

RANDOM TRAIN SPLIT

In [ ]:

```
X_train,X_test,y_train,y_test = train_test_split(data,y_true,stratify = True,test_size=0.3)
print("Number of data points in train data :",X_train.shape)
print("Number of data points in test data :",X_test.shape)
```

In [ ]:

```
print("-"*10,"Distribution of output variable in train data","-"*10)
train_distr = Counter(y_train)
train_len = len(y_train)
print("Class 0: ",int(train_distr[0])/train_len,"Class 1: ", int(train_distr[1])/train_len)
print("-"*10, "Distribution of output variable in test data", "-"*10)
test_distr = Counter(y_test)
test_len = len(y_test)
print("Class 0: ",int(test_distr[0])/test_len, "Class 1: ",int(test_distr[1])/test_len)
```

In [ ]:

```
def plot_confusion_matrix(test_y,predict_y):
    C= confusion_matrix(test_y,predict_y)
    A=(((C.T)/C.sum(axis=1))).T)
    B=(C/C.sum(axis=0))
    plt.figure(figsize=(20,4))

    labels = [1,2]
    cmap=sns.light_palette("blue")
    plt.subplot(1, 3, 1)
    plt.subplot(1, 3, 1)
    sns.heatmap(C, annot=True, cmap=cmap, fmt=".3f", xticklabels=labels, yticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.title("Confusion matrix")
    plt.subplot(1, 3, 2)
    sns.heatmap(B, annot=True, cmap=cmap, fmt=".3f", xticklabels=labels, yticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.title("Precision matrix")
    plt.subplot(1, 3, 3)
    # representing B in heatmap format
    sns.heatmap(A, annot=True, cmap=cmap, fmt=".3f", xticklabels=labels, yticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.title("Recall matrix")

    plt.show()
```

In [ ]:

```
predicted_y = np.zeros((test_len,2))
for i in range(test_len):
    rand_probs = np.random.rand(1,2)
    predicted_y[i]=((rand_probs/sum(sum(rand_probs)))[0])
print("log loss on test Data using random model",log_loss(y_test,predicted_y,eps=1e-15))

predicted_y =np.argmax(predicted_y, axis=1)
plot_confusion_matrix(y_test, predicted_y)
```

In [ ]:

```
alpha = [10 ** x for x in range (-5,2)]
log_error_array = []
for i in alpha:
    clf = SGDClassifier(loss='log',penalty='l2',alpha=i,random_state = 42)
    clf.fit(X_train,y_train)
    sig_clf = CalibratedClassifierCV(clf,method='sigmoid')
    sig_clf.fit(X_train,y_train)
    predicted_y = sig_clf.predicted_proba(X_test)
    log_error_array.append(log_loss(y_test,predicted_y,labels = clf.classes_,eps=1e-15))
    print('For values of alpha =',i,"the log loss is ",log_loss(y_test,predict_y,labels = clf.class
es_,eps=1e-15))
fig,ax = plt.subplots()
ax.plot(alpha,log_error_array,c='g')
for i,txt in enumerate(np.round(log_error_array,3)):
    ax.annotate((alpha[i],np.round(txt,3)),(alpha[i],log_error_array[i]))
plt.grid()
plt.title("Cross validation Error for each alpha")
plt.xlabel("ALPHA IS")
plt.ylabel("ERROR MEASURE")
plt.show()
best_alpha = np.argmin(log_error_array)
clf = SGDClassifier(alpha = alpha[best_alpha],penalty = 'l2',loss = 'log',random_state=42)
clf.fit(X_train,y_train)
sig_clf = CalibratedClassifierCV(clf,method = 'sigmoid')
sig_clf.fit(X_train,y_train)
predict_y = sig_clf.predict_proba(X_train)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train,
predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(X_test)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, p
redict_y, labels=clf.classes_, eps=1e-15))
predicted_y =np.argmax(predict_y,axis=1)
```

```
predicted_y =np.argmax(predict_y,axis=1)
print("Total number of data points :", len(predicted_y))
plot_confusion_matrix(y_test, predicted_y)
```

In [1]:

```
alpha = [10 ** x for x in range(-5, 2)] # hyperparam for SGD classifier.

# read more about SGDClassifier() at http://scikit-
learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# ------------------------------
# default parameters
# SGDClassifier(loss='hinge', penalty='l2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_i
ter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0
=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...]) Fit linear model with Stochastic Gradient Descent.
# predict(X) Predict class labels for samples in X.

#--------------------------------
# video link:
#--------------------------------


log_error_array=[]
for i in alpha:
    clf = SGDClassifier(alpha=i, penalty='l1', loss='hinge', random_state=42)
    clf.fit(X_train, y_train)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(X_train, y_train)
    predict_y = sig_clf.predict_proba(X_test)
    log_error_array.append(log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
    print('For values of alpha = ', i, "The log loss is:",log_loss(y_test, predict_y, labels=clf.cl
asses_, eps=1e-15))

fig, ax = plt.subplots()
ax.plot(alpha, log_error_array,c='g')
for i, txt in enumerate(np.round(log_error_array,3)):
    ax.annotate((alpha[i],np.round(txt,3)), (alpha[i],log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()


best_alpha = np.argmin(log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l1', loss='hinge', random_state=42)
clf.fit(X_train, y_train)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(X_train, y_train)

predict_y = sig_clf.predict_proba(X_train)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train,
predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(X_test)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, p
redict_y, labels=clf.classes_, eps=1e-15))
predicted_y =np.argmax(predict_y,axis=1)
print("Total number of data points :", len(predicted_y))
plot_confusion_matrix(y_test, predicted_y)
```

Out[1]:

```
(3, 2)
```

XGBOOST

In [ ]:

```
import xgboost as xgb
params = {}
params['objective'] = 'binary:logistic'
params['eval_metric'] = 'logloss
params['eta']   = 0.02
params['max_depth'] = 4
d_train = xgb.DMatrix(X_train,label=y_train)
d_test = xgb.DMatrix(X_test,label=y_test)
watchlist = [(d_train,'train'),(d_test,'valid')]
bst = xgb.train(params,d_train,400,watchlist,early_stopping_rounds =20,verbose_equal=10 )
xgdmat = xgb.DMatrix(X_train,y_train)
predict_y  = bst.predict(d_test)
print("The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

In [ ]:

```
predicted_y =np.array(predict_y>0.5,dtype=int)
print("Total number of data points :", len(predicted_y))
plot_confusion_matrix(y_test, predicted_y)
```

USING ONLY TFIDF VECTORIZER

In [76]:

```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.feature_extraction.text import CountVectorizer
# merge texts
questions = list(df['question1']) + list(df['question2'])

tfidf = TfidfVectorizer(lowercase=False, )
#tfidf.fit_transform(questions)

#word2tfidf = dict(zip(tfidf.get_feature_names(), tfidf.idf_))
```

In [78]:

```
q1_tfidf_feat = tfidf.fit_transform(df['question1'])
q2_tfidf_feat = tfidf.fit_transform(df['question2'])
```

In [79]:

```
q1_tfidf_feat
```

Out[79]:

```
<404290x84717 sparse matrix of type '<class 'numpy.float64'>'
 with 4045225 stored elements in Compressed Sparse Row format>
```

In [80]:

```
q2_tfidf_feat
```

Out[80]:

```
<404290x78351 sparse matrix of type '<class 'numpy.float64'>'
 with 4101330 stored elements in Compressed Sparse Row format>
```

In [81]:

```
df1 = dfnlp.drop(['qid1','qid2','question1','question2'],axis=1)
df2 = dfppro.drop(['qid1','qid2','question1','question2','is_duplicate'],axis=1)
df3 = df.drop(['qid1','qid2','question1','question2','is_duplicate'],axis=1)
```

In [82]:

```
df1.head()
```

Out[82]:

	id	is_duplicate	cwc_min	cwc_max	csc_min	csc_max	ctc_min	ctc_max	last_word_eq	first_word_eq	abs_len_diff
0	0	0	0.999980	0.833319	0.999983	0.999983	0.916659	0.785709	0.0	1.0	2.0
1	1	0	0.799984	0.399996	0.749981	0.599988	0.699993	0.466664	0.0	1.0	5.0
2	2	0	0.399992	0.333328	0.399992	0.249997	0.399996	0.285712	0.0	1.0	4.0
3	3	0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.0	2.0
4	4	0	0.399992	0.199998	0.999950	0.666644	0.571420	0.307690	0.0	1.0	6.0

In [83]:

```
df2.head()
```

Out[83]:

	id	freq_id1	freq_id2	q1len	q2len	q1_n_words	q2_n_words	Word_common	word_Total	word_share	freq_qid1	freq_
0	0	1	1	66	57	14	12	10.0	23.0	0.434783	1	1
1	1	4	1	51	88	8	13	4.0	20.0	0.200000	4	1
2	2	1	1	73	59	14	10	4.0	24.0	0.166667	1	1
3	3	1	1	50	65	11	9	0.0	19.0	0.000000	1	1
4	4	3	1	76	39	13	7	2.0	20.0	0.100000	3	1

In [84]:

```
result  = df1.merge(df2, on='id',how='left')
```

In [85]:

```
result.head()
```

Out[85]:

	id	is_duplicate	cwc_min	cwc_max	csc_min	csc_max	ctc_min	ctc_max	last_word_eq	first_word_eq	...	q2len	q1_
0	0	0	0.999980	0.833319	0.999983	0.999983	0.916659	0.785709	0.0	1.0	...	57	14
1	1	0	0.799984	0.399996	0.749981	0.599988	0.699993	0.466664	0.0	1.0	...	88	8
2	2	0	0.399992	0.333328	0.399992	0.249997	0.399996	0.285712	0.0	1.0	...	59	14
3	3	0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.0	...	65	11
4	4	0	0.399992	0.199998	0.999950	0.666644	0.571420	0.307690	0.0	1.0	...	39	13

5 rows × 30 columns

In [86]:

```
if not os.path.isfile('final_features_latest.csv'):


    result1  = df1.merge(df2, on='id',how='left')
    result1.to_csv('final_features_latest.csv')
```

In [87]:

```
pretfidf_features = pd.read_csv('final_features_latest.csv')
```

In [88]:

```
pretfidf_features.head(2)
```

Out[88]:

	Unnamed: 0	id	is_duplicate	cwc_min	cwc_max	csc_min	csc_max	ctc_min	ctc_max	last_word_eq	...	q2len	q1_n_w
0	0	0	0	0.999980	0.833319	0.999983	0.999983	0.916659	0.785709	0.0	...	57	14
1	1	1	0	0.799984	0.399996	0.749981	0.599988	0.699993	0.466664	0.0	...	88	8

2 rows × 31 columns

In [94]:

```
pretfidf_features.drop(['Unnamed: 0','id','is_duplicate'],axis=1,inplace=True)
```

In [95]:

```
pretfidf_features.head(2)
```

Out[95]:

	cwc_min	cwc_max	csc_min	csc_max	ctc_min	ctc_max	last_word_eq	first_word_eq	abs_len_diff	mean_len	...	q2l
0	0.999980	0.833319	0.999983	0.999983	0.916659	0.785709	0.0	1.0	2.0	13.0	...	57
1	0.799984	0.399996	0.749981	0.599988	0.699993	0.466664	0.0	1.0	5.0	12.5	...	88

2 rows × 28 columns

In [96]:

```
import scipy
```

In [97]:

```
q1_tfidf_feat
```

Out[97]:

```
<404290x84717 sparse matrix of type '<class 'numpy.float64'>'
 with 4045225 stored elements in Compressed Sparse Row format>
```

In [98]:

```
final_questions =  scipy.sparse.hstack((q1_tfidf_feat,q2_tfidf_feat))
```

In [99]:

```
final_new_feat = scipy.sparse.hstack((pretfidf_features,final_questions))
```

In [100]:

```
#http://www.sqlitetutorial.net/sqlite-python/create-tables/
def create_connection(db_file):
    """ create a database connection to the SQLite database
        specified by db_file
    :param db_file: database file
    :return: Connection object or None
    """
    try:
        conn = sqlite3.connect(db_file)
        return conn
```

```
        return conn
    except Error as e:
        print(e)

    return None


def checkTableExists(dbcon):
    cursr = dbcon.cursor()
    str = "select name from sqlite_master where type='table'"
    table_names = cursr.execute(str)
    print("Tables in the databse:")
    tables =table_names.fetchall()
    print(tables[0][0])
    return(len(tables))
```

In [101]:

```
y_true = df1['is_duplicate']
y_true = list(map(int,y_true.values))
```

In [102]:

```
X_train,X_test,y_train,y_test = train_test_split(final_new_feat,y_true,test_size=0.3)
print("Number of data points in train data :",X_train.shape)
print("Number of data points in test data :",X_test.shape)
```

```
Number of data points in train data : (283003, 163096)
Number of data points in test data : (121287, 163096)
```

In [103]:

```
print("-"*10,"Distribution of output variable in train data","-"*10)
train_distr = Counter(y_train)
train_len = len(y_train)
print("Class 0: ",int(train_distr[0])/train_len,"Class 1: ", int(train_distr[1])/train_len)
print("-"*10, "Distribution of output variable in test data", "-"*10)
test_distr = Counter(y_test)
test_len = len(y_test)
print("Class 0: ",int(test_distr[0])/test_len, "Class 1: ",int(test_distr[1])/test_len)
```

```
---------- Distribution of output variable in train data ----------
Class 0:  0.6305763543142652 Class 1:  0.36942364568573477
---------- Distribution of output variable in test data ----------
Class 0:  0.6313289965124045 Class 1:  0.3686710034875955
```

In [104]:

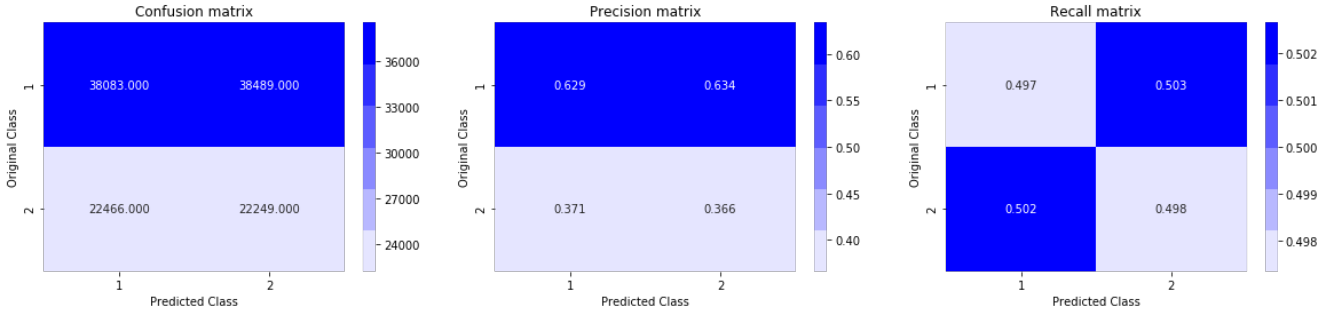
```
def plot_confusion_matrix(test_y,predict_y):
    C= confusion_matrix(test_y,predict_y)
    A=(((C.T)/C.sum(axis=1))).T
    B=(C/C.sum(axis=0))
    plt.figure(figsize=(20,4))

    labels = [1,2]
    cmap=sns.light_palette("blue")
    plt.subplot(1, 3, 1)
    plt.subplot(1, 3, 1)
    sns.heatmap(C, annot=True, cmap=cmap, fmt=".3f", xticklabels=labels, yticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.title("Confusion matrix")
    plt.subplot(1, 3, 2)
    sns.heatmap(B, annot=True, cmap=cmap, fmt=".3f", xticklabels=labels, yticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.title("Precision matrix")
    plt.subplot(1, 3, 3)
    # representing B in heatmap format
    sns.heatmap(A, annot=True, cmap=cmap, fmt=".3f", xticklabels=labels, yticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.title("Recall matrix")
```

```
    plt.title("Recall matrix")

    plt.show()
```

In [105]:

```
predicted_y = np.zeros((test_len,2))
for i in range(test_len):
    rand_probs = np.random.rand(1,2)
    predicted_y[i]=((rand_probs/sum(sum(rand_probs)))[0])
print("log loss on test Data using random model",log_loss(y_test,predicted_y,eps=1e-15))

predicted_y =np.argmax(predicted_y, axis=1)
plot_confusion_matrix(y_test, predicted_y)
```

```
log loss on test Data using random model 0.8892522827971946
```

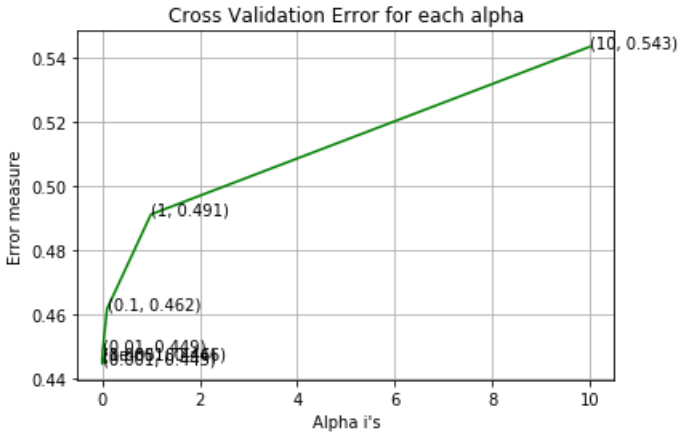


In [106]:

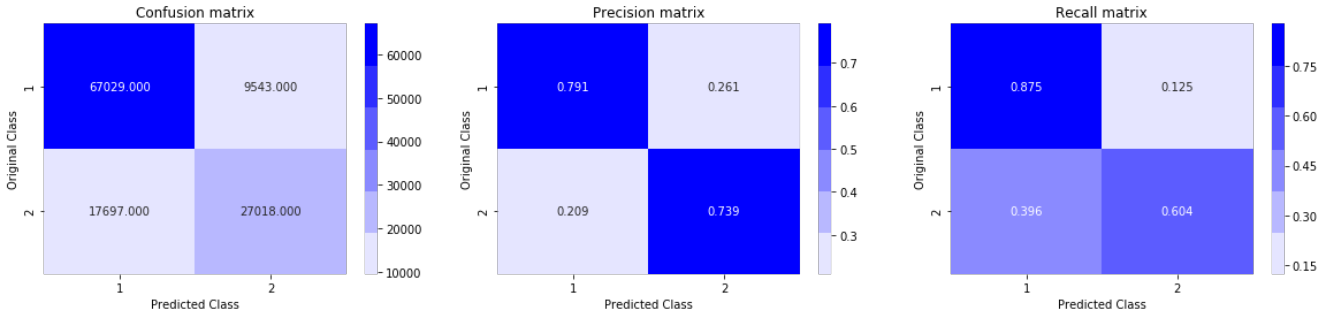
```
alpha = [10 ** x for x in range(-5, 2)] # hyperparam for SGD classifier.

# read more about SGDClassifier() at http://scikit-
learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# ------------------------------
# default parameters
# SGDClassifier(loss='hinge', penalty='l2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_i
ter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0
=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, …]) Fit linear model with Stochastic Gradient Descent.
# predict(X) Predict class labels for samples in X.

#-------------------------------
# video link:
#-------------------------------


log_error_array=[]
for i in alpha:
    clf = SGDClassifier(alpha=i, penalty='l2', loss='log', random_state=42)
    clf.fit(X_train, y_train)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(X_train, y_train)
    predict_y = sig_clf.predict_proba(X_test)
    log_error_array.append(log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
    print('For values of alpha = ', i, "The log loss is:",log_loss(y_test, predict_y, labels=clf.cl
asses_, eps=1e-15))

fig, ax = plt.subplots()
ax.plot(alpha, log_error_array,c='g')
for i, txt in enumerate(np.round(log_error_array,3)):
    ax.annotate((alpha[i],np.round(txt,3)), (alpha[i],log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()
```

```
best_alpha = np.argmin(log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
clf.fit(X_train, y_train)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(X_train, y_train)

predict_y = sig_clf.predict_proba(X_train)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train,
predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(X_test)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, p
redict_y, labels=clf.classes_, eps=1e-15))
predicted_y =np.argmax(predict_y,axis=1)
print("Total number of data points :", len(predicted_y))
plot_confusion_matrix(y_test, predicted_y)
```

```
For values of alpha =  1e-05 The log loss is: 0.44615619465082784
For values of alpha =  0.0001 The log loss is: 0.4463370362043058
For values of alpha =  0.001 The log loss is: 0.444546332298342
For values of alpha =  0.01 The log loss is: 0.4485842607801396
For values of alpha =  0.1 The log loss is: 0.4616826132639312
For values of alpha =  1 The log loss is: 0.4910739607729717
For values of alpha =  10 The log loss is: 0.5432746780586721
```




```
For values of best alpha =  0.001 The train log loss is: 0.4434305265834193
For values of best alpha =  0.001 The test log loss is: 0.444546332298342
Total number of data points : 121287
```

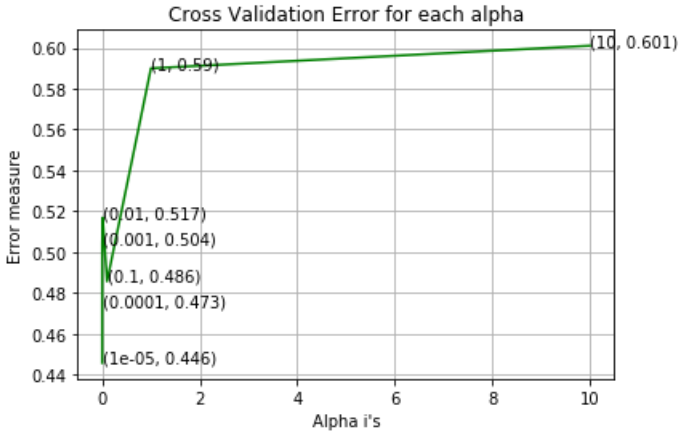


In [107]:

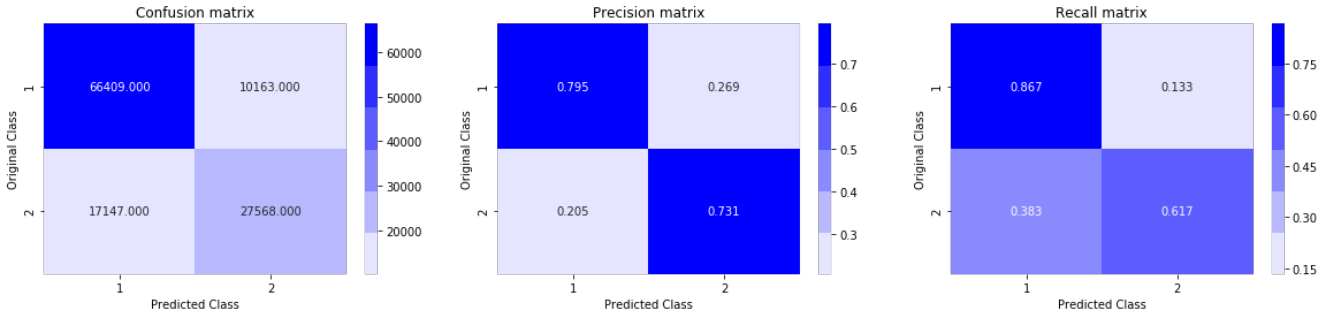
```
log_error_array=[]
for i in alpha:
    clf = SGDClassifier(alpha=i, penalty='l1', loss='hinge', random_state=42)
    clf.fit(X_train, y_train)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(X_train, y_train)
    predict_y = sig_clf.predict_proba(X_test)
    log_error_array.append(log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
    print('For values of alpha = ', i, "The log loss is:",log_loss(y_test, predict_y, labels=clf.cl
asses_, eps=1e-15))

fig, ax = plt.subplots()
```

```
fig, ax = plt.subplots()
ax.plot(alpha, log_error_array,c='g')
for i, txt in enumerate(np.round(log_error_array,3)):
    ax.annotate((alpha[i],np.round(txt,3)), (alpha[i],log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()


best_alpha = np.argmin(log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l1', loss='hinge', random_state=42)
clf.fit(X_train, y_train)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(X_train, y_train)

predict_y = sig_clf.predict_proba(X_train)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train,
predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(X_test)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, p
redict_y, labels=clf.classes_, eps=1e-15))
predicted_y =np.argmax(predict_y,axis=1)
print("Total number of data points :", len(predicted_y))
plot_confusion_matrix(y_test, predicted_y)
```

```
For values of alpha =  1e-05 The log loss is: 0.44582361201434234
For values of alpha =  0.0001 The log loss is: 0.4732786444548065
For values of alpha =  0.001 The log loss is: 0.5041098841634398
For values of alpha =  0.01 The log loss is: 0.5169188979345268
For values of alpha =  0.1 The log loss is: 0.4856891583733125
For values of alpha =  1 The log loss is: 0.5899486921528809
For values of alpha =  10 The log loss is: 0.6009966596966851
```




```
For values of best alpha =  1e-05 The train log loss is: 0.44506206860274683
For values of best alpha =  1e-05 The test log loss is: 0.44582361201434234
Total number of data points : 121287
```



XGBOOST

In [110]:

```
import xgboost as xgb
params = {}
params['objective'] = 'binary:logistic'
params['eval_metric'] = 'logloss'
params['eta']  = 0.02
params['max_depth'] = 4
d_train = xgb.DMatrix(X_train,label=y_train)
d_test = xgb.DMatrix(X_test,label=y_test)
watchlist = [(d_train,'train'),(d_test,'valid')]
bst = xgb.train(params,d_train,400,watchlist,early_stopping_rounds =20,verbose_eval=10 )
xgdmat = xgb.DMatrix(X_train,y_train)
predict_y  = bst.predict(d_test)
print("The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```
[13:58:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[0] train-logloss:0.684646 valid-logloss:0.684646
Multiple eval metrics have been passed: 'valid-logloss' will be used for early stopping.

Will train until valid-logloss hasn't improved in 20 rounds.
[13:58:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[10] train-logloss:0.615022 valid-logloss:0.615042
[13:58:46] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:53] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[20] train-logloss:0.564386 valid-logloss:0.564489
[13:58:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:57] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:58:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:01] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
```

[13:59:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:04] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:05] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[30] train-logloss:0.526002 valid-logloss:0.526226
[13:59:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:13] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[40] train-logloss:0.496746 valid-logloss:0.497039
[13:59:17] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:20] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:23] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:24] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[50] train-logloss:0.473747 valid-logloss:0.474076
[13:59:27] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[60] train-logloss:0.455539 valid-logloss:0.455925
[13:59:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,

1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[13:59:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:46] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[70] train-logloss:0.440899 valid-logloss:0.441275
[13:59:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[13:59:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[13:59:53] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[13:59:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[80] train-logloss:0.428927 valid-logloss:0.429368
[13:59:57] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[13:59:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:01] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:04] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:05] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[90] train-logloss:0.419362 valid-logloss:0.419852
[14:00:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:09] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:13] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[100] train-logloss:0.411371 valid-logloss:0.411905

[100] train-logloss:0.411571 valid-logloss:0.411965
[14:00:17] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:20] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:23] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:24] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[110] train-logloss:0.404602 valid-logloss:0.405216
[14:00:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:27] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[120] train-logloss:0.398916 valid-logloss:0.399608
[14:00:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[130] train-logloss:0.394102 valid-logloss:0.394851
[14:00:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:46] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4

1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[140] train-logloss:0.390216 valid-logloss:0.391016
[14:00:53] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:57] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:00:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:00:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:01] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[150] train-logloss:0.386857 valid-logloss:0.387704
[14:01:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:04] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:05] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:05] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:01:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:09] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[160] train-logloss:0.383917 valid-logloss:0.384807
[14:01:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:13] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:01:17] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:20] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[170] train-logloss:0.38122 valid-logloss:0.382144
[14:01:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:23] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:24] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,

[14:01:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:27] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[180] train-logloss:0.378947 valid-logloss:0.379917
[14:01:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[190] train-logloss:0.37696 valid-logloss:0.37802
[14:01:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:46] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[200] train-logloss:0.375325 valid-logloss:0.376461
[14:01:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:01:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:53] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:57] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:01:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[210] train-logloss:0.373682 valid-logloss:0.374873

```
[210] train-logloss:0.373682 valid-logloss:0.374875
[14:01:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:01] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:01] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:04] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:05] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[220] train-logloss:0.371902 valid-logloss:0.373142
[14:02:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:09] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:13] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[230] train-logloss:0.370183 valid-logloss:0.371481
[14:02:17] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:20] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:23] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:24] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[240] train-logloss:0.368587 valid-logloss:0.369941
[14:02:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:27] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:02:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:02:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
```

1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:02:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[250] train-logloss:0.3671 valid-logloss:0.368487
[14:02:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:02:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:02:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:02:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[260] train-logloss:0.365462 valid-logloss:0.366906
[14:02:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:02:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:46] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:02:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:02:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:53] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[270] train-logloss:0.364303 valid-logloss:0.365798
[14:02:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:02:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:57] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:02:57] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:02:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:02:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:01] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[280] train-logloss:0.363311 valid-logloss:0.364823
[14:03:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:04] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:05] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4

[14:03:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:03:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:03:09] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[290] train-logloss:0.362366 valid-logloss:0.363914
[14:03:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 20 extra nodes, 0 pruned nodes, max_depth=4
[14:03:13] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:03:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:17] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:03:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:03:20] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[300] train-logloss:0.361484 valid-logloss:0.36307
[14:03:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 20 extra nodes, 0 pruned nodes, max_depth=4
[14:03:23] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:24] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:27] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:03:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[310] train-logloss:0.360429 valid-logloss:0.362054
[14:03:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:03:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:03:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 20 extra nodes, 0 pruned nodes, max_depth=4
[14:03:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4

```
[320] train-logloss:0.359588 valid-logloss:0.361255
[14:03:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 20 extra nodes, 0 pruned nodes, max_depth=4
[14:03:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 20 extra nodes, 0 pruned nodes, max_depth=4
[14:03:46] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:03:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[330] train-logloss:0.358732 valid-logloss:0.360432
[14:03:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:53] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:03:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:57] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 20 extra nodes, 0 pruned nodes, max_depth=4
[340] train-logloss:0.357982 valid-logloss:0.359725
[14:03:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:03:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:03:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:04:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:01] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 20 extra nodes, 0 pruned nodes, max_depth=4
[14:04:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:04] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:05] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:04:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[350] train-logloss:0.357285 valid-logloss:0.359071
[14:04:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:04:09] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:04:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
```

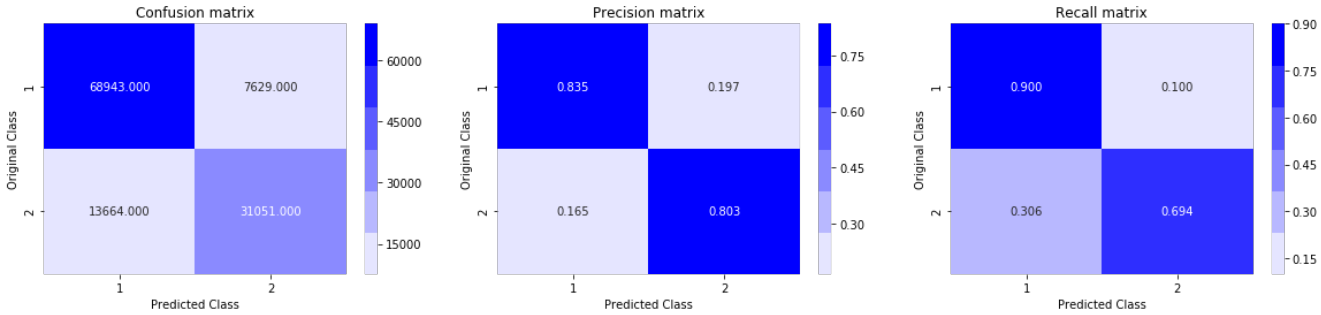
```
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:13] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:04:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 20 extra nodes, 0 pruned nodes, max_depth=4
[360] train-logloss:0.356618 valid-logloss:0.358441
[14:04:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:17] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:20] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:04:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:04:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:23] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:04:24] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[370] train-logloss:0.356003 valid-logloss:0.357869
[14:04:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 18 extra nodes, 0 pruned nodes, max_depth=4
[14:04:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:27] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:04:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:04:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:04:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[380] train-logloss:0.355223 valid-logloss:0.357122
[14:04:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:04:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 18 extra nodes, 0 pruned nodes, max_depth=4
[14:04:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 20 extra nodes, 0 pruned nodes, max_depth=4
[14:04:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:04:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 18 extra nodes, 0 pruned nodes, max_depth=4
[390] train-logloss:0.354589 valid-logloss:0.35655
[14:04:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=4
[14:04:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
```

```
[14:04:46] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=4
[14:04:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 22 extra nodes, 0 pruned nodes, max_depth=4
[14:04:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[14:04:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 18 extra nodes, 0 pruned nodes, max_depth=4
[14:04:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=4
[14:04:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=4
[399] train-logloss:0.354 valid-logloss:0.356004
The test log loss is: 0.3560038216169174
```

In [111]:

```
predicted_y =np.array(predict_y>0.5,dtype=int)
print("Total number of data points :", len(predicted_y))
plot_confusion_matrix(y_test, predicted_y)
```

```
Total number of data points : 121287
```



Random search parameter tuning for xgboost

In [ ]:

```
from sklearn.model_selection import RandomizedSearchCV
import xgboost as xgb
from xgboost import XGBClassifier
xgb1 = XGBClassifier(objective='binary:logistic')

params = {'learning_rate':[0.01,0.1,0.04,0.03,0.2,0.3],'max_depth':[3,4,5,6,7,8]}
rsv = RandomizedSearchCV(xgb1,params,verbose=0)
rsv.fit(X_train,y_train)
```

In [122]:

```
print(rsv.best_params_)
```

```
{'max_depth': 8, 'learning_rate': 0.3}
```

In [121]:

```
print(rsv.best_score_)
```

```
0.8465669975229945
```

In [123]:

```
import xgboost as xgb
params = {}
params['objective'] = 'binary:logistic'
params['eval_metric'] = 'logloss'
params['learning_rate']  = 0.3
params['max_depth'] = 8
```

```
params['max_depth'] = 8
d_train = xgb.DMatrix(X_train,label=y_train)
d_test = xgb.DMatrix(X_test,label=y_test)
watchlist = [(d_train,'train'),(d_test,'valid')]
bst = xgb.train(params,d_train,400,watchlist,early_stopping_rounds =20,verbose_eval=20 )
xgdmat = xgb.DMatrix(X_train,y_train)
predict_y  = bst.predict(d_test)
print("The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```
[16:35:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 428 extra nodes, 0 pruned nodes, max_depth=8
[0] train-logloss:0.563224 valid-logloss:0.564318
Multiple eval metrics have been passed: 'valid-logloss' will be used for early stopping.

Will train until valid-logloss hasn't improved in 20 rounds.
[16:35:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 438 extra nodes, 0 pruned nodes, max_depth=8
[1] train-logloss:0.49116 valid-logloss:0.493404
[16:35:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 436 extra nodes, 0 pruned nodes, max_depth=8
[2] train-logloss:0.444831 valid-logloss:0.448006
[16:35:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 402 extra nodes, 0 pruned nodes, max_depth=8
[3] train-logloss:0.413217 valid-logloss:0.417429
[16:35:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 432 extra nodes, 0 pruned nodes, max_depth=8
[4] train-logloss:0.391572 valid-logloss:0.39691
[16:35:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 424 extra nodes, 0 pruned nodes, max_depth=8
[5] train-logloss:0.375336 valid-logloss:0.381728
[16:36:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 308 extra nodes, 0 pruned nodes, max_depth=8
[6] train-logloss:0.363832 valid-logloss:0.371003
[16:36:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 304 extra nodes, 0 pruned nodes, max_depth=8
[7] train-logloss:0.355512 valid-logloss:0.363323
[16:36:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 364 extra nodes, 0 pruned nodes, max_depth=8
[8] train-logloss:0.348159 valid-logloss:0.357038
[16:36:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 310 extra nodes, 0 pruned nodes, max_depth=8
[9] train-logloss:0.342908 valid-logloss:0.352556
[16:36:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 236 extra nodes, 0 pruned nodes, max_depth=8
[10] train-logloss:0.338708 valid-logloss:0.348946
[16:36:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 192 extra nodes, 0 pruned nodes, max_depth=8
[11] train-logloss:0.335612 valid-logloss:0.346227
[16:36:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 294 extra nodes, 0 pruned nodes, max_depth=8
[12] train-logloss:0.331781 valid-logloss:0.34298
[16:36:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 218 extra nodes, 0 pruned nodes, max_depth=8
[13] train-logloss:0.329309 valid-logloss:0.340998
[16:36:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 152 extra nodes, 0 pruned nodes, max_depth=8
[14] train-logloss:0.327789 valid-logloss:0.339962
[16:36:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 162 extra nodes, 0 pruned nodes, max_depth=8
[15] train-logloss:0.325945 valid-logloss:0.338442
[16:36:46] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 152 extra nodes, 0 pruned nodes, max_depth=8
[16] train-logloss:0.324376 valid-logloss:0.33709
[16:36:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 136 extra nodes, 0 pruned nodes, max_depth=8
[17] train-logloss:0.322711 valid-logloss:0.335844
[16:36:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 166 extra nodes, 0 pruned nodes, max_depth=8
[18] train-logloss:0.32138 valid-logloss:0.335022
[16:36:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 238 extra nodes, 0 pruned nodes, max_depth=8
[19] train-logloss:0.320018 valid-logloss:0.334141
[16:37:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 160 extra nodes, 0 pruned nodes, max_depth=8
[20] train-logloss:0.318949 valid-logloss:0.333538
[16:37:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 170 extra nodes, 0 pruned nodes, max depth=8
```

```
[21] train-logloss:0.317797 valid-logloss:0.332757
[16:37:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 140 extra nodes, 0 pruned nodes, max_depth=8
[22] train-logloss:0.316705 valid-logloss:0.331916
[16:37:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 188 extra nodes, 0 pruned nodes, max_depth=8
[23] train-logloss:0.315658 valid-logloss:0.33136
[16:37:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 150 extra nodes, 0 pruned nodes, max_depth=8
[24] train-logloss:0.314888 valid-logloss:0.330904
[16:37:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 170 extra nodes, 0 pruned nodes, max_depth=8
[25] train-logloss:0.314046 valid-logloss:0.330544
[16:37:27] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 198 extra nodes, 0 pruned nodes, max_depth=8
[26] train-logloss:0.312704 valid-logloss:0.329911
[16:37:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 184 extra nodes, 0 pruned nodes, max_depth=8
[27] train-logloss:0.311526 valid-logloss:0.329161
[16:37:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 98 extra nodes, 0 pruned nodes, max_depth=8
[28] train-logloss:0.310966 valid-logloss:0.328812
[16:37:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 162 extra nodes, 0 pruned nodes, max_depth=8
[29] train-logloss:0.310222 valid-logloss:0.328493
[16:37:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 146 extra nodes, 0 pruned nodes, max_depth=8
[30] train-logloss:0.309446 valid-logloss:0.328018
[16:37:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 158 extra nodes, 0 pruned nodes, max_depth=8
[31] train-logloss:0.308721 valid-logloss:0.327662
[16:37:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 170 extra nodes, 0 pruned nodes, max_depth=8
[32] train-logloss:0.307798 valid-logloss:0.327128
[16:37:57] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 180 extra nodes, 0 pruned nodes, max_depth=8
[33] train-logloss:0.306138 valid-logloss:0.325872
[16:38:01] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 196 extra nodes, 0 pruned nodes, max_depth=8
[34] train-logloss:0.305325 valid-logloss:0.325547
[16:38:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 156 extra nodes, 0 pruned nodes, max_depth=8
[35] train-logloss:0.304604 valid-logloss:0.325177
[16:38:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 134 extra nodes, 0 pruned nodes, max_depth=8
[36] train-logloss:0.304024 valid-logloss:0.324898
[16:38:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 142 extra nodes, 0 pruned nodes, max_depth=8
[37] train-logloss:0.303477 valid-logloss:0.32472
[16:38:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 182 extra nodes, 0 pruned nodes, max_depth=8
[38] train-logloss:0.30285 valid-logloss:0.324527
[16:38:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 146 extra nodes, 0 pruned nodes, max_depth=8
[39] train-logloss:0.302033 valid-logloss:0.323928
[16:38:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 86 extra nodes, 0 pruned nodes, max_depth=8
[40] train-logloss:0.301593 valid-logloss:0.323713
[16:38:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 104 extra nodes, 0 pruned nodes, max_depth=8
[41] train-logloss:0.301052 valid-logloss:0.323454
[16:38:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 164 extra nodes, 0 pruned nodes, max_depth=8
[42] train-logloss:0.300456 valid-logloss:0.323266
[16:38:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 160 extra nodes, 0 pruned nodes, max_depth=8
[43] train-logloss:0.299858 valid-logloss:0.323042
[16:38:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 152 extra nodes, 0 pruned nodes, max_depth=8
[44] train-logloss:0.299043 valid-logloss:0.322552
[16:38:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 114 extra nodes, 0 pruned nodes, max_depth=8
[45] train-logloss:0.298542 valid-logloss:0.322298
[16:38:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 114 extra nodes, 0 pruned nodes, max_depth=8
[46] train-logloss:0.298128 valid-logloss:0.322115
[16:38:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
```

1 roots, 100 extra nodes, 0 pruned nodes, max_depth=8
[47] train-logloss:0.297456 valid-logloss:0.321747
[16:38:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[48] train-logloss:0.297176 valid-logloss:0.32162
[16:39:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 142 extra nodes, 0 pruned nodes, max_depth=8
[49] train-logloss:0.296472 valid-logloss:0.321342
[16:39:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 142 extra nodes, 0 pruned nodes, max_depth=8
[50] train-logloss:0.295827 valid-logloss:0.320999
[16:39:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 178 extra nodes, 0 pruned nodes, max_depth=8
[51] train-logloss:0.294735 valid-logloss:0.320511
[16:39:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 232 extra nodes, 0 pruned nodes, max_depth=8
[52] train-logloss:0.29377 valid-logloss:0.32021
[16:39:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 148 extra nodes, 0 pruned nodes, max_depth=8
[53] train-logloss:0.293262 valid-logloss:0.320028
[16:39:23] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 168 extra nodes, 0 pruned nodes, max_depth=8
[54] train-logloss:0.292649 valid-logloss:0.319847
[16:39:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[55] train-logloss:0.292357 valid-logloss:0.319701
[16:39:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 124 extra nodes, 0 pruned nodes, max_depth=8
[56] train-logloss:0.291861 valid-logloss:0.319482
[16:39:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 204 extra nodes, 0 pruned nodes, max_depth=8
[57] train-logloss:0.29063 valid-logloss:0.318778
[16:39:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 310 extra nodes, 0 pruned nodes, max_depth=8
[58] train-logloss:0.28902 valid-logloss:0.317996
[16:39:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 130 extra nodes, 0 pruned nodes, max_depth=8
[59] train-logloss:0.2886 valid-logloss:0.317806
[16:39:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 118 extra nodes, 0 pruned nodes, max_depth=8
[60] train-logloss:0.288223 valid-logloss:0.317631
[16:39:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 108 extra nodes, 0 pruned nodes, max_depth=8
[61] train-logloss:0.287913 valid-logloss:0.317518
[16:39:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 106 extra nodes, 0 pruned nodes, max_depth=8
[62] train-logloss:0.287558 valid-logloss:0.317321
[16:40:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 84 extra nodes, 0 pruned nodes, max_depth=8
[63] train-logloss:0.287265 valid-logloss:0.317233
[16:40:04] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 74 extra nodes, 0 pruned nodes, max_depth=8
[64] train-logloss:0.28695 valid-logloss:0.317104
[16:40:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 160 extra nodes, 0 pruned nodes, max_depth=8
[65] train-logloss:0.286431 valid-logloss:0.316943
[16:40:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 96 extra nodes, 0 pruned nodes, max_depth=8
[66] train-logloss:0.286155 valid-logloss:0.316814
[16:40:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 68 extra nodes, 0 pruned nodes, max_depth=8
[67] train-logloss:0.285929 valid-logloss:0.316748
[16:40:20] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[68] train-logloss:0.285649 valid-logloss:0.316608
[16:40:23] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 200 extra nodes, 0 pruned nodes, max_depth=8
[69] train-logloss:0.28505 valid-logloss:0.316426
[16:40:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 80 extra nodes, 0 pruned nodes, max_depth=8
[70] train-logloss:0.284659 valid-logloss:0.316145
[16:40:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 104 extra nodes, 0 pruned nodes, max_depth=8
[71] train-logloss:0.284341 valid-logloss:0.316031
[16:40:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 168 extra nodes, 0 pruned nodes, max_depth=8
[72] train-logloss:0.283386 valid-logloss:0.315492

[16:40:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 150 extra nodes, 0 pruned nodes, max_depth=8
[73] train-logloss:0.282809 valid-logloss:0.315287
[16:40:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 80 extra nodes, 0 pruned nodes, max_depth=8
[74] train-logloss:0.282522 valid-logloss:0.315209
[16:40:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[75] train-logloss:0.282257 valid-logloss:0.315167
[16:40:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 114 extra nodes, 0 pruned nodes, max_depth=8
[76] train-logloss:0.28189 valid-logloss:0.315018
[16:40:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 112 extra nodes, 0 pruned nodes, max_depth=8
[77] train-logloss:0.281534 valid-logloss:0.314926
[16:40:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 142 extra nodes, 0 pruned nodes, max_depth=8
[78] train-logloss:0.280768 valid-logloss:0.31468
[16:41:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 132 extra nodes, 0 pruned nodes, max_depth=8
[79] train-logloss:0.280324 valid-logloss:0.314581
[16:41:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[80] train-logloss:0.280152 valid-logloss:0.314466
[16:41:09] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 74 extra nodes, 0 pruned nodes, max_depth=8
[81] train-logloss:0.279939 valid-logloss:0.314392
[16:41:13] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 82 extra nodes, 0 pruned nodes, max_depth=8
[82] train-logloss:0.279678 valid-logloss:0.314325
[16:41:17] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 72 extra nodes, 0 pruned nodes, max_depth=8
[83] train-logloss:0.279415 valid-logloss:0.314232
[16:41:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 128 extra nodes, 0 pruned nodes, max_depth=8
[84] train-logloss:0.27894 valid-logloss:0.314019
[16:41:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 142 extra nodes, 0 pruned nodes, max_depth=8
[85] train-logloss:0.278538 valid-logloss:0.313972
[16:41:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 96 extra nodes, 0 pruned nodes, max_depth=8
[86] train-logloss:0.278188 valid-logloss:0.313798
[16:41:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 100 extra nodes, 0 pruned nodes, max_depth=8
[87] train-logloss:0.277878 valid-logloss:0.313708
[16:41:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 86 extra nodes, 0 pruned nodes, max_depth=8
[88] train-logloss:0.277603 valid-logloss:0.31362
[16:41:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[89] train-logloss:0.277395 valid-logloss:0.313552
[16:41:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[90] train-logloss:0.277135 valid-logloss:0.313485
[16:41:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 96 extra nodes, 0 pruned nodes, max_depth=8
[91] train-logloss:0.276853 valid-logloss:0.313467
[16:41:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 66 extra nodes, 0 pruned nodes, max_depth=8
[92] train-logloss:0.27666 valid-logloss:0.313359
[16:41:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 138 extra nodes, 0 pruned nodes, max_depth=8
[93] train-logloss:0.276269 valid-logloss:0.313289
[16:41:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 80 extra nodes, 0 pruned nodes, max_depth=8
[94] train-logloss:0.275995 valid-logloss:0.31321
[16:42:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 106 extra nodes, 0 pruned nodes, max_depth=8
[95] train-logloss:0.27573 valid-logloss:0.313149
[16:42:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=8
[96] train-logloss:0.275616 valid-logloss:0.313073
[16:42:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 74 extra nodes, 0 pruned nodes, max_depth=8
[97] train-logloss:0.275392 valid-logloss:0.313047
[16:42:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 118 extra nodes, 0 pruned nodes, max_depth=8

```
[98] train-logloss:0.274897 valid-logloss:0.312851
[16:42:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 74 extra nodes, 0 pruned nodes, max_depth=8
[99] train-logloss:0.274613 valid-logloss:0.312681
[16:42:23] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 78 extra nodes, 0 pruned nodes, max_depth=8
[100] train-logloss:0.274358 valid-logloss:0.312562
[16:42:27] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 126 extra nodes, 0 pruned nodes, max_depth=8
[101] train-logloss:0.274029 valid-logloss:0.312514
[16:42:31] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 148 extra nodes, 0 pruned nodes, max_depth=8
[102] train-logloss:0.273603 valid-logloss:0.312402
[16:42:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 114 extra nodes, 0 pruned nodes, max_depth=8
[103] train-logloss:0.273245 valid-logloss:0.312223
[16:42:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 100 extra nodes, 0 pruned nodes, max_depth=8
[104] train-logloss:0.272928 valid-logloss:0.312141
[16:42:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[105] train-logloss:0.272799 valid-logloss:0.312132
[16:42:46] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 138 extra nodes, 0 pruned nodes, max_depth=8
[106] train-logloss:0.272438 valid-logloss:0.312051
[16:42:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[107] train-logloss:0.272263 valid-logloss:0.311992
[16:42:53] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 150 extra nodes, 0 pruned nodes, max_depth=8
[108] train-logloss:0.271893 valid-logloss:0.31189
[16:42:57] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 88 extra nodes, 0 pruned nodes, max_depth=8
[109] train-logloss:0.271642 valid-logloss:0.311808
[16:43:01] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 82 extra nodes, 0 pruned nodes, max_depth=8
[110] train-logloss:0.271421 valid-logloss:0.311722
[16:43:05] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[111] train-logloss:0.271252 valid-logloss:0.311665
[16:43:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 114 extra nodes, 0 pruned nodes, max_depth=8
[112] train-logloss:0.270943 valid-logloss:0.311565
[16:43:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[113] train-logloss:0.270718 valid-logloss:0.31149
[16:43:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 112 extra nodes, 0 pruned nodes, max_depth=8
[114] train-logloss:0.270423 valid-logloss:0.311407
[16:43:20] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 100 extra nodes, 0 pruned nodes, max_depth=8
[115] train-logloss:0.270174 valid-logloss:0.311314
[16:43:24] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 116 extra nodes, 0 pruned nodes, max_depth=8
[116] train-logloss:0.269902 valid-logloss:0.31122
[16:43:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 94 extra nodes, 0 pruned nodes, max_depth=8
[117] train-logloss:0.269656 valid-logloss:0.311134
[16:43:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 128 extra nodes, 0 pruned nodes, max_depth=8
[118] train-logloss:0.269289 valid-logloss:0.311085
[16:43:35] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 46 extra nodes, 0 pruned nodes, max_depth=8
[119] train-logloss:0.269145 valid-logloss:0.311022
[16:43:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 90 extra nodes, 0 pruned nodes, max_depth=8
[120] train-logloss:0.268932 valid-logloss:0.311011
[16:43:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 90 extra nodes, 0 pruned nodes, max_depth=8
[121] train-logloss:0.268716 valid-logloss:0.310953
[16:43:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 110 extra nodes, 0 pruned nodes, max_depth=8
[122] train-logloss:0.268136 valid-logloss:0.310594
[16:43:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 92 extra nodes, 0 pruned nodes, max_depth=8
[123] train-logloss:0.267921 valid-logloss:0.310528
[16:43:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
```

[16:43:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 90 extra nodes, 0 pruned nodes, max_depth=8
[124] train-logloss:0.267648 valid-logloss:0.310519
[16:43:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[125] train-logloss:0.267476 valid-logloss:0.310485
[16:44:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[126] train-logloss:0.267303 valid-logloss:0.31043
[16:44:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[127] train-logloss:0.267091 valid-logloss:0.31038
[16:44:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 78 extra nodes, 0 pruned nodes, max_depth=8
[128] train-logloss:0.266881 valid-logloss:0.31034
[16:44:13] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 82 extra nodes, 0 pruned nodes, max_depth=8
[129] train-logloss:0.26669 valid-logloss:0.310303
[16:44:17] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[130] train-logloss:0.266574 valid-logloss:0.310272
[16:44:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 96 extra nodes, 0 pruned nodes, max_depth=8
[131] train-logloss:0.26634 valid-logloss:0.310198
[16:44:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[132] train-logloss:0.266153 valid-logloss:0.31015
[16:44:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 46 extra nodes, 0 pruned nodes, max_depth=8
[133] train-logloss:0.265904 valid-logloss:0.310012
[16:44:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 268 extra nodes, 0 pruned nodes, max_depth=8
[134] train-logloss:0.26497 valid-logloss:0.309721
[16:44:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 130 extra nodes, 0 pruned nodes, max_depth=8
[135] train-logloss:0.264566 valid-logloss:0.3096
[16:44:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 108 extra nodes, 0 pruned nodes, max_depth=8
[136] train-logloss:0.264307 valid-logloss:0.309561
[16:44:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[137] train-logloss:0.264143 valid-logloss:0.309507
[16:44:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 110 extra nodes, 0 pruned nodes, max_depth=8
[138] train-logloss:0.263745 valid-logloss:0.309341
[16:44:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 92 extra nodes, 0 pruned nodes, max_depth=8
[139] train-logloss:0.263556 valid-logloss:0.309319
[16:44:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 84 extra nodes, 0 pruned nodes, max_depth=8
[140] train-logloss:0.263344 valid-logloss:0.309211
[16:44:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 44 extra nodes, 0 pruned nodes, max_depth=8
[141] train-logloss:0.263212 valid-logloss:0.309132
[16:45:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[142] train-logloss:0.263076 valid-logloss:0.309094
[16:45:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[143] train-logloss:0.262923 valid-logloss:0.309046
[16:45:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 172 extra nodes, 0 pruned nodes, max_depth=8
[144] train-logloss:0.26232 valid-logloss:0.308854
[16:45:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 100 extra nodes, 0 pruned nodes, max_depth=8
[145] train-logloss:0.26199 valid-logloss:0.308745
[16:45:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[146] train-logloss:0.261817 valid-logloss:0.30871
[16:45:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 116 extra nodes, 0 pruned nodes, max_depth=8
[147] train-logloss:0.261385 valid-logloss:0.308518
[16:45:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 84 extra nodes, 0 pruned nodes, max_depth=8
[148] train-logloss:0.261157 valid-logloss:0.308444
[16:45:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 66 extra nodes, 0 pruned nodes, max_depth=8
[149] train-logloss:0.260988 valid-logloss:0.308402

[149] train-logloss:0.260966 valid-logloss:0.308402
[16:45:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[150] train-logloss:0.260872 valid-logloss:0.30837
[16:45:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 46 extra nodes, 0 pruned nodes, max_depth=8
[151] train-logloss:0.260758 valid-logloss:0.308355
[16:45:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[152] train-logloss:0.260578 valid-logloss:0.30829
[16:45:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[153] train-logloss:0.260437 valid-logloss:0.308253
[16:45:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 74 extra nodes, 0 pruned nodes, max_depth=8
[154] train-logloss:0.26024 valid-logloss:0.30819
[16:45:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 110 extra nodes, 0 pruned nodes, max_depth=8
[155] train-logloss:0.26 valid-logloss:0.308122
[16:45:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[156] train-logloss:0.259853 valid-logloss:0.308093
[16:45:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[157] train-logloss:0.259672 valid-logloss:0.308047
[16:46:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[158] train-logloss:0.259564 valid-logloss:0.308039
[16:46:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[159] train-logloss:0.259434 valid-logloss:0.307994
[16:46:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[160] train-logloss:0.259329 valid-logloss:0.307952
[16:46:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 82 extra nodes, 0 pruned nodes, max_depth=8
[161] train-logloss:0.259141 valid-logloss:0.307916
[16:46:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[162] train-logloss:0.259021 valid-logloss:0.30787
[16:46:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[163] train-logloss:0.258886 valid-logloss:0.307847
[16:46:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[164] train-logloss:0.258712 valid-logloss:0.307762
[16:46:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[165] train-logloss:0.258574 valid-logloss:0.307704
[16:46:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 82 extra nodes, 0 pruned nodes, max_depth=8
[166] train-logloss:0.258354 valid-logloss:0.307665
[16:46:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 68 extra nodes, 0 pruned nodes, max_depth=8
[167] train-logloss:0.258218 valid-logloss:0.30764
[16:46:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 112 extra nodes, 0 pruned nodes, max_depth=8
[168] train-logloss:0.257945 valid-logloss:0.307543
[16:46:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 84 extra nodes, 0 pruned nodes, max_depth=8
[169] train-logloss:0.257751 valid-logloss:0.307488
[16:46:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 62 extra nodes, 0 pruned nodes, max_depth=8
[170] train-logloss:0.257586 valid-logloss:0.30743
[16:46:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[171] train-logloss:0.257436 valid-logloss:0.307346
[16:46:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 42 extra nodes, 0 pruned nodes, max_depth=8
[172] train-logloss:0.257334 valid-logloss:0.307317
[16:47:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 92 extra nodes, 0 pruned nodes, max_depth=8
[173] train-logloss:0.257176 valid-logloss:0.307261
[16:47:04] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 108 extra nodes, 0 pruned nodes, max_depth=8
[174] train-logloss:0.256952 valid-logloss:0.307187
[16:47:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 84 extra nodes, 0 pruned nodes, max_depth=8

```
1 roots, 84 extra nodes, 0 pruned nodes, max_depth=8
[175] train-logloss:0.256754 valid-logloss:0.307153
[16:47:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[176] train-logloss:0.256622 valid-logloss:0.307105
[16:47:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 84 extra nodes, 0 pruned nodes, max_depth=8
[177] train-logloss:0.256228 valid-logloss:0.30693
[16:47:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 110 extra nodes, 0 pruned nodes, max_depth=8
[178] train-logloss:0.255984 valid-logloss:0.306883
[16:47:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[179] train-logloss:0.2558 valid-logloss:0.306826
[16:47:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 80 extra nodes, 0 pruned nodes, max_depth=8
[180] train-logloss:0.255581 valid-logloss:0.306802
[16:47:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 86 extra nodes, 0 pruned nodes, max_depth=8
[181] train-logloss:0.255371 valid-logloss:0.306713
[16:47:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[182] train-logloss:0.255236 valid-logloss:0.306703
[16:47:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 94 extra nodes, 0 pruned nodes, max_depth=8
[183] train-logloss:0.254933 valid-logloss:0.30663
[16:47:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[184] train-logloss:0.254759 valid-logloss:0.306571
[16:47:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 46 extra nodes, 0 pruned nodes, max_depth=8
[185] train-logloss:0.254648 valid-logloss:0.30654
[16:47:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 90 extra nodes, 0 pruned nodes, max_depth=8
[186] train-logloss:0.254453 valid-logloss:0.306468
[16:47:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 106 extra nodes, 0 pruned nodes, max_depth=8
[187] train-logloss:0.254219 valid-logloss:0.306439
[16:47:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 32 extra nodes, 0 pruned nodes, max_depth=8
[188] train-logloss:0.254127 valid-logloss:0.306408
[16:48:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[189] train-logloss:0.25399 valid-logloss:0.306376
[16:48:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 78 extra nodes, 0 pruned nodes, max_depth=8
[190] train-logloss:0.253814 valid-logloss:0.306345
[16:48:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[191] train-logloss:0.253654 valid-logloss:0.306273
[16:48:12] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 160 extra nodes, 0 pruned nodes, max_depth=8
[192] train-logloss:0.253207 valid-logloss:0.306162
[16:48:16] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 134 extra nodes, 0 pruned nodes, max_depth=8
[193] train-logloss:0.252731 valid-logloss:0.30594
[16:48:20] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 88 extra nodes, 0 pruned nodes, max_depth=8
[194] train-logloss:0.252476 valid-logloss:0.305794
[16:48:24] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 78 extra nodes, 0 pruned nodes, max_depth=8
[195] train-logloss:0.2523 valid-logloss:0.305751
[16:48:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 62 extra nodes, 0 pruned nodes, max_depth=8
[196] train-logloss:0.252181 valid-logloss:0.305716
[16:48:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[197] train-logloss:0.252084 valid-logloss:0.305706
[16:48:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 108 extra nodes, 0 pruned nodes, max_depth=8
[198] train-logloss:0.251799 valid-logloss:0.305632
[16:48:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 84 extra nodes, 0 pruned nodes, max_depth=8
[199] train-logloss:0.251605 valid-logloss:0.305597
[16:48:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[200] train-logloss:0.251396 valid-logloss:0.305529
[16:48:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
```

[16:48:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[201] train-logloss:0.251134 valid-logloss:0.305479
[16:48:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 128 extra nodes, 0 pruned nodes, max_depth=8
[202] train-logloss:0.250811 valid-logloss:0.305413
[16:48:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[203] train-logloss:0.250638 valid-logloss:0.305343
[16:48:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 130 extra nodes, 0 pruned nodes, max_depth=8
[204] train-logloss:0.250376 valid-logloss:0.305326
[16:49:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 66 extra nodes, 0 pruned nodes, max_depth=8
[205] train-logloss:0.250242 valid-logloss:0.305274
[16:49:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 32 extra nodes, 0 pruned nodes, max_depth=8
[206] train-logloss:0.250159 valid-logloss:0.305249
[16:49:09] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 106 extra nodes, 0 pruned nodes, max_depth=8
[207] train-logloss:0.249951 valid-logloss:0.305223
[16:49:13] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[208] train-logloss:0.249857 valid-logloss:0.305202
[16:49:17] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[209] train-logloss:0.24976 valid-logloss:0.305195
[16:49:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[210] train-logloss:0.249652 valid-logloss:0.305196
[16:49:24] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[211] train-logloss:0.249552 valid-logloss:0.305159
[16:49:28] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 54 extra nodes, 0 pruned nodes, max_depth=8
[212] train-logloss:0.24944 valid-logloss:0.305124
[16:49:32] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 42 extra nodes, 0 pruned nodes, max_depth=8
[213] train-logloss:0.249341 valid-logloss:0.3051
[16:49:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 86 extra nodes, 0 pruned nodes, max_depth=8
[214] train-logloss:0.249156 valid-logloss:0.305059
[16:49:39] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 44 extra nodes, 0 pruned nodes, max_depth=8
[215] train-logloss:0.249059 valid-logloss:0.305019
[16:49:43] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[216] train-logloss:0.248959 valid-logloss:0.305004
[16:49:47] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[217] train-logloss:0.248826 valid-logloss:0.304969
[16:49:51] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[218] train-logloss:0.248736 valid-logloss:0.30493
[16:49:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 48 extra nodes, 0 pruned nodes, max_depth=8
[219] train-logloss:0.248642 valid-logloss:0.304889
[16:49:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 172 extra nodes, 0 pruned nodes, max_depth=8
[220] train-logloss:0.247997 valid-logloss:0.304702
[16:50:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 90 extra nodes, 0 pruned nodes, max_depth=8
[221] train-logloss:0.247797 valid-logloss:0.304658
[16:50:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[222] train-logloss:0.247722 valid-logloss:0.304676
[16:50:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=8
[223] train-logloss:0.247649 valid-logloss:0.304662
[16:50:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[224] train-logloss:0.247535 valid-logloss:0.304652
[16:50:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[225] train-logloss:0.247393 valid-logloss:0.304614
[16:50:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[226] train-logloss:0.247262 valid-logloss:0.304575

[226] train-logloss:0.247262 valid-logloss:0.304575
[16:50:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[227] train-logloss:0.247171 valid-logloss:0.304538
[16:50:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=8
[228] train-logloss:0.247106 valid-logloss:0.304511
[16:50:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 94 extra nodes, 0 pruned nodes, max_depth=8
[229] train-logloss:0.246892 valid-logloss:0.304479
[16:50:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[230] train-logloss:0.246741 valid-logloss:0.304468
[16:50:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 74 extra nodes, 0 pruned nodes, max_depth=8
[231] train-logloss:0.246535 valid-logloss:0.304423
[16:50:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 34 extra nodes, 0 pruned nodes, max_depth=8
[232] train-logloss:0.246462 valid-logloss:0.304438
[16:50:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 90 extra nodes, 0 pruned nodes, max_depth=8
[233] train-logloss:0.246265 valid-logloss:0.304431
[16:50:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 106 extra nodes, 0 pruned nodes, max_depth=8
[234] train-logloss:0.246017 valid-logloss:0.304382
[16:50:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 68 extra nodes, 0 pruned nodes, max_depth=8
[235] train-logloss:0.245895 valid-logloss:0.304389
[16:50:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[236] train-logloss:0.245794 valid-logloss:0.304393
[16:51:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[237] train-logloss:0.245671 valid-logloss:0.304368
[16:51:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 54 extra nodes, 0 pruned nodes, max_depth=8
[238] train-logloss:0.245556 valid-logloss:0.304356
[16:51:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 54 extra nodes, 0 pruned nodes, max_depth=8
[239] train-logloss:0.245456 valid-logloss:0.304372
[16:51:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 90 extra nodes, 0 pruned nodes, max_depth=8
[240] train-logloss:0.245283 valid-logloss:0.304376
[16:51:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[241] train-logloss:0.245198 valid-logloss:0.30437
[16:51:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[242] train-logloss:0.245089 valid-logloss:0.304346
[16:51:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[243] train-logloss:0.244962 valid-logloss:0.304292
[16:51:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[244] train-logloss:0.244841 valid-logloss:0.304267
[16:51:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 38 extra nodes, 0 pruned nodes, max_depth=8
[245] train-logloss:0.24476 valid-logloss:0.304252
[16:51:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[246] train-logloss:0.244689 valid-logloss:0.304214
[16:51:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 62 extra nodes, 0 pruned nodes, max_depth=8
[247] train-logloss:0.244549 valid-logloss:0.304184
[16:51:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 88 extra nodes, 0 pruned nodes, max_depth=8
[248] train-logloss:0.244365 valid-logloss:0.304165
[16:51:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 66 extra nodes, 0 pruned nodes, max_depth=8
[249] train-logloss:0.244219 valid-logloss:0.304116
[16:51:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 48 extra nodes, 0 pruned nodes, max_depth=8
[250] train-logloss:0.244093 valid-logloss:0.304064
[16:51:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 100 extra nodes, 0 pruned nodes, max_depth=8
[251] train-logloss:0.243878 valid-logloss:0.304018
[16:51:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,

1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[252] train-logloss:0.24378 valid-logloss:0.303988
[16:52:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[253] train-logloss:0.243664 valid-logloss:0.30395
[16:52:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[254] train-logloss:0.243557 valid-logloss:0.303921
[16:52:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[255] train-logloss:0.243461 valid-logloss:0.303915
[16:52:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 68 extra nodes, 0 pruned nodes, max_depth=8
[256] train-logloss:0.243314 valid-logloss:0.303881
[16:52:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[257] train-logloss:0.243149 valid-logloss:0.303847
[16:52:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[258] train-logloss:0.243053 valid-logloss:0.303807
[16:52:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 38 extra nodes, 0 pruned nodes, max_depth=8
[259] train-logloss:0.242957 valid-logloss:0.303754
[16:52:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 66 extra nodes, 0 pruned nodes, max_depth=8
[260] train-logloss:0.242833 valid-logloss:0.303707
[16:52:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[261] train-logloss:0.2427 valid-logloss:0.303681
[16:52:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[262] train-logloss:0.242635 valid-logloss:0.303658
[16:52:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[263] train-logloss:0.24253 valid-logloss:0.303636
[16:52:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 32 extra nodes, 0 pruned nodes, max_depth=8
[264] train-logloss:0.242458 valid-logloss:0.30364
[16:52:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 32 extra nodes, 0 pruned nodes, max_depth=8
[265] train-logloss:0.242388 valid-logloss:0.303639
[16:52:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=8
[266] train-logloss:0.242327 valid-logloss:0.303617
[16:52:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 72 extra nodes, 0 pruned nodes, max_depth=8
[267] train-logloss:0.242175 valid-logloss:0.303574
[16:52:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 66 extra nodes, 0 pruned nodes, max_depth=8
[268] train-logloss:0.242057 valid-logloss:0.303556
[16:53:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[269] train-logloss:0.241872 valid-logloss:0.303512
[16:53:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[270] train-logloss:0.241767 valid-logloss:0.303483
[16:53:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 82 extra nodes, 0 pruned nodes, max_depth=8
[271] train-logloss:0.241614 valid-logloss:0.303488
[16:53:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[272] train-logloss:0.241502 valid-logloss:0.303447
[16:53:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[273] train-logloss:0.2414 valid-logloss:0.303428
[16:53:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[274] train-logloss:0.241287 valid-logloss:0.303423
[16:53:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[275] train-logloss:0.241178 valid-logloss:0.30337
[16:53:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[276] train-logloss:0.241055 valid-logloss:0.303361
[16:53:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[277] train-logloss:0.240984 valid-logloss:0.303345

```
[16:53:36] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[278] train-logloss:0.240891 valid-logloss:0.303319
[16:53:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[279] train-logloss:0.240793 valid-logloss:0.303292
[16:53:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 54 extra nodes, 0 pruned nodes, max_depth=8
[280] train-logloss:0.240696 valid-logloss:0.303277
[16:53:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 92 extra nodes, 0 pruned nodes, max_depth=8
[281] train-logloss:0.240502 valid-logloss:0.303263
[16:53:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[282] train-logloss:0.240308 valid-logloss:0.303203
[16:53:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[283] train-logloss:0.240182 valid-logloss:0.303152
[16:53:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 96 extra nodes, 0 pruned nodes, max_depth=8
[284] train-logloss:0.240041 valid-logloss:0.303132
[16:54:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=8
[285] train-logloss:0.239984 valid-logloss:0.303128
[16:54:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 82 extra nodes, 0 pruned nodes, max_depth=8
[286] train-logloss:0.239819 valid-logloss:0.303081
[16:54:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 68 extra nodes, 0 pruned nodes, max_depth=8
[287] train-logloss:0.23965 valid-logloss:0.303069
[16:54:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 184 extra nodes, 0 pruned nodes, max_depth=8
[288] train-logloss:0.238959 valid-logloss:0.302823
[16:54:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 80 extra nodes, 0 pruned nodes, max_depth=8
[289] train-logloss:0.238836 valid-logloss:0.30279
[16:54:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[290] train-logloss:0.238755 valid-logloss:0.302755
[16:54:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[291] train-logloss:0.238641 valid-logloss:0.302742
[16:54:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[292] train-logloss:0.238503 valid-logloss:0.302697
[16:54:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 44 extra nodes, 0 pruned nodes, max_depth=8
[293] train-logloss:0.238408 valid-logloss:0.302676
[16:54:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 34 extra nodes, 0 pruned nodes, max_depth=8
[294] train-logloss:0.238346 valid-logloss:0.302671
[16:54:40] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 54 extra nodes, 0 pruned nodes, max_depth=8
[295] train-logloss:0.238239 valid-logloss:0.302643
[16:54:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 44 extra nodes, 0 pruned nodes, max_depth=8
[296] train-logloss:0.238148 valid-logloss:0.302631
[16:54:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 52 extra nodes, 0 pruned nodes, max_depth=8
[297] train-logloss:0.238052 valid-logloss:0.302622
[16:54:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 110 extra nodes, 0 pruned nodes, max_depth=8
[298] train-logloss:0.237851 valid-logloss:0.302587
[16:54:55] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=8
[299] train-logloss:0.237798 valid-logloss:0.302571
[16:54:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 24 extra nodes, 0 pruned nodes, max_depth=8
[300] train-logloss:0.237738 valid-logloss:0.302538
[16:55:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[301] train-logloss:0.237621 valid-logloss:0.302569
[16:55:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 34 extra nodes, 0 pruned nodes, max_depth=8
[302] train-logloss:0.237556 valid-logloss:0.302542
[16:55:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
```

```
[303] train-logloss:0.237427 valid-logloss:0.302524
[16:55:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[304] train-logloss:0.237318 valid-logloss:0.302504
[16:55:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=8
[305] train-logloss:0.237255 valid-logloss:0.302504
[16:55:21] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[306] train-logloss:0.237089 valid-logloss:0.302431
[16:55:25] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 98 extra nodes, 0 pruned nodes, max_depth=8
[307] train-logloss:0.236901 valid-logloss:0.302365
[16:55:29] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[308] train-logloss:0.236727 valid-logloss:0.302324
[16:55:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 214 extra nodes, 0 pruned nodes, max_depth=8
[309] train-logloss:0.236038 valid-logloss:0.302178
[16:55:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 102 extra nodes, 0 pruned nodes, max_depth=8
[310] train-logloss:0.235866 valid-logloss:0.302132
[16:55:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 78 extra nodes, 0 pruned nodes, max_depth=8
[311] train-logloss:0.235725 valid-logloss:0.302086
[16:55:44] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 44 extra nodes, 0 pruned nodes, max_depth=8
[312] train-logloss:0.235635 valid-logloss:0.302084
[16:55:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 58 extra nodes, 0 pruned nodes, max_depth=8
[313] train-logloss:0.235508 valid-logloss:0.302044
[16:55:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[314] train-logloss:0.235433 valid-logloss:0.302016
[16:55:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 68 extra nodes, 0 pruned nodes, max_depth=8
[315] train-logloss:0.235296 valid-logloss:0.30198
[16:55:59] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 66 extra nodes, 0 pruned nodes, max_depth=8
[316] train-logloss:0.235172 valid-logloss:0.301971
[16:56:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 78 extra nodes, 0 pruned nodes, max_depth=8
[317] train-logloss:0.235046 valid-logloss:0.301968
[16:56:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 72 extra nodes, 0 pruned nodes, max_depth=8
[318] train-logloss:0.234894 valid-logloss:0.301958
[16:56:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[319] train-logloss:0.23478 valid-logloss:0.301933
[16:56:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=8
[320] train-logloss:0.234726 valid-logloss:0.301908
[16:56:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[321] train-logloss:0.234619 valid-logloss:0.301899
[16:56:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 42 extra nodes, 0 pruned nodes, max_depth=8
[322] train-logloss:0.234546 valid-logloss:0.301894
[16:56:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 98 extra nodes, 0 pruned nodes, max_depth=8
[323] train-logloss:0.234367 valid-logloss:0.301815
[16:56:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 62 extra nodes, 0 pruned nodes, max_depth=8
[324] train-logloss:0.234267 valid-logloss:0.301773
[16:56:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 86 extra nodes, 0 pruned nodes, max_depth=8
[325] train-logloss:0.23408 valid-logloss:0.301752
[16:56:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 72 extra nodes, 0 pruned nodes, max_depth=8
[326] train-logloss:0.233821 valid-logloss:0.301627
[16:56:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 98 extra nodes, 0 pruned nodes, max_depth=8
[327] train-logloss:0.233461 valid-logloss:0.301399
[16:56:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 90 extra nodes, 0 pruned nodes, max_depth=8
[328] train-logloss:0.233325 valid-logloss:0.30137
[16:56:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
```

```
1 roots, 106 extra nodes, 0 pruned nodes, max_depth=8
[329] train-logloss:0.233054 valid-logloss:0.301288
[16:56:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 116 extra nodes, 0 pruned nodes, max_depth=8
[330] train-logloss:0.232854 valid-logloss:0.301237
[16:56:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 78 extra nodes, 0 pruned nodes, max_depth=8
[331] train-logloss:0.232643 valid-logloss:0.301181
[16:57:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[332] train-logloss:0.232535 valid-logloss:0.301179
[16:57:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[333] train-logloss:0.232409 valid-logloss:0.30111
[16:57:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 80 extra nodes, 0 pruned nodes, max_depth=8
[334] train-logloss:0.232243 valid-logloss:0.301048
[16:57:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[335] train-logloss:0.232125 valid-logloss:0.301002
[16:57:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 46 extra nodes, 0 pruned nodes, max_depth=8
[336] train-logloss:0.232032 valid-logloss:0.301035
[16:57:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 66 extra nodes, 0 pruned nodes, max_depth=8
[337] train-logloss:0.23193 valid-logloss:0.301049
[16:57:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 40 extra nodes, 0 pruned nodes, max_depth=8
[338] train-logloss:0.231855 valid-logloss:0.301034
[16:57:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[339] train-logloss:0.231709 valid-logloss:0.301022
[16:57:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 62 extra nodes, 0 pruned nodes, max_depth=8
[340] train-logloss:0.231608 valid-logloss:0.301
[16:57:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 72 extra nodes, 0 pruned nodes, max_depth=8
[341] train-logloss:0.231452 valid-logloss:0.300975
[16:57:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 90 extra nodes, 0 pruned nodes, max_depth=8
[342] train-logloss:0.231291 valid-logloss:0.300925
[16:57:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[343] train-logloss:0.230876 valid-logloss:0.300683
[16:57:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[344] train-logloss:0.230776 valid-logloss:0.300651
[16:57:50] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 50 extra nodes, 0 pruned nodes, max_depth=8
[345] train-logloss:0.230694 valid-logloss:0.300616
[16:57:54] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 294 extra nodes, 0 pruned nodes, max_depth=8
[346] train-logloss:0.229732 valid-logloss:0.300322
[16:57:58] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 282 extra nodes, 0 pruned nodes, max_depth=8
[347] train-logloss:0.228765 valid-logloss:0.299965
[16:58:02] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 146 extra nodes, 0 pruned nodes, max_depth=8
[348] train-logloss:0.228547 valid-logloss:0.299935
[16:58:06] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 32 extra nodes, 0 pruned nodes, max_depth=8
[349] train-logloss:0.228489 valid-logloss:0.299929
[16:58:10] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 36 extra nodes, 0 pruned nodes, max_depth=8
[350] train-logloss:0.228421 valid-logloss:0.299908
[16:58:14] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 44 extra nodes, 0 pruned nodes, max_depth=8
[351] train-logloss:0.22833 valid-logloss:0.299898
[16:58:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 68 extra nodes, 0 pruned nodes, max_depth=8
[352] train-logloss:0.22819 valid-logloss:0.299885
[16:58:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 34 extra nodes, 0 pruned nodes, max_depth=8
[353] train-logloss:0.228132 valid-logloss:0.299883
[16:58:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 42 extra nodes, 0 pruned nodes, max_depth=8
[354] train-logloss:0.228067 valid-logloss:0.299896
```

```
[16:58:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=8
[355] train-logloss:0.228017 valid-logloss:0.29988
[16:58:34] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 42 extra nodes, 0 pruned nodes, max_depth=8
[356] train-logloss:0.227948 valid-logloss:0.299877
[16:58:38] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 88 extra nodes, 0 pruned nodes, max_depth=8
[357] train-logloss:0.227801 valid-logloss:0.299895
[16:58:42] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 194 extra nodes, 0 pruned nodes, max_depth=8
[358] train-logloss:0.227154 valid-logloss:0.299639
[16:58:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 72 extra nodes, 0 pruned nodes, max_depth=8
[359] train-logloss:0.227019 valid-logloss:0.299592
[16:58:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[360] train-logloss:0.226926 valid-logloss:0.299599
[16:58:53] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[361] train-logloss:0.226801 valid-logloss:0.299567
[16:58:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=8
[362] train-logloss:0.226748 valid-logloss:0.299541
[16:59:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=8
[363] train-logloss:0.226693 valid-logloss:0.299538
[16:59:04] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 42 extra nodes, 0 pruned nodes, max_depth=8
[364] train-logloss:0.226623 valid-logloss:0.299548
[16:59:08] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 34 extra nodes, 0 pruned nodes, max_depth=8
[365] train-logloss:0.226554 valid-logloss:0.299532
[16:59:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 54 extra nodes, 0 pruned nodes, max_depth=8
[366] train-logloss:0.226468 valid-logloss:0.29952
[16:59:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 38 extra nodes, 0 pruned nodes, max_depth=8
[367] train-logloss:0.226405 valid-logloss:0.299518
[16:59:19] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 62 extra nodes, 0 pruned nodes, max_depth=8
[368] train-logloss:0.226304 valid-logloss:0.29951
[16:59:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 46 extra nodes, 0 pruned nodes, max_depth=8
[369] train-logloss:0.226225 valid-logloss:0.299524
[16:59:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 42 extra nodes, 0 pruned nodes, max_depth=8
[370] train-logloss:0.226155 valid-logloss:0.299491
[16:59:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 28 extra nodes, 0 pruned nodes, max_depth=8
[371] train-logloss:0.226103 valid-logloss:0.299483
[16:59:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 80 extra nodes, 0 pruned nodes, max_depth=8
[372] train-logloss:0.225965 valid-logloss:0.299471
[16:59:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 44 extra nodes, 0 pruned nodes, max_depth=8
[373] train-logloss:0.2259 valid-logloss:0.299456
[16:59:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 88 extra nodes, 0 pruned nodes, max_depth=8
[374] train-logloss:0.225731 valid-logloss:0.299456
[16:59:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 62 extra nodes, 0 pruned nodes, max_depth=8
[375] train-logloss:0.225644 valid-logloss:0.299444
[16:59:49] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 62 extra nodes, 0 pruned nodes, max_depth=8
[376] train-logloss:0.225541 valid-logloss:0.299448
[16:59:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 26 extra nodes, 0 pruned nodes, max_depth=8
[377] train-logloss:0.225495 valid-logloss:0.299428
[16:59:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 42 extra nodes, 0 pruned nodes, max_depth=8
[378] train-logloss:0.225409 valid-logloss:0.299414
[17:00:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 34 extra nodes, 0 pruned nodes, max_depth=8
[379] train-logloss:0.225356 valid-logloss:0.2994
[17:00:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 78 extra nodes, 0 pruned nodes, max_depth=8
```

```
[380] train-logloss:0.225215 valid-logloss:0.299387
[17:00:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 64 extra nodes, 0 pruned nodes, max_depth=8
[381] train-logloss:0.225099 valid-logloss:0.299402
[17:00:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 112 extra nodes, 0 pruned nodes, max_depth=8
[382] train-logloss:0.224772 valid-logloss:0.299312
[17:00:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 38 extra nodes, 0 pruned nodes, max_depth=8
[383] train-logloss:0.224702 valid-logloss:0.299294
[17:00:18] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 32 extra nodes, 0 pruned nodes, max_depth=8
[384] train-logloss:0.224651 valid-logloss:0.299282
[17:00:22] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 96 extra nodes, 0 pruned nodes, max_depth=8
[385] train-logloss:0.224499 valid-logloss:0.299266
[17:00:26] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 76 extra nodes, 0 pruned nodes, max_depth=8
[386] train-logloss:0.224368 valid-logloss:0.299255
[17:00:30] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=8
[387] train-logloss:0.224312 valid-logloss:0.29925
[17:00:33] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 30 extra nodes, 0 pruned nodes, max_depth=8
[388] train-logloss:0.224258 valid-logloss:0.299225
[17:00:37] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 54 extra nodes, 0 pruned nodes, max_depth=8
[389] train-logloss:0.224173 valid-logloss:0.299186
[17:00:41] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 48 extra nodes, 0 pruned nodes, max_depth=8
[390] train-logloss:0.224088 valid-logloss:0.299165
[17:00:45] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 54 extra nodes, 0 pruned nodes, max_depth=8
[391] train-logloss:0.223991 valid-logloss:0.299127
[17:00:48] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 16 extra nodes, 0 pruned nodes, max_depth=8
[392] train-logloss:0.223952 valid-logloss:0.299129
[17:00:52] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 56 extra nodes, 0 pruned nodes, max_depth=8
[393] train-logloss:0.223876 valid-logloss:0.299095
[17:00:56] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 86 extra nodes, 0 pruned nodes, max_depth=8
[394] train-logloss:0.22375 valid-logloss:0.299083
[17:01:00] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[395] train-logloss:0.223632 valid-logloss:0.299053
[17:01:03] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 70 extra nodes, 0 pruned nodes, max_depth=8
[396] train-logloss:0.223468 valid-logloss:0.299025
[17:01:07] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 34 extra nodes, 0 pruned nodes, max_depth=8
[397] train-logloss:0.2234 valid-logloss:0.299016
[17:01:11] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 38 extra nodes, 0 pruned nodes, max_depth=8
[398] train-logloss:0.223297 valid-logloss:0.298949
[17:01:15] C:\Users\Administrator\Desktop\xgboost\src\tree\updater_prune.cc:74: tree pruning end,
1 roots, 60 extra nodes, 0 pruned nodes, max_depth=8
[399] train-logloss:0.223162 valid-logloss:0.298909
The test log loss is: 0.2989092178404534
```

In [125]:

```
print("The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```
The test log loss is: 0.2989092178404534
```

OBSERVATION

1. using only tfidf vectorizer gave really good result
2. the log loss decreased to 0.44 using logistic regression
3. by using any random model the loss was 0.88 so the improved results is better by 50%
4. linear svm gave almost same result to that of logistic regression i.e 0.44

4. linear svm gave almost same result to that of logistic regression i.e 0.44
5. xgboost gave fairly better result as it reduced the test log loss to 0.29