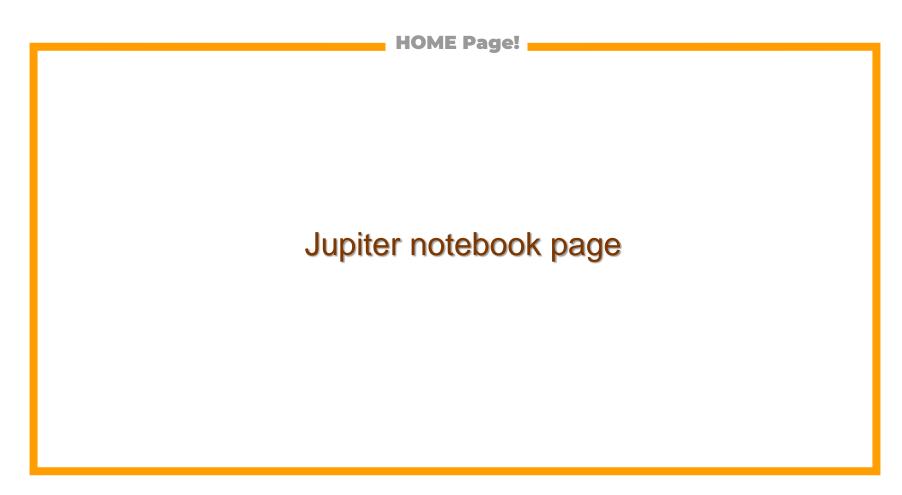
News Articles Sorting

Wireframe Documentation



```
In [9]: import pandas as pd #pandas Library for data frame
 datatrain=pd.read csv(r"C:\Users\welcome\Desktop\BBC News Train.csv") #traindata
 datatest=pd.read csv(r"C:\Users\welcome\Desktop\BBC News Test.csv") #test data
 import re
 import nltk
 nltk.download('stopwords')
 from nltk.corpus import stopwords
 from nltk.stem.porter import PorterStemmer
 corpus = []
 corpust = []
 for i in range(0, 1490):
                                                                   #makina traindata
  text = re.sub('[^a-zA-Z]', ' ', datatrain['Text'][i])
  text = text.lower()
  text = text.split()
  ps = PorterStemmer()
  all stopwords = stopwords.words('english')
  all stopwords.remove('not')
  text = [ps.stem(word) for word in text if not word in set(all stopwords)]
  text = ' '.join(text)
  corpus.append(text)
 for i in range(0, 735):
                                                                  #makina testdata te.
  textt = re.sub('[^a-zA-Z]', ' ', datatest['Text'][i])
  textt = textt.lower()
  textt = textt.split()
  ps = PorterStemmer()
  all_stopwords = stopwords.words('english')
  all stopwords.remove('not')
  textt = [ps.stem(word) for word in textt if not word in set(all_stopwords)]
  textt = ' '.ioin(textt)
  corpust.append(textt)
 from sklearn.feature extraction.text import CountVectorizer #feature extraction
 cv = CountVectorizer()
X = cv.fit_transform(corpus).toarray() #fit and transform features in train data
                                            # transform features of test data
 x = cv.transform(corpust).toarray()
y = datatrain.iloc[:,-1].values
 from sklearn.preprocessing import LabelEncoder # Label for category
 a=LabelEncoder()
Y=a.fit transform(v)
 from sklearn.linear model import LogisticRegression #logistic regression for class
 lg=LogisticRegression(random state=0).fit(X,Y)
 e=lg.predict(x)
 print(a.inverse transform(e)) #predicting
 # datatest["Pr"]=a.inverse_transform(e) # prediction of test dataset goes to Finalt.
 # datatest.to csv("Finalt.csv".index=False)
```

| | | C2 | ▼ Jx sport | |
|---------------------|----|-----------|---|---------------|
| | 4 | Α | В | С |
| Predicted test file | 1 | ArticleId | Text | Category |
| | 2 | 1018 | qpr keeper day heads for preston queens park rang | sport |
| | 3 | 1319 | software watching while you work software that ca | tech |
| | 4 | 1138 | d arcy injury adds to ireland woe gordon d arcy has | sport |
| | 5 | 459 | india s reliance family feud heats up the ongoing p | business |
| | 6 | 1020 | boro suffer morrison injury blow middlesbrough m | sport |
| | 7 | 51 | lewsey puzzle over disallowed try england s josh le | sport |
| | 8 | 2025 | blair blasts tory spending plans tony blair has laund | politics |
| | 9 | 1479 | former ni minister scott dies former northern irela | politics |
| | 10 | 27 | career honour for actor dicaprio actor leonardo dica | entertainment |
| | 11 | 397 | tsunami to hit sri lanka banks sri lanka s banks face | business |
| | 12 | 1644 | us economy still growing says fed most areas of the | business |
| | 13 | 263 | digital uk driven by net and tv the uk s adoption of | tech |
| | 14 | 765 | blunkett tells of love and pain david blunkett has s | politics |
| | 15 | 2134 | ibm puts cash behind linux push ibm is spending \$3 | tech |
| | 16 | 297 | cage film s third week at us top nicolas cage movie | entertainment |
| | 17 | 1712 | souness backs smith for scotland graeme souness b | sport |
| | 18 | 1631 | ukip mep attacked german empire a uk independe | politics |
| | 19 | 942 | cheaper chip for mobiles a mobile phone chip which | tech |
| | 20 | 1549 | usher leads soul train shortlist chart-topping r&b st | entertainment |
| | 21 | 516 | india to deport bollywood actress india has ordere | entertainment |
| | 22 | 2215 | bush budget seeks deep cutbacks president bush h | business |
| | 23 | 531 | labour s core support takes stock tony blair has told | politics |
| | 24 | 1541 | hodges announces rugby retirement scarlets and u | sport |
| | 25 | 1340 | latin america sees strong growth latin america s ec | business |



Vaibhav Chopra