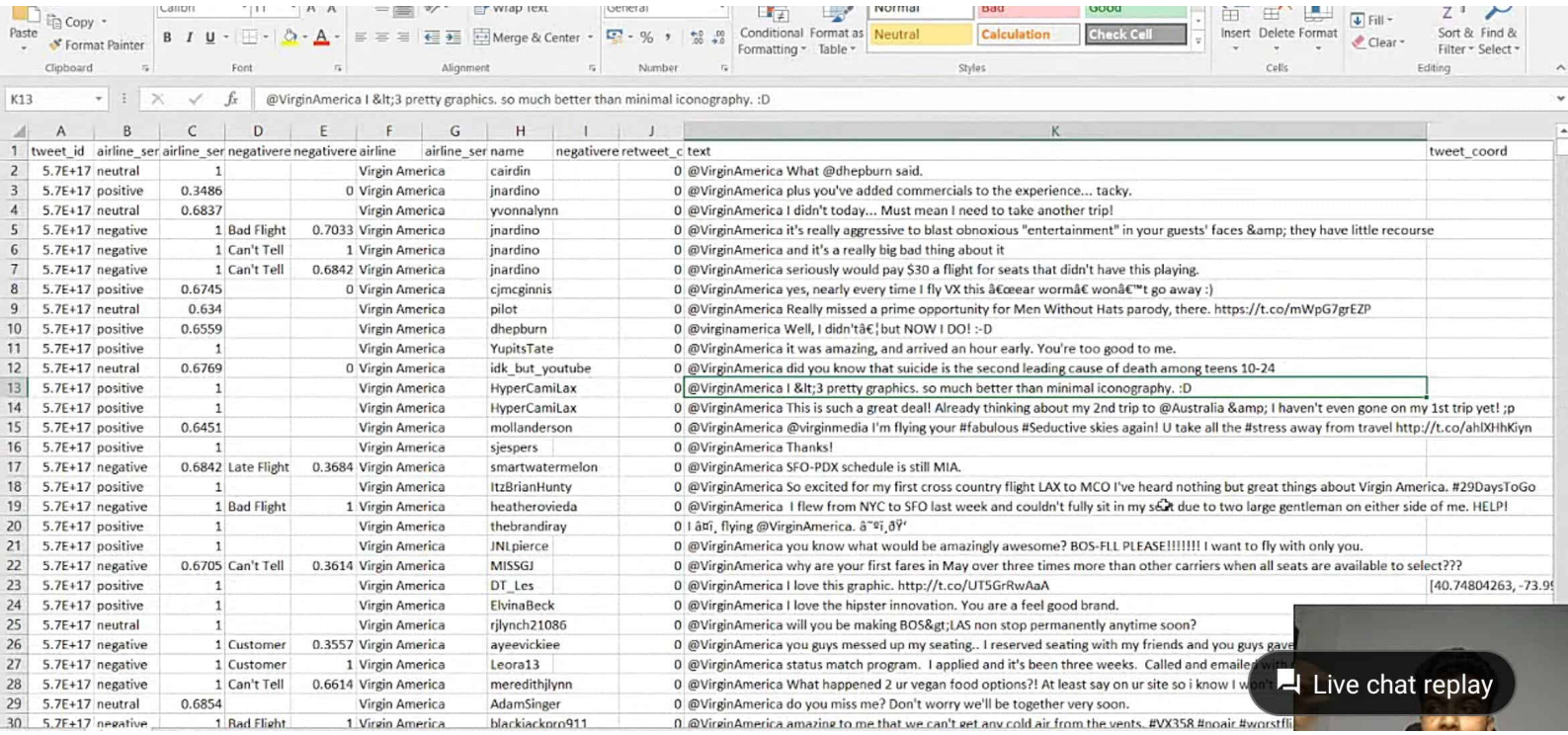
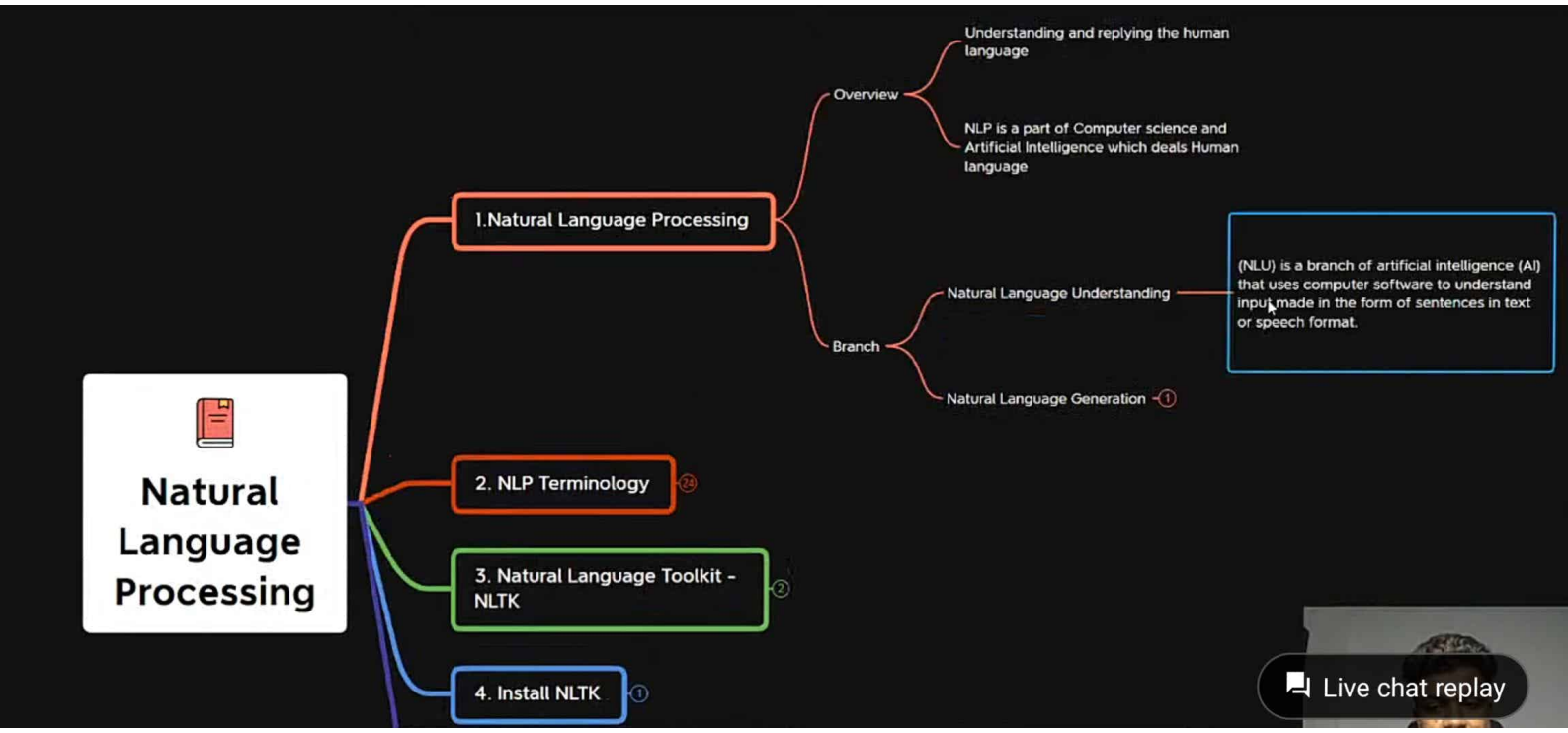
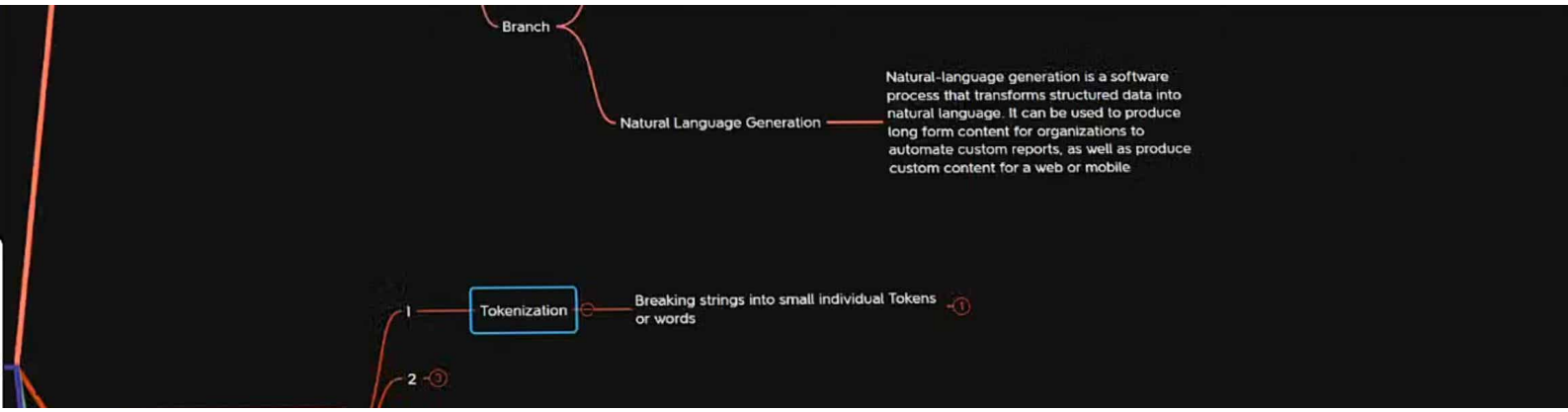
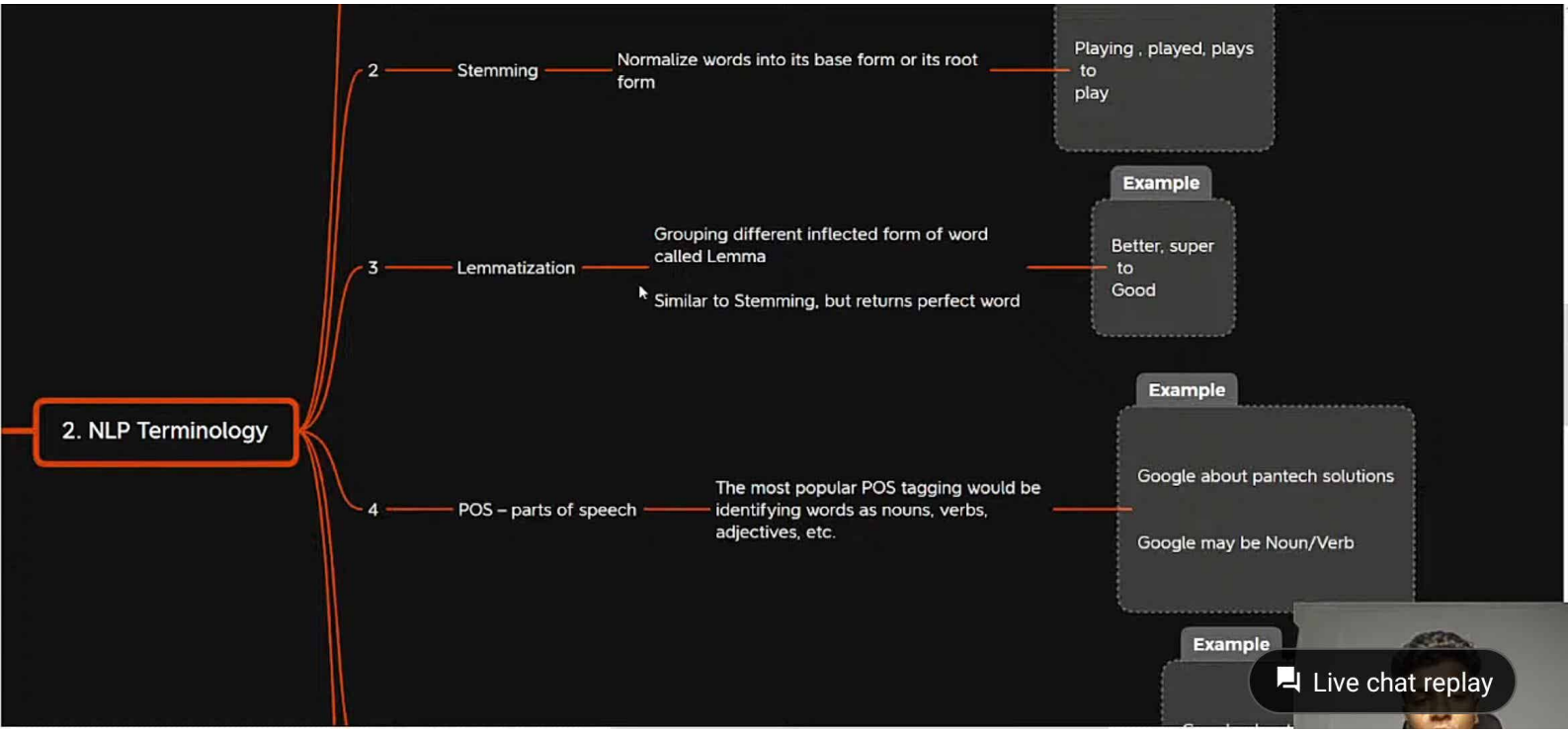
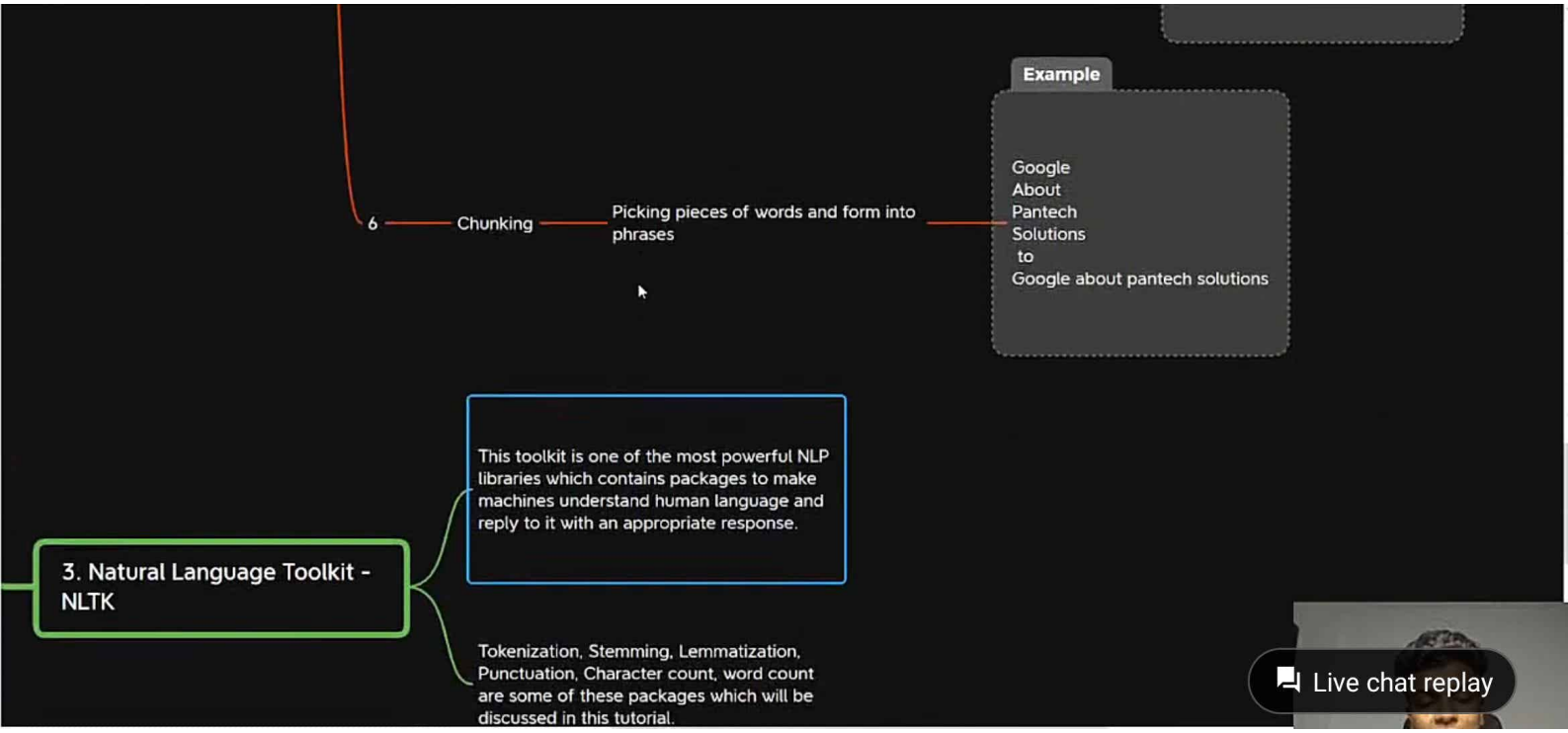


	A	B	C	D	E	F	G	H	I	J	K	L
1	tweet_id	airline_ser	airline_ser	negativere	negativere	airline	airline_ser	name	negativere	retweet_c	text	tweet_coord
2	5.7E+17	neutral	1			Virgin America		cairdin		0	@VirginAmerica What @dhepburn said.	
3	5.7E+17	positive	0.3486		0	Virgin America		jnardino		0	@VirginAmerica plus you've added commercials to the experience... tacky.	
4	5.7E+17	neutral	0.6837			Virgin America		yvonnalynn		0	@VirginAmerica I didn't today... Must mean I need to take another trip!	
5	5.7E+17	negative	1	Bad Flight	0.7033	Virgin America		jnardino		0	@VirginAmerica it's really aggressive to blast obnoxious "entertainment" in your guests' faces & they have little recourse	
6	5.7E+17	negative	1	Can't Tell	1	Virgin America		jnardino		0	@VirginAmerica and it's a really big bad thing about it	
7	5.7E+17	negative	1	Can't Tell	0.6842	Virgin America		jnardino		0	@VirginAmerica seriously would pay $30 a flight for seats that didn't have this playing.	
8	5.7E+17	positive	0.6745		0	Virgin America		cjmcginnis		0	@VirginAmerica yes, nearly every time I fly VX this â€œear wormâ€ wonâ€™t go away :)	
9	5.7E+17	neutral	0.634			Virgin America		pilot		0	@VirginAmerica Really missed a prime opportunity for Men Without Hats parody, there. https://t.co/mWpG7grEZP	
10	5.7E+17	positive	0.6559			Virgin America		dhepburn		0	@virginamerica Well, I didn'tâ€¦ but NOW I DO! :-D	
11	5.7E+17	positive	1			Virgin America		YupitsTate		0	@VirginAmerica it was amazing, and arrived an hour early. You're too good to me.	
12	5.7E+17	neutral	0.6769		0	Virgin America		idk_but_youtube		0	@VirginAmerica did you know that suicide is the second leading cause of death among teens 10-24	
13	5.7E+17	positive	1			Virgin America		HyperCamiLax		0	@VirginAmerica I <3 pretty graphics. so much better than minimal iconography. :D	
14	5.7E+17	positive	1			Virgin America		HyperCamiLax		0	@VirginAmerica This is such a great deal! Already thinking about my 2nd trip to @Australia & I haven't even gone on my 1st trip yet! ;p	
15	5.7E+17	positive	0.6451			Virgin America		mollanderson		0	@VirginAmerica @virginmedia I'm flying your #fabulous #Seductive skies again! U take all the #stress away from travel http://t.co/ahlXHhKiyn	
16	5.7E+17	positive	1			Virgin America		sjespers		0	@VirginAmerica Thanks!	
17	5.7E+17	negative	0.6842	Late Flight	0.3684	Virgin America		smartwatermelon		0	@VirginAmerica SFO-PDX schedule is still MIA.	
18	5.7E+17	positive	1			Virgin America		ItzBrianHunty		0	@VirginAmerica So excited for my first cross country flight LAX to MCO I've heard nothing but great things about Virgin America. #29DaysToGo	
19	5.7E+17	negative	1	Bad Flight	1	Virgin America		heatherovieda		0	@VirginAmerica I flew from NYC to SFO last week and couldn't fully sit in my seat due to two large gentleman on either side of me. HELP!	
20	5.7E+17	positive	1			Virgin America		thebrandiray		0	I âū¡ flying @VirginAmerica. â˜¤ï¸ğŸŸ	
21	5.7E+17	positive	1			Virgin America		JNLpierce		0	@VirginAmerica you know what would be amazingly awesome? BOS-FLL PLEASE!!!!!!! I want to fly with only you.	
22	5.7E+17	negative	0.6705	Can't Tell	0.3614	Virgin America		MISSGJ		0	@VirginAmerica why are your first fares in May over three times more than other carriers when all seats are available to select???	
23	5.7E+17	positive	1			Virgin America		DT_Les		0	@VirginAmerica I love this graphic. http://t.co/UT5GrRwAaA	[40.74804263, -73.9
24	5.7E+17	positive	1			Virgin America		ElvinaBeck		0	@VirginAmerica I love the hipster innovation. You are a feel good brand.	
25	5.7E+17	neutral	1			Virgin America		rjlynch21086		0	@VirginAmerica will you be making BOS>LAS non stop permanently anytime soon?	
26	5.7E+17	negative	1	Customer	0.3557	Virgin America		ayeevickiee		0	@VirginAmerica you guys messed up my seating.. I reserved seating with my friends and you guys gave	
27	5.7E+17	negative	1	Customer	1	Virgin America		Leora13		0	@VirginAmerica status match program. I applied and it's been three weeks. Called and emailed w fl	
28	5.7E+17	negative	1	Can't Tell	0.6614	Virgin America		meredithjlynn		0	@VirginAmerica What happened 2 ur vegan food options?! At least say on ur site so i know I won't	
29	5.7E+17	neutral	0.6854			Virgin America		AdamSinger		0	@VirginAmerica do you miss me? Don't worry we'll be together very soon.	
30	5.7E+17	negative	1	Bad Flight	1	Virgin America		blackjackpro911		0	@VirginAmerica amazing to me that we can't get any cold air from the vents. #VX358 #noair #worstfli	


Natural Language Processing
1.Natural Language Processing
Overview
Understanding and replying the human language
NLP is a part of Computer science and Artificial Intelligence which deals Human language
Branch
Natural Language Understanding
(NLU) is a branch of artificial intelligence (AI) that uses computer software to understand input made in the form of sentences in text or speech format.
Natural Language Generation
2. NLP Terminology
3. Natural Language Toolkit - NLTK
4. Install NLTK
Live chat replay


Branch
Natural Language Generation
Natural-language generation is a software process that transforms structured data into natural language. It can be used to produce long form content for organizations to automate custom reports, as well as produce custom content for a web or mobile
1
Tokenization
Breaking strings into small individual Tokens or words
2

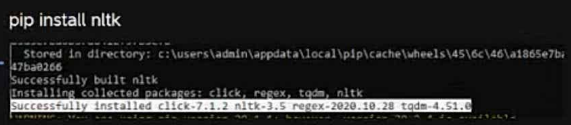

2. NLP Terminology
2
Stemming
Normalize words into its base form or its root form
Playing , played, plays to play
Example
3
Lemmatization
Grouping different inflected form of word called Lemma
Similar to Stemming, but returns perfect word
Better, super to Good
Example
4
POS – parts of speech
The most popular POS tagging would be identifying words as nouns, verbs, adjectives, etc.
Google about pantech solutions
Google may be Noun/Verb
Example
Live chat replay


Example
6
Chunking
Picking pieces of words and form into phrases
Google
About
Pantech
Solutions
to
Google about pantech solutions
This toolkit is one of the most powerful NLP libraries which contains packages to make machines understand human language and reply to it with an appropriate response.
3. Natural Language Toolkit - NLTK
Tokenization, Stemming, Lemmatization, Punctuation, Character count, word count are some of these packages which will be discussed in this tutorial.
Live chat replay

Tokenization, Stemming, Lemmatization, Punctuation, Character count, word count are some of these packages which will be discussed in this tutorial.

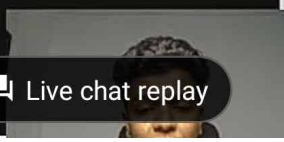
4. Install NLTK

pip install nltk

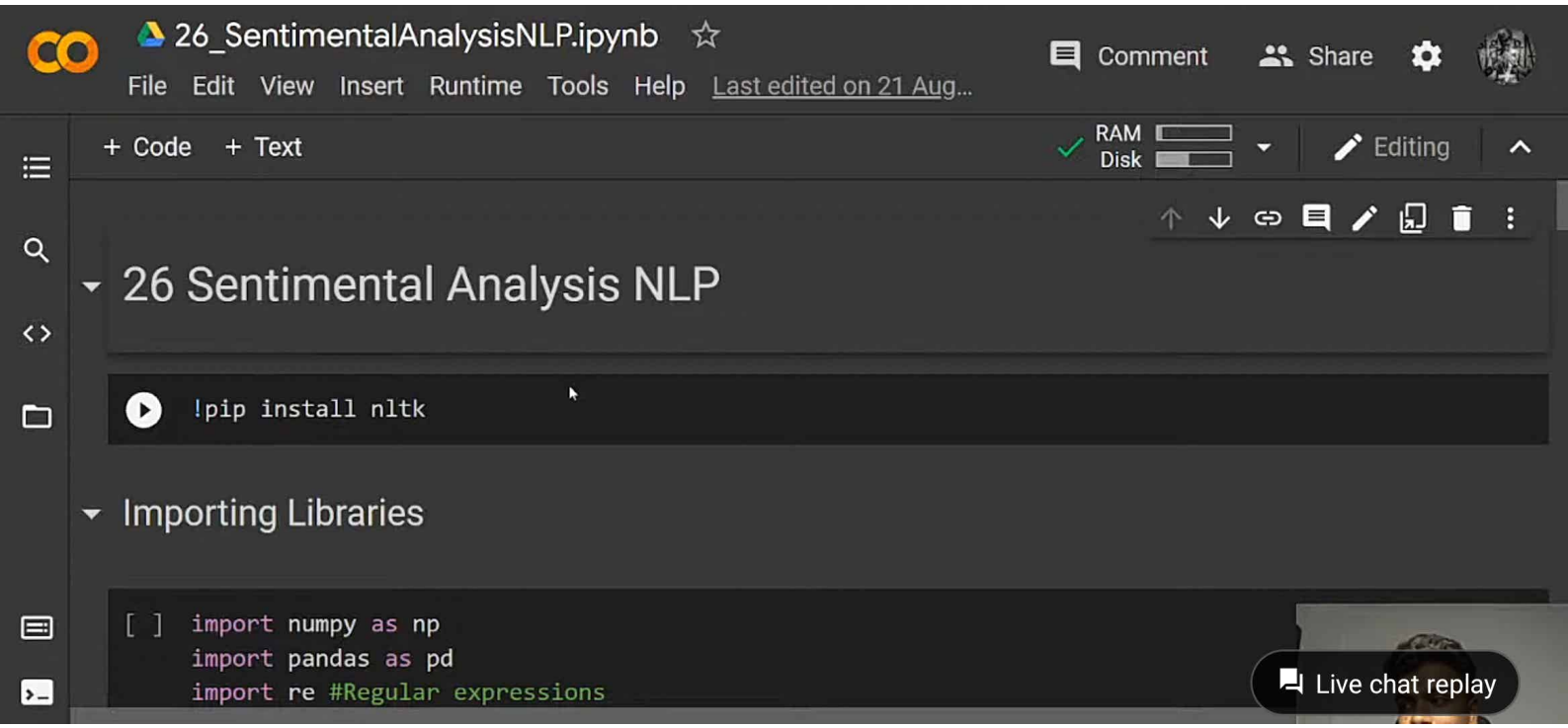


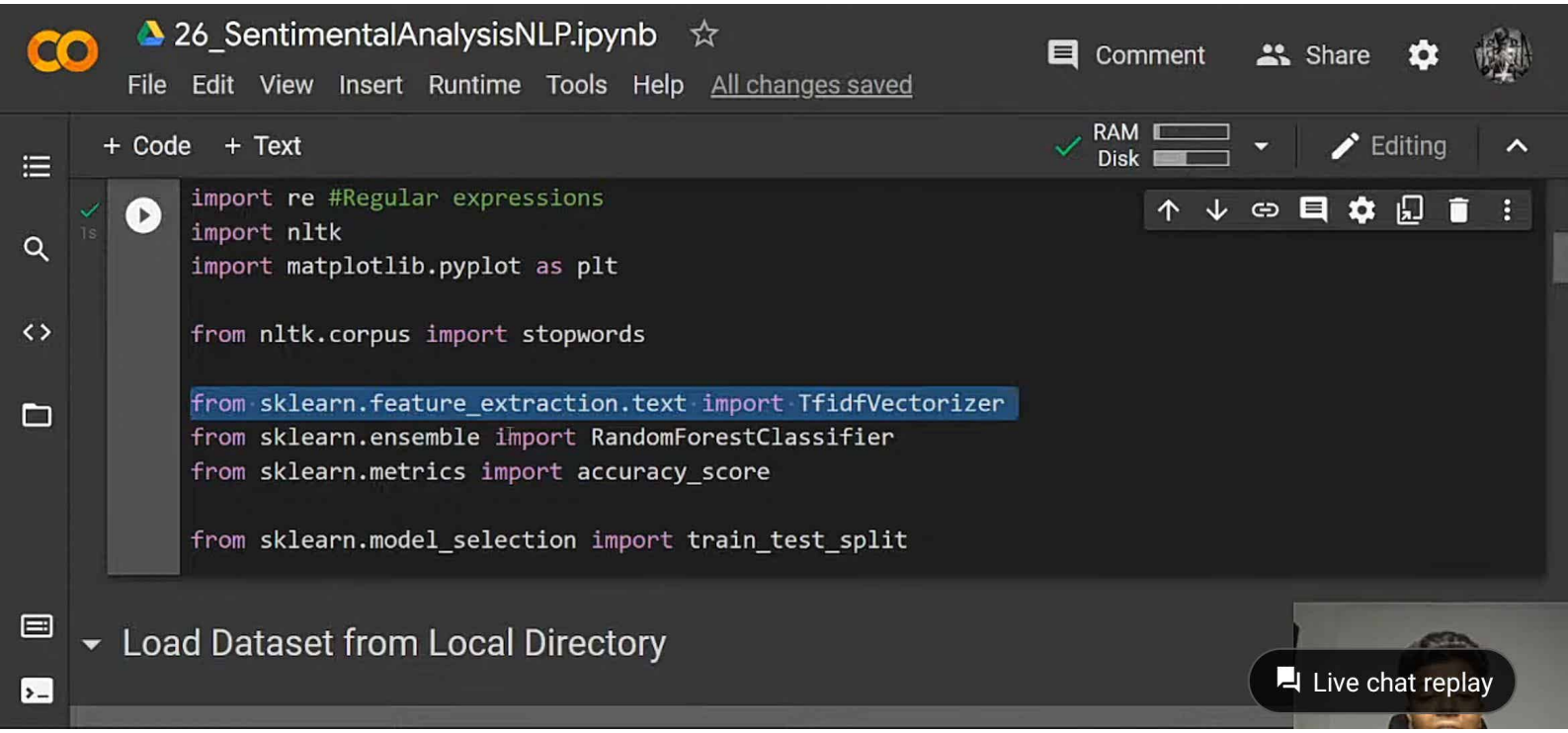
5.Feature extraction in Text

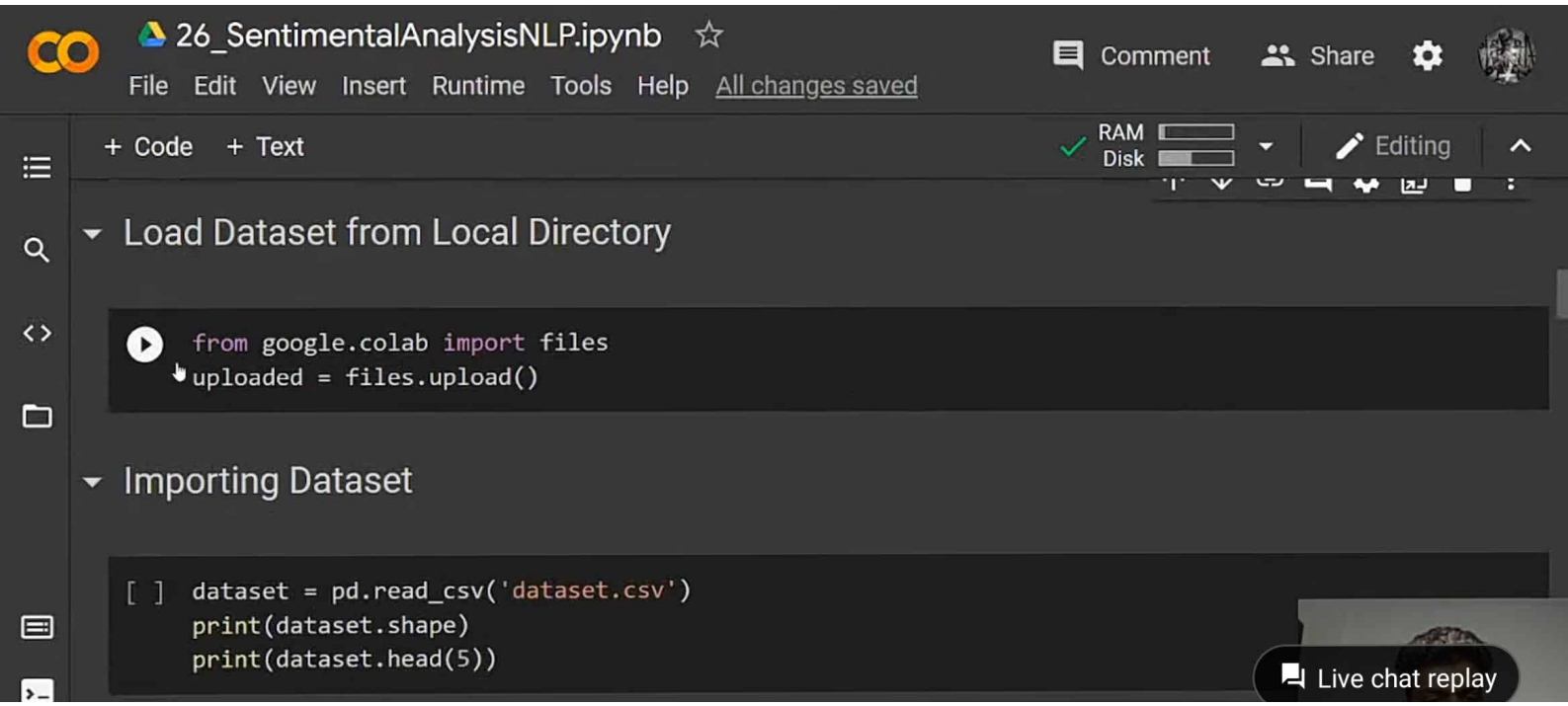
- CountVectorizer
- HashingVectorizer
- TfidfVectorizer

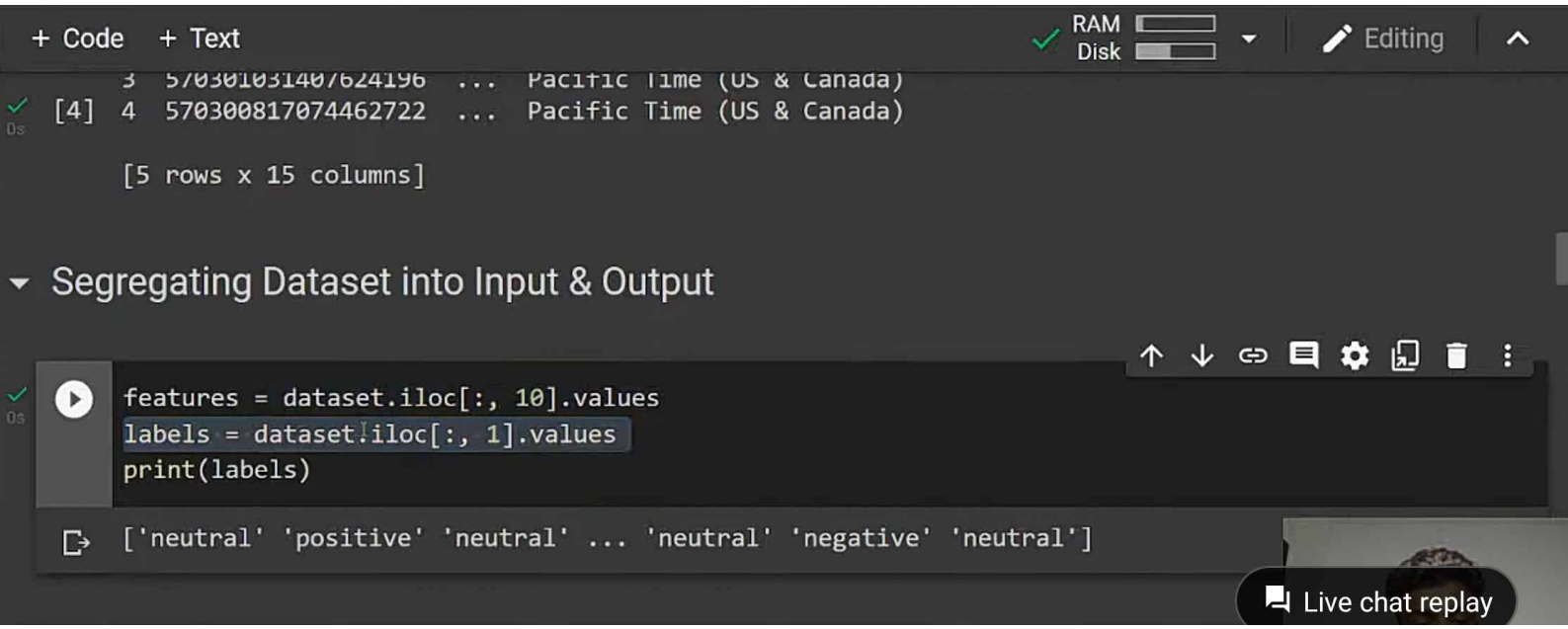


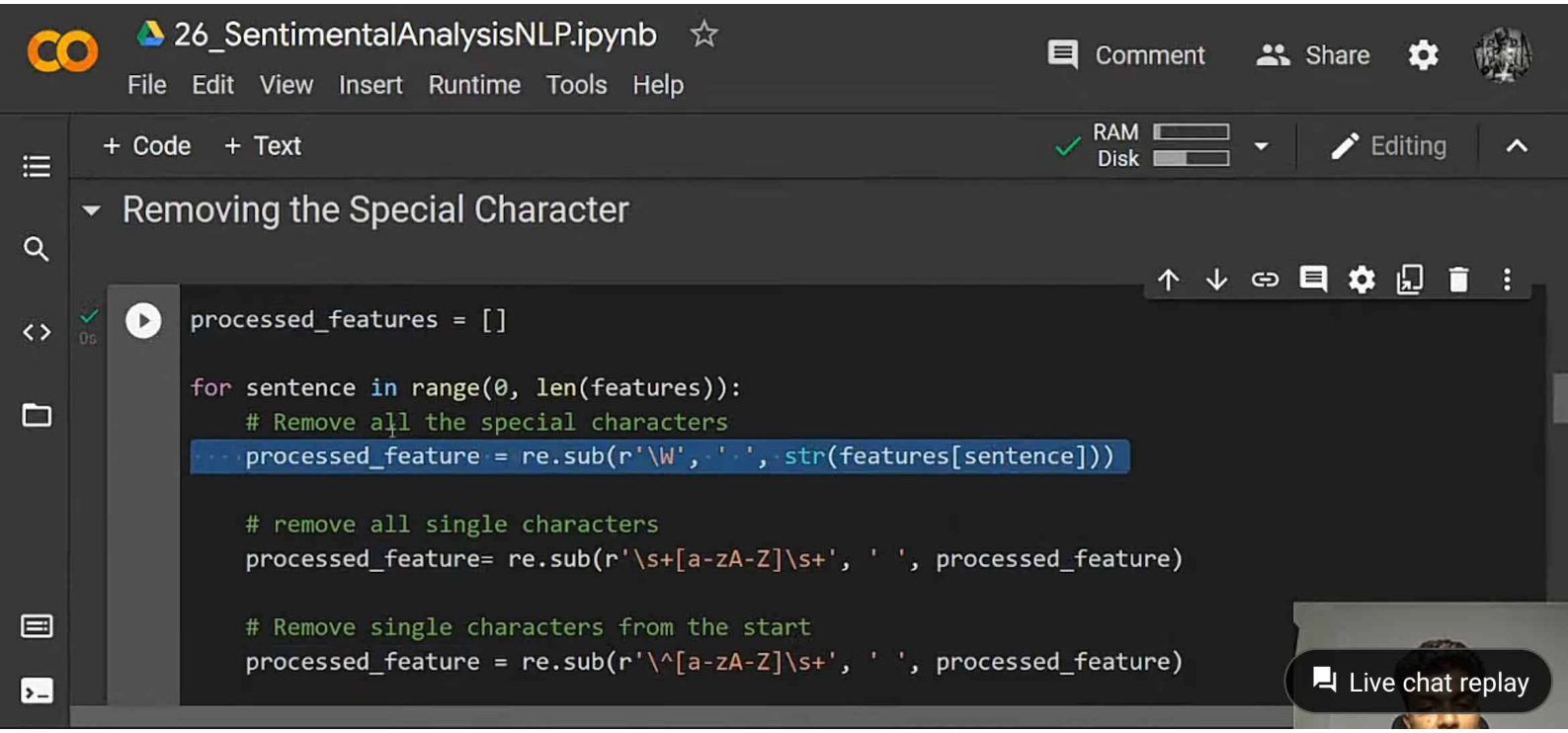
Live chat replay

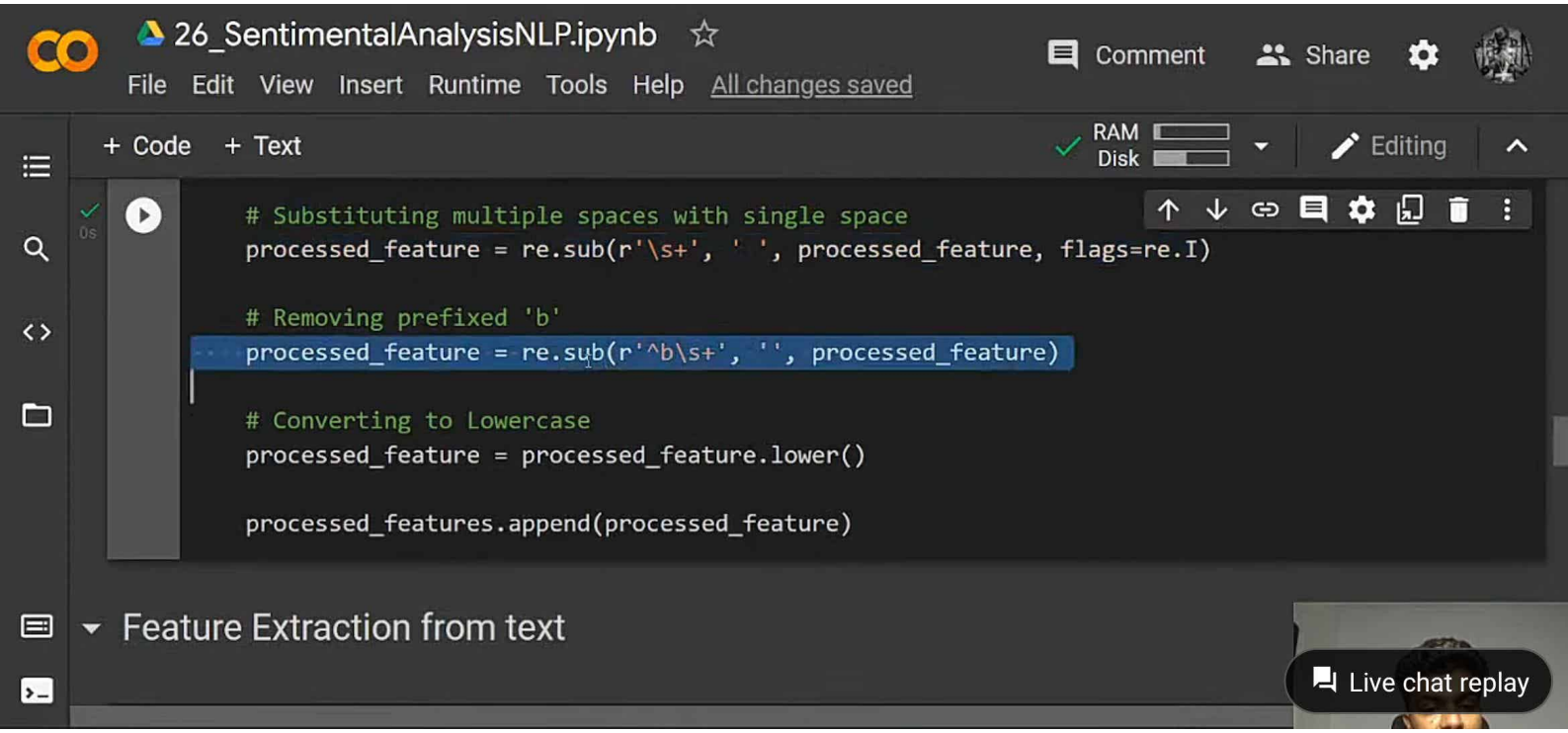

26_SentimentalAnalysisNLP.ipynb
File Edit View Insert Runtime Tools Help Last edited on 21 Aug...
Comment
Share
+ Code + Text
RAM
Disk
Editing
26 Sentimental Analysis NLP
!pip install nltk
Importing Libraries
[ ] import numpy as np
import pandas as pd
import re #Regular expressions
Live chat replay

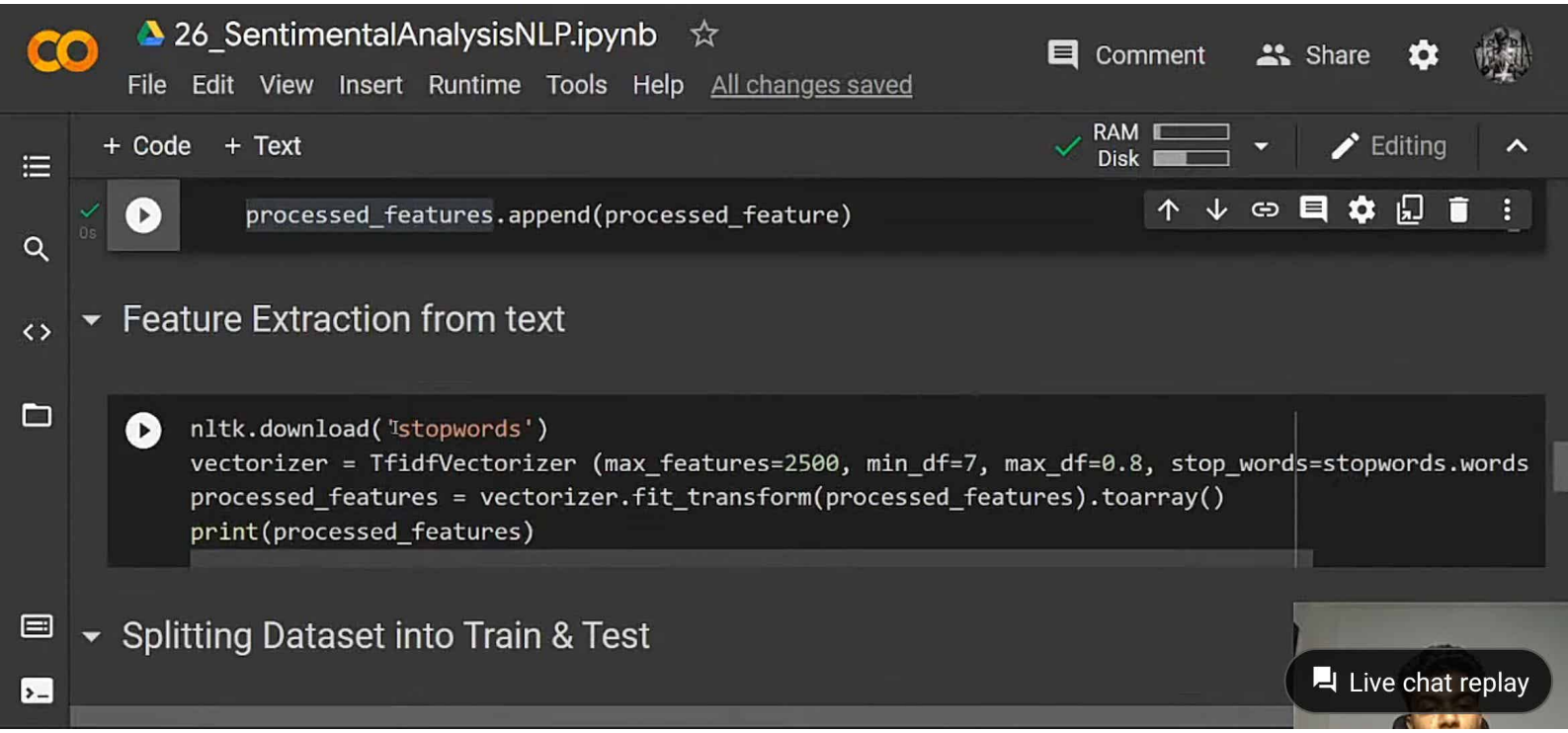

26_SentimentalAnalysisNLP.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
import re #Regular expressions
import nltk
import matplotlib.pyplot as plt
from nltk.corpus import stopwords
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score
from sklearn.model_selection import train_test_split
Load Dataset from Local Directory
Live chat replay

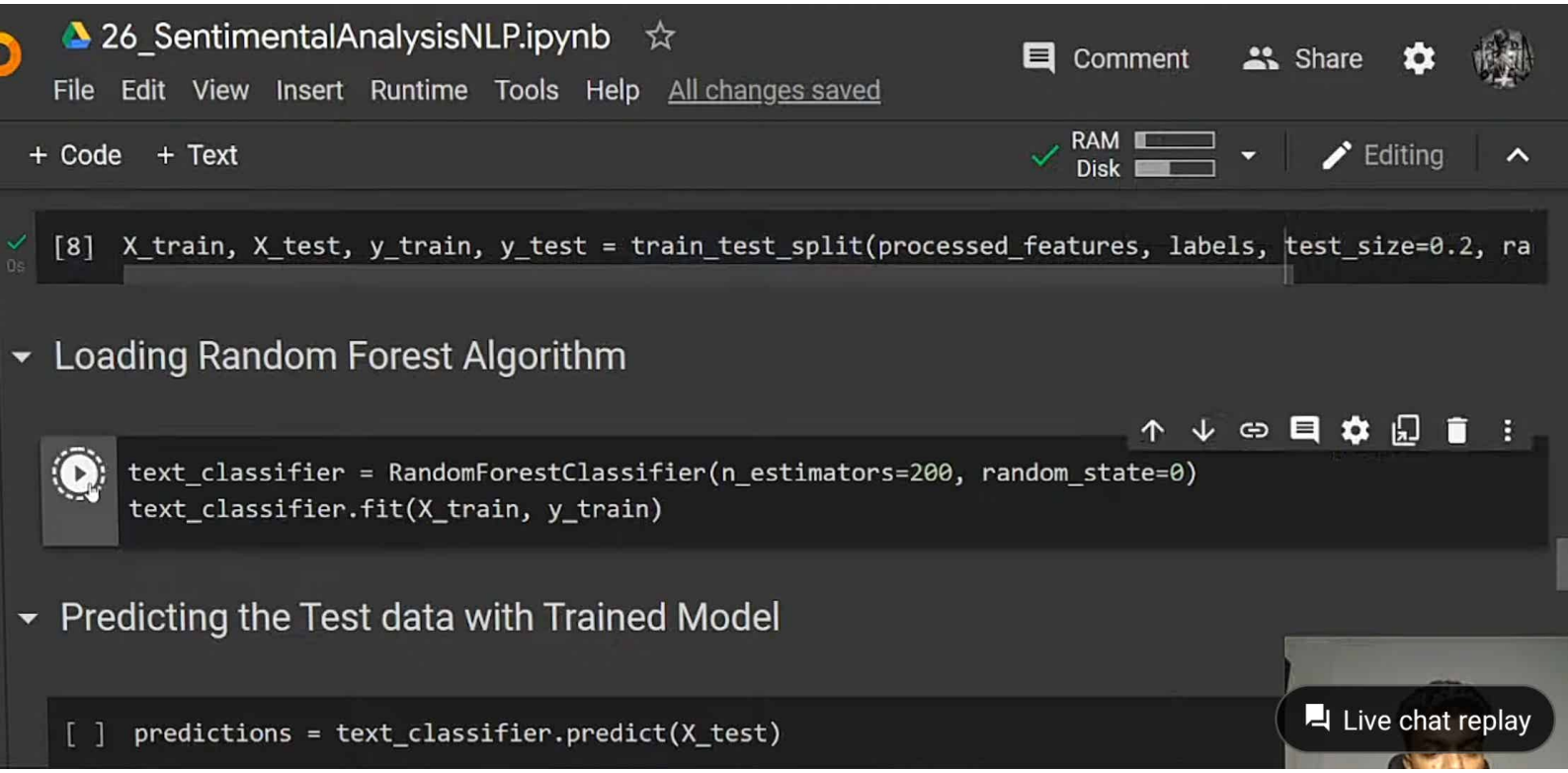

26_SentimentalAnalysisNLP.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
Load Dataset from Local Directory
from google.colab import files
uploaded = files.upload()
Importing Dataset
[ ] dataset = pd.read_csv('dataset.csv')
print(dataset.shape)
print(dataset.head(5))
Live chat replay

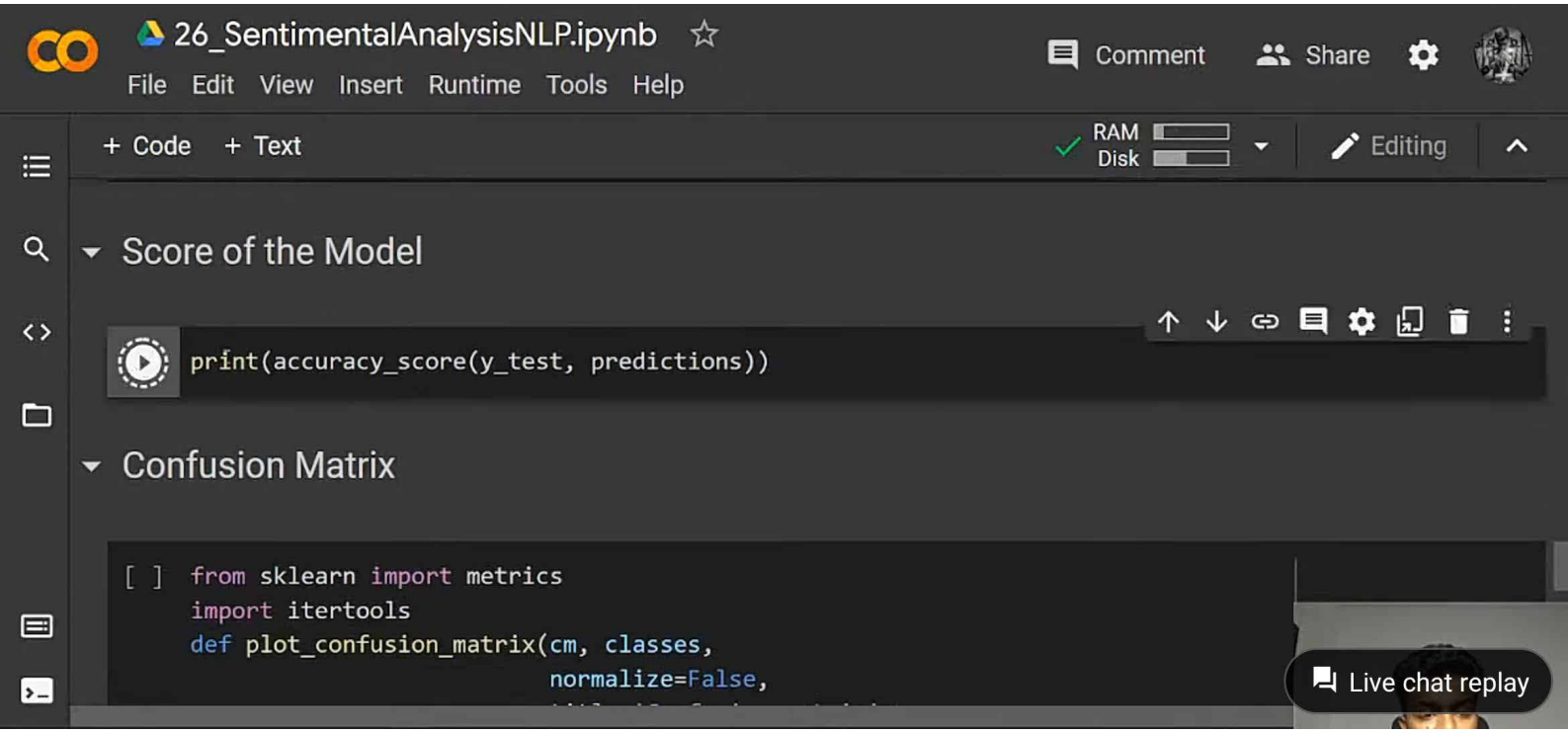

+ Code
+ Text
RAM
Disk
Editing
3 570301031407624196 ... Pacific Time (US & Canada)
[4] 4 570300817074462722 ... Pacific Time (US & Canada)
[5 rows x 15 columns]
Segregating Dataset into Input & Output
features = dataset.iloc[:, 10].values
labels = dataset.iloc[:, 1].values
print(labels)
['neutral' 'positive' 'neutral' ... 'neutral' 'negative' 'neutral']
Live chat replay

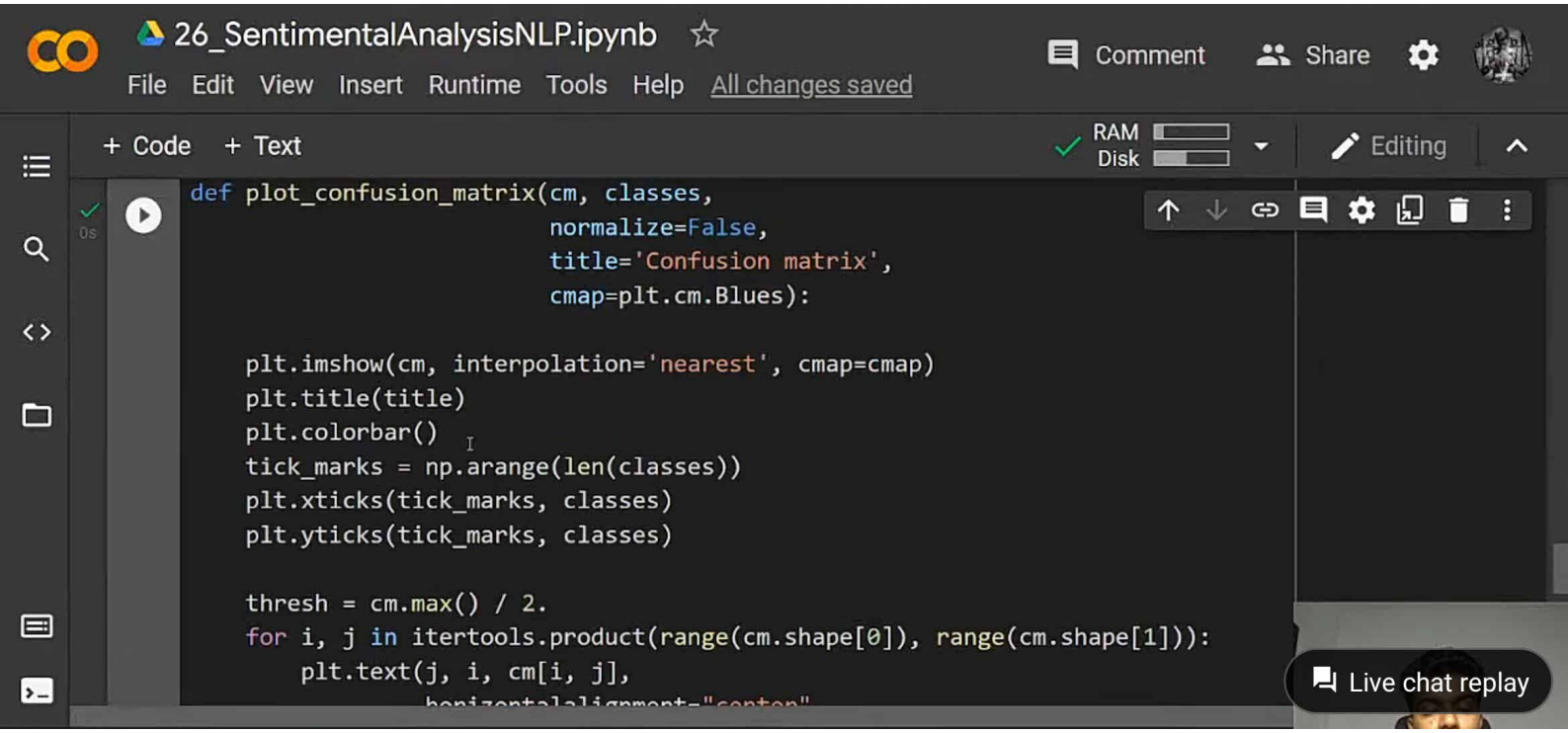

26_SentimentalAnalysisNLP.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
+ Code
+ Text
RAM
Disk
Editing
Removing the Special Character
processed_features = []
for sentence in range(0, len(features)):
# Remove all the special characters
processed_feature = re.sub(r'\W', ' ', str(features[sentence]))
# remove all single characters
processed_feature= re.sub(r'\s+[a-zA-Z]\s+', ' ', processed_feature)
# Remove single characters from the start
processed_feature = re.sub(r'\^[a-zA-Z]\s+', ' ', processed_feature)
Live chat replay


26_SentimentalAnalysisNLP.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
# Substituting multiple spaces with single space
processed_feature = re.sub(r'\s+', ' ', processed_feature, flags=re.I)
# Removing prefixed 'b'
processed_feature = re.sub(r'^b\s+', '', processed_feature)
# Converting to Lowercase
processed_feature = processed_feature.lower()
processed_features.append(processed_feature)
Feature Extraction from text
Live chat replay


26_SentimentalAnalysisNLP.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
processed_features.append(processed_feature)
Feature Extraction from text
nltk.download('stopwords')
vectorizer = TfidfVectorizer (max_features=2500, min_df=7, max_df=0.8, stop_words=stopwords.words
processed_features = vectorizer.fit_transform(processed_features).toarray()
print(processed_features)
Splitting Dataset into Train & Test
Live chat replay


26_SentimentalAnalysisNLP.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
[8] X_train, X_test, y_train, y_test = train_test_split(processed_features, labels, test_size=0.2, ra
Loading Random Forest Algorithm
text_classifier = RandomForestClassifier(n_estimators=200, random_state=0)
text_classifier.fit(X_train, y_train)
Predicting the Test data with Trained Model
[ ] predictions = text_classifier.predict(X_test)
Live chat replay


26_SentimentalAnalysisNLP.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
+ Code + Text
RAM
Disk
Editing
Score of the Model
print(accuracy_score(y_test, predictions))
Confusion Matrix
[ ] from sklearn import metrics
import itertools
def plot_confusion_matrix(cm, classes,
normalize=False,
Live chat replay


26_SentimentalAnalysisNLP.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
def plot_confusion_matrix(cm, classes,
normalize=False,
title='Confusion matrix',
cmap=plt.cm.Blues):
plt.imshow(cm, interpolation='nearest', cmap=cmap)
plt.title(title)
plt.colorbar()
tick_marks = np.arange(len(classes))
plt.xticks(tick_marks, classes)
plt.yticks(tick_marks, classes)
thresh = cm.max() / 2.
for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
plt.text(j, i, cm[i, j],
Live chat replay

26_SentimentalAnalysisNLP.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

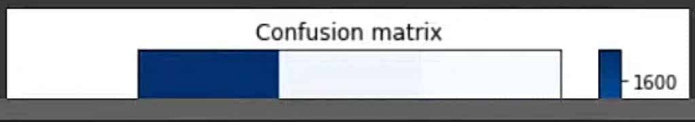
+ Code + Text

RAM Disk Editing

```
    thresh = cm.max() / 2.
    for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
        plt.text(j, i, cm[i, j],
                 horizontalalignment="center",
                 color="white" if cm[i, j] > thresh else "black")

    plt.tight_layout()
    plt.ylabel('True label')
    plt.xlabel('Predicted label')

cm = metrics.confusion_matrix(y_test, predictions, labels=['negative', 'neutral', 'positive'])
plot_confusion_matrix(cm, classes=['negative', 'neutral', 'positive'])
```



Live chat replay