ML Pipeline Preparation

Follow the instructions below to help you create your ML pipeline.

1. Import libraries and load data from database.

- Import Python libraries
- Load dataset from database with read_sql_table_(https://pandas.pydata.org/pandas-docs/stable/generated/pandas.read_sql_table.html)
- Define feature and target variables X and Y

In []:			

In [34]:

```
# import libraries
import nltk
nltk.download('punkt')
nltk.download('wordnet')
#nltk.download()
import pandas as pd
from sqlalchemy import create engine
from nltk.tokenize import word_tokenize
from nltk.stem.porter import PorterStemmer
from nltk.stem.wordnet import WordNetLemmatizer
from sklearn.feature extraction.text import TfidfVectorizer
from nltk.corpus import stopwords
from sklearn.ensemble import RandomForestClassifier
from sklearn.model selection import train test split
import re
import numpy as np
from sklearn.metrics import confusion matrix
from sklearn.pipeline import Pipeline, FeatureUnion
from sklearn.feature_extraction.text import CountVectorizer, TfidfTransformer
from sklearn import multioutput
from sklearn.metrics import classification_report
from sklearn.model selection import GridSearchCV
from sklearn.linear model import SGDClassifier
from sklearn.svm import SVC
nltk.download('stopwords')
from self_transformers import StartingVerbExtractor
from sklearn.metrics import fbeta_score, make_scorer
import pickle
import matplotlib.pyplot as plt
%matplotlib inline
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]
              Package punkt is already up-to-date!
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk data]
              Package wordnet is already up-to-date!
[nltk data] Downloading package stopwords to /root/nltk data...
[nltk data]
              Package stopwords is already up-to-date!
[nltk data] Downloading package punkt to /root/nltk data...
[nltk data]
              Package punkt is already up-to-date!
[nltk_data] Downloading package wordnet to /root/nltk_data...
              Package wordnet is already up-to-date!
[nltk data]
[nltk data] Downloading package averaged perceptron tagger to
[nltk data]
                /root/nltk data...
[nltk data]
              Unzipping taggers/averaged perceptron tagger.zip.
```

In [2]:

```
# Load data from database
engine = create_engine('sqlite:///DisasterResponse.db')
df = pd.read_sql ('SELECT * FROM MessagesCategories', engine)
#display (df.head (n=10))
X = df ['message']
y = df.iloc[:,4:]
y.head (n=3)
```

Out[2]:

	related	request	offer	aid_related	medical_help	medical_products	search_and_rescue	Sŧ
0	1	0	0	0	0	0	0	
1	1	0	0	1	0	0	0	
2	1	0	0	0	0	0	0	
3 r	3 rows × 36 columns							
4								>

2. Write a tokenization function to process your text data

In [3]:

```
def tokenize(text):
    """Tokenization function. Receives as input raw text which afterwards normalized, s
top words removed, stemmed and lemmatized.
    Returns tokenized text"""

# Normalize text
text = re.sub(r"[^a-zA-Z0-9]", " ", text.lower())

stop_words = stopwords.words("english")

#tokenize
words = word_tokenize (text)

#stemming
stemmed = [PorterStemmer().stem(w) for w in words]

#lemmatizing
words_lemmed = [WordNetLemmatizer().lemmatize(w) for w in stemmed if w not in stop_words]

return words_lemmed
```

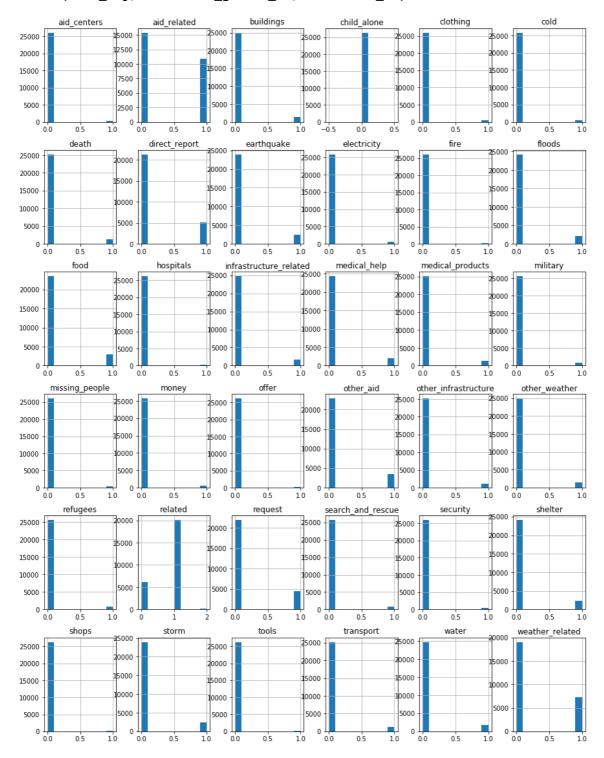
In [4]:

```
#let's take a look to the possible values distribution within classes

#making size of figure bigger
fig = plt.figure(figsize = (15,20))
ax = fig.gca()
y.hist(ax = ax)
plt.show()
```

/opt/conda/lib/python3.6/site-packages/IPython/core/interactiveshell.py:29 61: UserWarning: To output multiple subplots, the figure containing the passed axes is being cleared

exec(code_obj, self.user_global_ns, self.user_ns)



3. Build a machine learning pipeline

This machine pipeline should take in the <code>message</code> column as input and output classification results on the other 36 categories in the dataset. You may find the MultiOutputClassifier (MultiOutputClassifier. helpful for predicting multiple target variables.

In [5]:

In []:

4. Train pipeline

- · Split data into train and test sets
- · Train pipeline

```
In [6]:
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state = 22)
```

In [7]:

```
# train classifier
pipeline.fit(X_train, y_train)
```

Out[7]:

```
Pipeline(memory=None,
    steps=[('vect', CountVectorizer(analyzer='word', binary=False, decode
_error='strict',
    dtype=<class 'numpy.int64'>, encoding='utf-8', input='content',
    lowercase=True, max_df=1.0, max_features=None, min_df=1,
    ngram_range=(1, 1), preprocessor=None, stop_words=None,
    strip...oob_score=False, random_state=None, verbose=0,
        warm_start=False),
    n_jobs=1))])
```

5. Test your model

Report the f1 score, precision and recall for each output category of the dataset. You can do this by iterating through the columns and calling sklearn's classification_report on each.

In [8]:

```
!pip install -U scikit-learn
Collecting scikit-learn
  Downloading https://files.pythonhosted.org/packages/f5/ef/bcd79e8d59250d
6e8478eb1290dc6e05be42b3be8a86e3954146adbc171a/scikit_learn-0.24.2-cp36-cp
36m-manylinux1 x86 64.whl (20.0MB)
    100% |
                                          | 20.0MB 1.6MB/s eta 0:00:01
Requirement already satisfied, skipping upgrade: scipy>=0.19.1 in /opt/con
da/lib/python3.6/site-packages (from scikit-learn) (1.2.1)
Collecting threadpoolctl>=2.0.0 (from scikit-learn)
  Downloading https://files.pythonhosted.org/packages/61/cf/6e354304bcb9c6
413c4e02a747b600061c21d38ba51e7e544ac7bc66aecc/threadpoolctl-3.1.0-py3-non
e-any.whl
Collecting numpy>=1.13.3 (from scikit-learn)
  Downloading https://files.pythonhosted.org/packages/45/b2/6c7545bb7a3875
4d63048c7696804a0d947328125d81bf12beaa692c3ae3/numpy-1.19.5-cp36-cp36m-man
ylinux1_x86_64.whl (13.4MB)
    100%
                                          | 13.4MB 1.7MB/s eta 0:00:01
Requirement already satisfied, skipping upgrade: joblib>=0.11 in /opt/cond
a/lib/python3.6/site-packages (from scikit-learn) (0.11)
tensorflow 1.3.0 requires tensorflow-tensorboard<0.2.0,>=0.1.0, which is n
ot installed.
Installing collected packages: threadpoolctl, numpy, scikit-learn
  Found existing installation: numpy 1.12.1
    Uninstalling numpy-1.12.1:
      Successfully uninstalled numpy-1.12.1
  Found existing installation: scikit-learn 0.19.1
    Uninstalling scikit-learn-0.19.1:
      Successfully uninstalled scikit-learn-0.19.1
Successfully installed numpy-1.19.5 scikit-learn-0.24.2 threadpoolctl-3.1.
```

In [20]:

```
y_pred = pipeline.predict(X_test)
    # print the metrics
category_names = list(df.columns[4:])
for i, col in enumerate(category_names):
    print('{} category metrics: '.format(col))
    print(classification_report(y_test.iloc[:,i], y_pred[:,i]))
```

rela	ted categ	gory metrics:			
		precision	recall	f1-score	support
	0	0.64	0.47	0.54	1541
	1	0.85			
	2	0.36	0.35		
avg ,	/ total	0.79	0.81	0.80	6597
reque	est categ	gory metrics:		C4	
		precision	recall	†1-score	support
	0	0.90	0.98	0.93	5480
	1	0.78	0.44		1117
	_	0.70	0.44	0.30	
avg ,	/ total	0.88	0.88	0.87	6597
offei	r categor	ry metrics:		_	
		precision	recall	f1-score	support
	0	0.99	1.00	1.00	6561
	1	1.00	0.06		36
	1	1.00	0.00	0.11	30
avg	/ total	0.99	0.99	0.99	6597
aid_ı	related o	category metr	rics:		
		precision	recall	f1-score	support
	0	0.75	0.86		
	1	0.75	0.59	0.66	2740
av.a	/ total	0.75	0.75	0.74	6597
avg	COCAI	0.73	0.75	0.74	0557
medi	cal help	category met	trics:		
	- '	precision		f1-score	support
					• •
	0	0.93	0.99	0.96	6083
	1	0.55	0.10	0.17	514
avg ,	/ total	0.90	0.92	0.90	6597
meal	ca1_produ	ucts category			
		precision	recall	+1-score	support
	0	0.96	1.00	0.98	6280
	1	0.79	0.10		
		0.73	0.10	0.17	317
avg	/ total	0.95	0.96	0.94	6597
0 .					
sear	ch_and_re	escue categor	ry metric	s:	
		precision	recall	f1-score	support
	0	0.97	1.00		
	1	0.50	0.04	0.07	182
21/4	/ +0+21	0.06	0.97	0.96	6597
avg	, rorgi	0.96	Ø.97	Ø.96	/פכס
secui	ritv cate	egory metrics	5:		
	-, 55.00	precision		f1-score	support
			·		11
	0	0.98	1.00	0.99	6474

1/23, 8:17 PW				ML Pipeline Pr
1	0.00	0.00	0.00	123
avg / total	0.96	0.98	0.97	6597
military cat	egory metric	cs:		
	0 ,		f1-score	support
0	0.97	1.00	0.98	6374
1	0.68	0.08	0.14	223
avg / total	0.96	0.97	0.95	6597
child alone	category met	trics:		
_	precision		f1-score	support
0	1.00	1.00	1.00	6597
avg / total	1.00	1.00	1.00	6597
water catego	orv metrics:			
	precision	recall	f1-score	support
0	0.96	1.00	0.98	6214
1	0.80	0.31	0.45	383
_	0.00	0.31	0.43	262
avg / total	0.95	0.96	0.95	6597
food categor	ry metrics:			
	precision	recall	f1-score	support
0	0.94	0.99	0.96	5868
1	0.82	0.47	0.59	729
avg / total	0.92	0.93	0.92	6597
shelter cate	egory metrics	5:		
5	0 ,	recall	f1-score	support
	·			
0	0.94	0.99	0.96	6011
1	0.79	0.34	0.48	586
avg / total	0.93	0.93	0.92	6597
7				
clotning cat	egory metrion	recall	f1 ccono	cuppont
	precision	recarr	11-2001.6	support
0	0.99	1.00	0.99	6486
1	0.84	0.19	0.31	111
avg / total	0.98	0.99	0.98	6597
money catego	-	recall	f1-score	support
-	2 22	4 00	0.00	
0	0.98	1.00	0.99	6452
1	0.69	0.06	0.11	145
avg / total	0.97	0.98	0.97	6597
missing neor	ole category	metrics		
=23±118_bcot	precision		f1-score	support

0 1	0.99 0.00	1.00 0.00	1.00 0.00	6533 64
avg / total	0.98	0.99	0.99	6597
refugees cat		.cs: recall	f1-score	support
0 1	0.97 0.62	1.00 0.08	0.98 0.14	6377 220
avg / total	0.96	0.97	0.96	6597
death catego	ory metrics: precision		f1-score	support
0 1	0.96 0.79	1.00 0.16	0.98 0.26	6289 308
avg / total	0.95	0.96	0.95	6597
other_aid ca				
	precision	recall	f1-score	support
0	0.88	0.99	0.93	5744
1	0.47	0.06	0.10	853
avg / total	0.82	0.87	0.82	6597
infrastructu	_			
	precision	recall	f1-score	support
0	0.94 0.12	1.00 0.01	0.97	6207
1	0.12	0.01	0.01	390
avg / total	0.89	0.94	0.91	6597
transport ca				
	precision	recall	f1-score	support
0	0.96	1.00	0.98	6292
1	0.67	0.07	0.12	305
avg / total	0.94	0.96	0.94	6597
buildings ca	ategory metr	ics:		
	precision	recall	f1-score	support
0	0.95		0.98	6266
1	0.75	0.11	0.19	331
avg / total	0.94	0.95	0.94	6597
electricity				
	precision	recall	f1-score	support
0	0.98		0.99	
1	0.73	0.07	0.12	119
avg / total	0.98	0.98	0.98	6597

tools catego	rv metrics:			
5	precision	recall	f1-score	support
0	0.99	1.00	1.00	6562
1	0.00	0.00	0.00	35
avg / total	0.99	0.99	0.99	6597
hospitals ca				
	precision	recall	f1-score	support
0	0.99	1.00	0.99	6530
1	0.00	0.00	0.00	67
avg / total	0.98	0.99	0.98	6597
shops catego	-			
	precision	recall	f1-score	support
0	1.00	1.00	1.00	6572
1	0.00	0.00	0.00	25
avg / total	0.99	1.00	0.99	6597
aid_centers	category me	trics:		
_	precision		f1-score	support
0	0.99	1.00	0.99	6526
1	0.00	0.00	0.00	71
avg / total	0.98	0.99	0.98	6597
other_infras	tructure ca	tegory met	rics:	
	precision	recall	f1-score	support
0	0.96	1.00	0.98	6331
1	0.00	0.00	0.00	266
avg / total	0.92	0.96	0.94	6597
weather_rela	ted categor	y metrics:		
	precision	recall	f1-score	support
0	0.86	0.95	0.91	4755
1	0.84	0.61	0.70	1842
avg / total	0.86	0.86	0.85	6597
floods categ	ory metrics	:		
	precision	recall	f1-score	support
0	0.95	1.00	0.97	6079
1	0.90	0.37	0.53	518
avg / total	0.95	0.95	0.94	6597
storm catego	ry metrics:			
J	precision	recall	f1-score	support
0	0.95	0.98	0.97	5998

23, 8:17 PM				ML Pipeline P
1	0.73	0.46	0.57	599
avg / total	0.93	0.94	0.93	6597
fire categor	-			
	precision	recall	f1-score	support
0	0.99	1.00	0.99	6526
1	1.00	0.01	0.03	71
avg / total	0.99	0.99	0.98	6597
earthquake c	ategory met	trics:		
	precision	recall	f1-score	support
0	0.97		0.98	5990
1	0.89	0.74	0.81	607
avg / total	0.97	0.97	0.97	6597
cold categor	-			
	precision	recall	f1-score	support
0	0.98	1.00	0.99	6458
1	0.80	0.06	0.11	139
avg / total	0.98	0.98	0.97	6597
other_weathe		metrics:		
	precision	recall	f1-score	support
0	0.95	1.00	0.97	6252
1	0.42	0.03	0.05	345
avg / total	0.92	0.95	0.92	6597
direct_repor				
	precision	recall	f1-score	support
0	0.85			5302
1	0.75	0.31	0.44	1295
avg / total	0.83	0.84	0.82	6597

/opt/conda/lib/python3.6/site-packages/sklearn/metrics/classification.py:1 135: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples.

6. Improve your model

Use grid search to find better parameters.

In [24]:

In [25]:

```
model = cv
```

In [26]:

model.fit(X_train, y_train)

```
Fitting 3 folds for each of 8 candidates, totalling 24 fits
[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=10,
vect__max_df=0.75
```

[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=1
0, vect__max_df=0.75, score=0.23040776110353192, total= 1.1min
[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=10, vect max df=0.75

[Parallel(n_jobs=1)]: Done 1 out of 1 | elapsed: 1.5min remaining: 0.0s

[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=1
0, vect__max_df=0.75, score=0.24514857489387507, total= 1.1min
[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=10, vect__max_df=0.75

[Parallel(n_jobs=1)]: Done 2 out of 2 | elapsed: 3.0min remaining:
0.0s

[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=1
0, vect__max_df=0.75, score=0.2427228623408126, total= 1.1min
[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=10, vect max df=1.0

[Parallel(n_jobs=1)]: Done 3 out of 3 | elapsed: 4.4min remaining: 0.0s

[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=1
0, vect__max_df=1.0, score=0.23010459299681674, total= 1.0min
[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=10, vect__max_df=1.0

[Parallel(n_jobs=1)]: Done 4 out of 4 | elapsed: 5.8min remaining:
0.0s

[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=1
0, vect__max_df=1.0, score=0.24620982413583992, total= 1.1min
[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=10, vect__max_df=1.0

[Parallel(n_jobs=1)]: Done 5 out of 5 | elapsed: 7.3min remaining: 0.0s

[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=1
0, vect__max_df=1.0, score=0.24408732565191024, total= 1.1min
[CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=20, vect__max_df=0.75

[Parallel(n_jobs=1)]: Done 6 out of 6 | elapsed: 8.8min remaining:
0.0s

- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=2 0, vect max df=0.75, score=0.24344398969228437, total= 1.5min
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=20,
 vect__max_df=0.75
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=2 0, vect__max_df=0.75, score=0.25651910248635534, total= 1.6min
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=20,
 vect__max_df=0.75
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=2
 0, vect__max_df=0.75, score=0.25272892662219526, total= 1.6min
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=20,
 vect__max_df=1.0
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=2
 0, vect__max_df=1.0, score=0.24283765347885403, total= 1.6min
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=20,
 vect__max_df=1.0
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=2
 0, vect__max_df=1.0, score=0.2548514251061249, total= 1.6min
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=20,
 vect__max_df=1.0
- [CV] clf__estimator__min_samples_split=2, clf__estimator__n_estimators=2
 0, vect__max_df=1.0, score=0.26303820497271074, total= 1.6min
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=10,
 vect__max_df=0.75
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=1 0, vect__max_df=0.75, score=0.23025617705017432, total= 58.1s
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=10,
 vect__max_df=0.75
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=1 0, vect__max_df=0.75, score=0.2348392965433596, total= 1.0min
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=10,
 vect__max_df=0.75
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=1 0, vect__max_df=0.75, score=0.23544572468162522, total= 1.0min
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=10,
 vect__max_df=1.0
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=1
 0, vect__max_df=1.0, score=0.23177201758375018, total= 59.4s
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=10,
 vect__max_df=1.0
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=1 0, vect max df=1.0, score=0.2340812613705276, total= 59.7s
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=10,
 vect__max_df=1.0
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=1
 0, vect__max_df=1.0, score=0.23832625833838691, total= 59.5s
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=20,
 vect__max_df=0.75
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=2 0, vect__max_df=0.75, score=0.23904805214491434, total= 1.4min
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=20,
 vect__max_df=0.75
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=2 0, vect max df=0.75, score=0.25227410551849605, total= 1.4min
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=20, vect__max_df=0.75
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=2 0, vect__max_df=0.75, score=0.25667070952092175, total= 1.4min
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=20,
 vect max df=1.0
- [CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=2

```
0, vect max df=1.0, score=0.23707745945126574, total= 1.5min
[CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=20,
vect max df=1.0
[CV] clf_estimator_min_samples_split=5, clf_estimator_n_estimators=2
0, vect max df=1.0, score=0.25667070952092175, total= 1.4min
[CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=20,
vect__max_df=1.0
[CV] clf__estimator__min_samples_split=5, clf__estimator__n_estimators=2
0, vect max df=1.0, score=0.24787750151607035, total= 1.4min
[Parallel(n jobs=1)]: Done 24 out of 24 | elapsed: 40.6min finished
Out[26]:
GridSearchCV(cv=None, error score='raise',
       estimator=Pipeline(memory=None,
     steps=[('vect', CountVectorizer(analyzer='word', binary=False, decode
_error='strict',
        dtype=<class 'numpy.int64'>, encoding='utf-8', input='content',
        lowercase=True, max_df=1.0, max_features=None, min_df=1,
        ngram_range=(1, 1), preprocessor=None, stop_words=None,
        strip...oob_score=False, random_state=None, verbose=0,
            warm start=False),
           n_jobs=1))]),
       fit_params=None, iid=True, n_jobs=1,
       param_grid={'vect__max_df': (0.75, 1.0), 'clf__estimator__n_estimat
ors': [10, 20], 'clf_estimator_min_samples_split': [2, 5]},
       pre_dispatch='2*n_jobs', refit=True, return_train_score='warn',
       scoring=None, verbose=7)
```

In []:

7. Test your model

Show the accuracy, precision, and recall of the tuned model.

Since this project focuses on code quality, process, and pipelines, there is no minimum performance metric needed to pass. However, make sure to fine tune your models for accuracy, precision and recall to make your project stand out - especially for your portfolio!

In [27]:

```
y_pred_tuned = model.predict(X_test)
#converting to a dataframe
#y_pred_tuned = pd.DataFrame(y_pred_tuned, columns = y_test.columns)

category_names = list(df.columns[4:])
for i, col in enumerate(category_names):
    print('{} category metrics: '.format(col))
    print(classification_report(y_test.iloc[:,i], y_pred_tuned[:,i]))
```

egory metric			
precision	recall	f1-score	support
9 9 71	0 11	0 55	1541
			65
0.50	0.51	0.43	03
0.81	0.82	0.80	6597
• •			
precision	recall	f1-score	support
0.00	0.00	0.04	F490
			5480
0.81	0.46	0.59	1117
0.88	0.89	0.88	6597
ory metrics:			
•		f1-score	support
precision		11 30010	заррог с
0.99	1.00	1.00	6561
			36
0.99	0.99	0.99	6597
d category me	etrics:		
		f1-score	support
0.78	0.85	0.81	3857
			2740
	0.77	0.77	6597
		C 4	
precision	recall	+1-score	support
a a o o o	a 00	0.06	6083
			514
L 0.59	0.09	0.13	514
0.90	0.92	0.90	6597
nducts catego	orv metrics	·	
•	,, ,, ,,,, ,, ,, ,, ,, ,, ,, ,, ,, ,, ,	•	
nrecision	recall		support
precision	recall	f1-score	support
·		f1-score	
precision 0.96 0.69	recall 1.00 0.09		
0.96	1.00	f1-score 0.98	6280
0.96	1.00	f1-score 0.98	6280
0.96 0.69 0.94	1.00 0.09 0.95	f1-score 0.98 0.15 0.94	6280 317
0.96 0.69 0.94 rescue cate	1.00 0.09 0.95 gory metric	f1-score 0.98 0.15 0.94	6280 317 6597
0.96 0.69 0.94 rescue cate	1.00 0.09 0.95	f1-score 0.98 0.15 0.94	6280 317
0.96 0.69 0.94 rescue categ	1.00 0.09 0.95 gory metric recall	f1-score 0.98 0.15 0.94 cs: f1-score	6280 317 6597 support
0.96 0.69 0.94 rescue categorecision 0.97	1.00 0.09 0.95 gory metric recall 1.00	f1-score 0.98 0.15 0.94 es: f1-score 0.99	6280 317 6597 support 6415
0.96 0.69 0.94 rescue categ	1.00 0.09 0.95 gory metric recall	f1-score 0.98 0.15 0.94 cs: f1-score	6280 317 6597 support
0.96 0.69 0.94 rescue categorecision 0.97	1.00 0.09 0.95 gory metric recall 1.00	f1-score 0.98 0.15 0.94 es: f1-score 0.99	6280 317 6597 support 6415
0.96 0.69 0.94 rescue categorecision 0.97 0.55 0.96	1.00 0.09 0.95 gory metric recall 1.00 0.07 0.97	f1-score 0.98 0.15 0.94 es: f1-score 0.99 0.12	6280 317 6597 support 6415 182
0.96 0.69 0.94 rescue cates precision 0.97 0.55 0.96	1.00 0.09 0.95 gory metric recall 1.00 0.07 0.97	f1-score 0.98 0.15 0.94 es: f1-score 0.99 0.12 0.96	6280 317 6597 support 6415 182 6597
0.96 0.69 0.94 rescue cates precision 0.97 0.55 0.96	1.00 0.09 0.95 gory metric recall 1.00 0.07 0.97	f1-score 0.98 0.15 0.94 es: f1-score 0.99 0.12 0.96	6280 317 6597 support 6415 182
0.96 0.69 0.94 rescue cates precision 0.97 0.55 0.96	1.00 0.09 0.95 gory metric recall 1.00 0.07 0.97	f1-score 0.98 0.15 0.94 es: f1-score 0.99 0.12 0.96	6280 317 6597 support 6415 182 6597
	precision 0	precision recall 0 0.71 0.44 1 0.84 0.94 2 0.38 0.51 1 0.81 0.82 1 0.90 0.98 1 0.81 0.46 1 0.88 0.89 2 0.99 1.00 1 0.99 1.00 1 0.99 0.99 2 category metrics: precision recall 2 0.99 0.99 3 category metrics: precision recall 3 0.78 0.85 1 0.76 0.66 1 0.77 0.77 3 category metrics: precision recall 3 0.78 0.85 1 0.76 0.66 3 0.77 0.77 4 category metrics: precision recall 6 0.77 0.77 5 category metrics: precision recall 7 0.79 0.79 6 0.99 0.99 7 0.99	precision recall f1-score 0 0.71 0.44 0.55 1 0.84 0.94 0.89 2 0.38 0.51 0.43 1 0.81 0.82 0.80 1 egory metrics: precision recall f1-score 0 0.90 0.98 0.94 1 0.81 0.46 0.59 1 0.88 0.89 0.88 1 0.89 1.00 1.00 1 1.00 0.06 0.11 1 0.99 0.99 0.99 1 category metrics: precision recall f1-score 0 0.78 0.85 0.81 1 0.76 0.66 0.71 1 0.77 0.77 0.77 1 p category metrics: precision recall f1-score 0 0.93 0.99 0.96 1 0.90 0.92 0.90

### A Composition of the composi	1/23, 8:17 PIVI				ML Pipeline Pr
military category metrics:	1	0.00	0.00	0.00	123
### Precision recall f1-score support ### 0	avg / total	0.96	0.98	0.97	6597
### Precision recall f1-score support ### 0	military cat	tegorv metric	cs:		
avg / total				f1-score	support
avg / total 0.96 0.97 0.96 6597 1.00 1.00 1.00 1.00 1.00 6597 1.00 1.00 1.00 6597 1.00 1.00 1.00 6597 1.00 1.00 1.00 6597 1.00 1.00 6597 1.00 1.00 6597 1.00 1.00 6597 1.00 1.00 6597 1.00 1.00 6597 1.00 1.00 6597 1.00 1.00 6597 1.00 1.00 6597 1.00 1.00 1.00 6597 1.00 1.00 1.00 1.00 6597 1.00 1	0	0.97	1.00	0.98	6374
child_alone category metrics: precision recall f1-score support 0 1.00 1.00 1.00 6597 avg / total 1.00 1.00 1.00 6597 water category metrics: precision recall f1-score support 0 0.96 1.00 0.98 6214 1 0.88 0.32 0.47 383 avg / total 0.96 0.96 0.95 6597 food category metrics: precision recall f1-score support 0 0.94 0.99 0.96 5868 1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.96 5868 1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6611 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 0.99 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	1	0.73	0.09	0.15	223
### Precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision preci	avg / total	0.96	0.97	0.96	6597
### Precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision recall f1-score support ### of the category metrics: precision preci	child alone	category met	trics:		
avg / total 1.00 1.00 1.00 6597 water category metrics: precision recall f1-score support 0 0.96 1.00 0.98 6214 1 0.88 0.32 0.47 383 avg / total 0.96 0.96 0.95 6597 food category metrics: precision recall f1-score support 0 0.94 0.99 0.96 5868 1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 <td< td=""><td>_</td><td></td><td></td><td>f1-score</td><td>support</td></td<>	_			f1-score	support
avg / total 1.00 1.00 1.00 6597 water category metrics: precision recall f1-score support 0 0.96 1.00 0.98 6214 1 0.88 0.32 0.47 383 avg / total 0.96 0.96 0.95 6597 food category metrics: precision recall f1-score support 0 0.94 0.99 0.96 5868 1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 <td< td=""><td></td><td></td><td></td><td></td><td></td></td<>					
water category metrics:	0	1.00	1.00	1.00	6597
## Precision recall f1-score support ## 0 0.96 1.00 0.98 6214 ## 1 0.88 0.32 0.47 383 ## 383 avg / total 0.96 0.96 0.95 6597 ## 500d category metrics:	avg / total	1.00	1.00	1.00	6597
## Precision recall f1-score support ## 0 0.96 1.00 0.98 6214 ## 1 0.88 0.32 0.47 383 ## 383 avg / total 0.96 0.96 0.95 6597 ## 500d category metrics:	water catego	orv metrics:			
1 0.88 0.32 0.47 383 avg / total 0.96 0.96 0.95 6597 food category metrics: precision recall f1-score support 0 0.94 0.99 0.96 5868 1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	water catego	-	recall	f1-score	support
1 0.88 0.32 0.47 383 avg / total 0.96 0.96 0.95 6597 food category metrics: precision recall f1-score support 0 0.94 0.99 0.96 5868 1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	0	0.06	1 00	0.00	624.4
avg / total 0.96 0.96 0.95 6597 food category metrics: precision recall f1-score support 0 0.94 0.99 0.96 5868 1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10					
food category metrics:	1	0.00	0.32	0.47	565
precision recall f1-score support 0 0.94 0.99 0.96 5868 1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	avg / total	0.96	0.96	0.95	6597
precision recall f1-score support 0 0.94 0.99 0.96 5868 1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	food categor	rv metrics:			
1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics: support 0.99 0.99 0.99 0.99	J	-	recall	f1-score	support
1 0.81 0.53 0.64 729 avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics: support 0.99 0.99 0.99 0.99	a	0 91	a 99	a 96	5868
avg / total 0.93 0.93 0.93 6597 shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:					
shelter category metrics: precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics: category metrics: category metrics: category metrics:		0.00	0.00		
precision recall f1-score support 0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	avg / total	0.93	0.93	0.93	6597
0 0.94 0.99 0.97 6011 1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	shelter cate	egory metrics	s:		
1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:		precision	recall	f1-score	support
1 0.83 0.41 0.55 586 avg / total 0.93 0.94 0.93 6597 clothing category metrics: precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	0	0.94	0.99	0.97	6011
clothing category metrics:	1	0.83	0.41	0.55	586
precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	avg / total	0.93	0.94	0.93	6597
precision recall f1-score support 0 0.99 1.00 0.99 6486 1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	clothing cat	togony motni			
1 0.88 0.13 0.22 111 avg / total 0.98 0.98 0.98 6597 money category metrics: precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	crothing car			f1-score	support
avg / total 0.98 0.98 0.98 6597 money category metrics: recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	0	0.99	1.00	0.99	6486
money category metrics: recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	1		0.13	0.22	111
precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	avg / total	0.98	0.98	0.98	6597
precision recall f1-score support 0 0.98 1.00 0.99 6452 1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	money catego	orv metrics:			
1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	morrey eacego	-	recall	f1-score	support
1 0.73 0.06 0.10 145 avg / total 0.97 0.98 0.97 6597 missing_people category metrics:	0	0.98	1.00	0.99	6452
missing_people category metrics:					
	avg / total	0.97	0.98	0.97	6597
	•	-1			
·	missing_peop	precision		f1-score	support

0 1	0.99 0.00	1.00 0.00	1.00 0.00	6533 64
avg / total	0.98	0.99	0.99	6597
refugees cat		.cs: recall	f1-score	support
0 1	0.97 0.62	1.00 0.07	0.98 0.12	6377 220
avg / total	0.96	0.97	0.95	6597
death catego	ory metrics: precision		f1-score	support
0 1	0.96 0.78	1.00 0.15	0.98 0.26	6289 308
avg / total	0.95	0.96	0.94	6597
other_aid ca			5 4	
	precision		f1-score	support
0 1	0.88 0.56	0.99 0.04	0.93 0.08	5744 853
avg / total	0.83	0.87	0.82	6597
infrastructu				
	precision	recall	f1-score	support
0 1	0.94 0.08	1.00 0.00	0.97 0.00	6207 390
avg / total	0.89	0.94	0.91	6597
transport ca	ategorv metr	ics:		
C. GSpo. C Co		recall	f1-score	support
0 1	0.96 0.70	1.00 0.10	0.98 0.18	6292 305
avg / total	0.95	0.96	0.94	6597
buildings ca		rics: recall	f1-score	support
0	0.96 0.82		0.98 0.28	6266 331
avg / total	0.95	0.96	0.94	6597
electricity	category me	etrics:		
		recall	f1-score	support
0 1	0.98 0.73	1.00 0.07	0.99 0.12	6478 119
avg / total	0.98	0.98	0.98	6597

tools catego	orv metrics:			
J	precision	recall	f1-score	support
0	0.99	1.00	1.00	6562
1	0.00	0.00	0.00	35
avg / total	0.99	0.99	0.99	6597
hospitals ca	• .			
	precision	recall	f1-score	support
0	0.99	1.00	0.99	6530
1	0.00	0.00	0.00	67
avg / total	0.98	0.99	0.98	6597
shops catego	-			
	precision	recall	f1-score	support
0	1.00	1.00	1.00	6572
1	0.00	0.00	0.00	25
avg / total	0.99	1.00	0.99	6597
aid_centers	category me	trics:		
_	precision		f1-score	support
0	0.99	1.00	0.99	6526
1	0.00	0.00	0.00	71
avg / total	0.98	0.99	0.98	6597
other_infras	tructure ca	tegory met	rics:	
	precision	recall	f1-score	support
0	0.96	1.00	0.98	6331
1	0.00	0.00	0.00	266
avg / total	0.92	0.96	0.94	6597
weather_rela	ted categor	y metrics:		
	precision	recall	f1-score	support
0	0.88	0.96	0.92	4755
1	0.86	0.66	0.74	1842
avg / total	0.87	0.87	0.87	6597
floods categ	ory metrics	:		
	precision	recall	f1-score	support
0	0.95	1.00	0.97	6079
1	0.91	0.44	0.59	518
avg / total	0.95	0.95	0.94	6597
storm catego	ry metrics:			
3	precision	recall	f1-score	support
0	0.94	0.99	0.96	5998

/23, 8:17 PM				ML Pipeline P
1	0.74	0.38	0.50	599
avg / total	0.92	0.93	0.92	6597
fire categor	-			
	precision	recall	f1-score	support
0	0.99		0.99	6526
1	0.67	0.03	0.05	71
avg / total	0.99	0.99	0.98	6597
earthquake c				
	precision	recall	f1-score	support
0	0.98		0.98	5990
1	0.90	0.76	0.82	607
avg / total	0.97	0.97	0.97	6597
cold categor	y metrics:			
	precision	recall	f1-score	support
0	0.98	1.00	0.99	6458
1	0.83	0.07	0.13	139
avg / total	0.98	0.98	0.97	6597
other_weathe				
	precision	recall	f1-score	support
0	0.95	1.00	0.97	6252
1	0.50	0.03	0.05	345
avg / total	0.93	0.95	0.93	6597
direct_repor				
	precision	recall	f1-score	support
0	0.86		0.91	5302
1	0.78	0.34	0.47	1295
avg / total	0.84	0.85	0.83	6597

/opt/conda/lib/python3.6/site-packages/sklearn/metrics/classification.py:1 135: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples.

8. Try improving your model further. Here are a few ideas:

- try other machine learning algorithms
- · add other features besides the TF-IDF

In [35]:

Out[35]:

```
Pipeline(memory=None,
    steps=[('features', FeatureUnion(n_jobs=1,
        transformer_list=[('text_pipeline', Pipeline(memory=None,
    steps=[('vect', CountVectorizer(analyzer='word', binary=False, decode
    _error='strict',
        dtype=<class 'numpy.int64'>, encoding='utf-8', input='content',
        lowercase=True, max_d...oob_score=False, random_state=None, verbos
e=0,
        warm_start=False),
        n_jobs=1))])
```

In [36]:

```
y_pred_upd = upd_pipeline.predict(X_test)
#converting to a dataframe
#y_pred_tuned = pd.DataFrame(y_pred_tuned, columns = y_test.columns)

category_names = list(df.columns[4:])
for i, col in enumerate(category_names):
    print('{} category metrics: '.format(col))
    print(classification_report(y_test.iloc[:,i], y_pred_upd[:,i]))
```

related	_	ry metrics			
	р	recision	recall	f1-score	support
	0	0.65	0.46	0.54	1541
	1			0.88	
	2	0.37			65
	2	0.57	0.43	0.40	03
avg / t	otal	0.79	0.81	0.80	6597
noguest	sat a sa	nu motoice			
request		ry metrics		£1 ccopo	cuppont
	Þ	recision	recarr	T1-Score	Support
	0	0.90	0.97	0.93	5480
	1	0.78	0.45	0.57	
	_				
avg / t	otal	0.88	0.89	0.87	6597
-CC					
orrer C		metrics:	maca11	£1 ccopo	cuppost
	þ	recision	recarr	T1-Score	Support
	0	0.99	1.00	1.00	6561
	1	1.00			
	-	2.00	0.00	3.11	50
avg / t	otal	0.99	0.99	0.99	6597
aid_rel		tegory met			
	р	recision	recall	f1-score	support
	0	0.75	0.00	0.00	2057
	0			0.80	
	1	0.75	0.60	0.66	2740
avg / t	otal	0.75	0.75	0.74	6597
medical	_help c	ategory me	trics:		
	р	recision	recall	f1-score	support
	0	0.93	0.99		
	1	0.56	0.12	0.19	514
, ,		0.00	0.00	0.00	6507
avg / to	отат	0.90	0.92	0.90	6597
medical	produc	ts category	v metrics	:	
		recision			support
	0	0.96	1.00	0.98	6280
	1			0.18	
avg / t	otal	0.94	0.95	0.94	6597
search_		cue catego			
	р	recision	recall	t1-score	support
	0	0.97	1 00	0.99	6415
	1	0.50			
	-	0.50	0.00	0.00	102
avg / to	otal	0.96	0.97	0.96	6597
J , -					
securit	-	ory metric			
	р	recision	recall	f1-score	support
	_				د
	0	0.98	1.00	0.99	6474

. /	20, 0.17 1 10				ME i ipellile i i				
	1	0.00	0.00	0.00	123				
	avg / total	0.96	0.98	0.97	6597				
	military category metrics:								
	military ca			f1-score	support				
	0	0.97	1.00	0.98	6374				
	1	0.58	0.07	0.12	223				
	avg / total	0.96	0.97	0.95	6597				
	child alone	category met	trics:						
	_	precision		f1-score	support				
	0	1.00	1.00	1.00	6597				
	ava / +a+a1	1 00	1 00	1 00	6507				
	avg / total	1.00	1.00	1.00	6597				
	water categ	ory metrics:							
	water catego	precision	recall	f1-score	support				
	0	0.96	1.00	0.98	6214				
	1	0.82	0.27	0.40	383				
	avg / total	0.95	0.95	0.94	6597				
	food categor	rv metrics:							
	Toda cacego	precision	recall	f1-score	support				
	0	0.94	0.99	0.96	5868				
	1	0.85	0.47	0.60	729				
	1	0.83	0.47	0.00	729				
	avg / total	0.93	0.93	0.92	6597				
	shelter cate	egory metrics	s:						
	51102001 000	precision		f1-score	support				
	0	2.04	0.00	0.06	C011				
	0	0.94	0.99	0.96	6011				
	1	0.80	0.34	0.47	586				
	avg / total	0.93	0.93	0.92	6597				
	clothing co	tegory metric							
	CIOCHIIII Ca	•		C4					
		precision	recall	+1-score	support				
	0	0.99	1.00	0.99	6486				
	1	0.90	0.17	0.29	111				
	avg / total	0.98	0.99	0.98	6597				
	money category metrics:								
	money cacego	precision	recall	f1-score	support				
	0	0.98	1.00	0.99	6452				
	1	0.70	0.05	0.09	145				
	1	0.70	6.63	0.03	143				
	avg / total	0.97	0.98	0.97	6597				
	• •	-1							
	missing_peop	ple category precision		f1-score	support				

0 1	0.99 0.00	1.00 0.00	1.00 0.00	6533 64		
avg / total	0.98	0.99	0.99	6597		
refugees cat		cs: recall	f1-score	support		
0 1	0.97 0.78	1.00 0.06	0.98 0.12	6377 220		
avg / total	0.96	0.97	0.95	6597		
death catego	ory metrics: precision		f1-score	support		
0 1	0.96 0.93	1.00 0.16	0.98 0.28	6289 308		
avg / total	0.96	0.96	0.95	6597		
other_aid ca			C1			
	precision		f1-score	support		
0 1	0.88 0.56	0.99 0.08	0.93 0.13	5744 853		
avg / total	0.84	0.87	0.83	6597		
infrastructu						
	p. cc1310		f1-score	support		
0 1	0.94 0.07	1.00 0.00	0.97 0.00	6207 390		
avg / total	0.89	0.94	0.91	6597		
transport category metrics:						
	precision	recall	†1-score	support		
0 1	0.96 0.70	1.00 0.09	0.98 0.15	6292 305		
avg / total		0.96	0.94	6597		
buildings ca						
	precision	recall	f1-score	support		
0 1	0.96 0.71	1.00 0.12	0.98 0.20	6266 331		
avg / total	0.94	0.95	0.94	6597		
electricity category metrics: precision recall f1-score support						
0 1	0.98 0.67	1.00 0.05	0.99 0.09	6478 119		
avg / total	0.98	0.98	0.97	6597		

tools category metrics:						
J	precision	recall	f1-score	support		
0	0.99	1.00	1.00	6562		
1	0.00	0.00	0.00	35		
avg / total	0.99	0.99	0.99	6597		
hospitals ca						
	precision	recall	f1-score	support		
0	0.99	1.00	0.99	6530		
1	0.00	0.00	0.00	67		
avg / total	0.98	0.99	0.98	6597		
shops catego	-					
	precision	recall	f1-score	support		
0	1.00	1.00	1.00	6572		
1	0.00	0.00	0.00	25		
avg / total	0.99	1.00	0.99	6597		
aid_centers	category me	trics:				
	precision	recall	f1-score	support		
0	0.99	1.00	0.99	6526		
1	0.00	0.00	0.00	71		
avg / total	0.98	0.99	0.98	6597		
other_infras	tructure ca	tegory met	rics:			
	precision	recall	f1-score	support		
0	0.96	1.00	0.98	6331		
1	0.17	0.00	0.01	266		
avg / total	0.93	0.96	0.94	6597		
weather_rela	ted categor	y metrics:				
	precision	recall	f1-score	support		
0	0.86	0.95	0.90	4755		
1	0.83	0.60	0.70	1842		
avg / total	0.85	0.85	0.85	6597		
floods category metrics:						
	precision	recall	f1-score	support		
0	0.95	1.00	0.97	6079		
1	0.90	0.39	0.55	518		
avg / total	0.95	0.95	0.94	6597		
storm catego	ry metrics:					
J	precision	recall	f1-score	support		
0	0.94	0.99	0.96	5998		

0.73	0.36	0.48	F00
		01.0	599
0.92	0.93	0.92	6597
y metrics:			
-	recall	f1-score	support
0.99	1.00	0.99	6526
0.75	0.04	0.08	71
0.99	0.99	0.98	6597
ategory met	rics:		
		f1-score	support
0.97	0.99	0.98	5990
0.89	0.69	0.77	607
0.96	0.96	0.96	6597
y metrics:			
	recall	f1-score	support
0.98	1.00	0.99	6458
0.77	0.12	0.21	139
0.98	0.98	0.97	6597
r category	metrics:		
		f1-score	support
0.95	1.00	0.97	6252
0.44	0.06	0.10	345
0.92	0.95	0.93	6597
t category	metrics:		
precision		f1-score	support
0.85	0.97	0.91	5302
0.73	0.32	0.45	1295
0.83	0.84	0.82	6597
	y metrics: precision 0.99 0.75 0.99 ategory metrics: precision 0.96 y metrics: precision 0.98 0.77 0.98 r category precision 0.95 0.44 0.92 t category precision 0.85 0.73	y metrics: precision	y metrics: precision recall f1-score 0.99 1.00 0.99 0.75 0.04 0.08 0.99 0.99 0.99 0.98 ategory metrics: precision recall f1-score 0.97 0.99 0.98 0.89 0.69 0.77 0.96 0.96 0.96 y metrics: precision recall f1-score 0.98 1.00 0.99 0.77 0.12 0.21 0.98 0.98 0.97 r category metrics: precision recall f1-score 0.95 1.00 0.97 0.44 0.06 0.10 0.92 0.95 0.93 t category metrics: precision recall f1-score 0.85 0.97 0.91 0.73 0.32 0.45

/opt/conda/lib/python3.6/site-packages/sklearn/metrics/classification.py:1 135: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples.

9. Export your model as a pickle file

```
In [37]:
```

```
pickle.dump(model, open('final_model.sav', 'wb'))
```

10. Use this notebook to complete train.py

Use the template file attached in the Resources folder to write a script that runs the steps above to create a database and export a model based on a new dataset specified by the user.

In []:			