

Lab 3 Assignment CS203.....

Team 31

Vansh Kumar (23110351)

Aditya Jain (23110016)

GitHub URL: <https://github.com/VanshOnGit/CS-203-Assignment-3>

Task 1: Environment Setup

- Execute the “history -c && history -w” command. It should clear all the history of the specific user

```
vansh@ubuntu:~$ history -c && history -w  
vansh@ubuntu:~$ _
```

This command clears the user's command history for the session and writes the changes to the history file.

- Execute the “history” command and take the screenshot.

```
vansh@ubuntu:~$ history -c && history -w  
vansh@ubuntu:~$ history  
  1  history  
vansh@ubuntu:~$ _
```

This command displays the user's current command history, which should now be empty

- Execute command “`sudo apt list | grep -i python3.10`”. If the package for Python 3.10 is installed, then it is removed.

```
vansh@ubuntu:~$ sudo apt list | grep -i pythonn3.10
WARNING: apt does not have a stable CLI interface. Use with caution in scripts.
vansh@ubuntu:~$ _
```

This command lists available packages matching Python 3.10 and removes it if installed.

- Download and install [Python 3.10](#) using tarball.

```
vansh@ubuntu:~$ wget https://www.python.org/ftp/python/3.10.16/Python-3.10.16.tgz
--2025-01-26 05:32:19-- https://www.python.org/ftp/python/3.10.16/Python-3.10.16.tgz
Resolving www.python.org (www.python.org)... 146.75.44.223, 2a04:4e42:48::223
Connecting to www.python.org (www.python.org)|146.75.44.223|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 25942994 (25M) [application/octet-stream]
Saving to: 'Python-3.10.16.tgz'

Python-3.10.16.tgz      100%[=====>] 24.74M  3.94MB/s  in 6.3s
2025-01-26 05:32:26 (3.92 MB/s) - 'Python-3.10.16.tgz' saved [25942994/25942994]

vansh@ubuntu:~$ _
```

This command downloads the tarball (compressed source code) for Python 3.10.

- This command extracts the downloaded tarball

```
vansh@ubuntu:~$ tar -xvzf Python-3.10.16.tgz_
Python-3.10.16/Tools/unicode/gencodec.py
Python-3.10.16/Tools/unicode/genmap_japanese.py
Python-3.10.16/Tools/unicode/genmap_korean.py
Python-3.10.16/Tools/unicode/genmap_schinese.py
Python-3.10.16/Tools/unicode/genmap_support.py
Python-3.10.16/Tools/unicode/genwincodec.py
Python-3.10.16/Tools/unicode/genwincodecs.bat
Python-3.10.16/Tools/unicode/listcodecs.py
Python-3.10.16/Tools/unicode/makeunicodedata.py
Python-3.10.16/Tools/unicode/mkstringprep.py
Python-3.10.16/Tools/unicode/python-mappings/
Python-3.10.16/Tools/unicode/python-mappings/CP1140.TXT
Python-3.10.16/Tools/unicode/python-mappings/CP273.TXT
Python-3.10.16/Tools/unicode/python-mappings/GB2312.TXT
Python-3.10.16/Tools/unicode/python-mappings/KOI8-U.TXT
Python-3.10.16/Tools/unicode/python-mappings/TIS-620.TXT
Python-3.10.16/Tools/unicode/python-mappings/diff/
Python-3.10.16/Tools/unicode/python-mappings/diff/jisx0213-2000-std.txt.diff
Python-3.10.16/Tools/unicode/python-mappings/diff/jisx0213-2004-std.txt.diff
Python-3.10.16/Tools/unicode/python-mappings/gb-18030-2000.xml
Python-3.10.16/Tools/unicode/python-mappings/jisx0213-2004-std.txt
Python-3.10.16/Tools/unittestgui/
Python-3.10.16/Tools/unittestgui/README.txt
Python-3.10.16/Tools/unittestgui/unittestgui.py
Python-3.10.16/aclocal.m4
Python-3.10.16/config.guess
Python-3.10.16/config.sub
Python-3.10.16/configure
Python-3.10.16/configure.ac
Python-3.10.16/install-sh
Python-3.10.16/pyconfig.h.in
Python-3.10.16/setup.py
```

Python3.10 --version

```
vansh@ubuntu:~$ ls
Desktop  Documents  Downloads  Music  Pictures  Public  Python-3.10.16  Python-3.10.16.tgz  snap  Templates  Videos
vansh@ubuntu:~$ cd Python-3.10.16
vansh@ubuntu:~/Python-3.10.16$ python3.10 --version
Python 3.10.16
vansh@ubuntu:~/Python-3.10.16$
```

This verifies that Python 3.10 was installed successfully

- Make Pip environment using newly installed python3.10

Create a virtual environment:

```
vansh@ubuntu:~/Python-3.10.16$ python3.10 -m venv labelstudio_env  
vansh@ubuntu:~/Python-3.10.16$ _
```

This creates a virtual environment named 'labelstudio_env' using Python 3.10.

Activate the virtual environment:

```
vansh@ubuntu:~/Python-3.10.16$ python3.10 -m venv labelstudio_env  
vansh@ubuntu:~/Python-3.10.16$ source labelstudio_env/bin/activate  
(labelstudio_env) vansh@ubuntu:~/Python-3.10.16$
```

This activates the virtual environment

- Install [label studio](#) using pip.

```
(labelstudio_env) vansh@ubuntu:~/Python-3.10.16$ pip install label-studio
Collecting label-studio
  Downloading label_studio-1.15.0-py3-none-any.whl.metadata (14 kB)
Collecting Django<4.3.0,>=4.2.13 (from label-studio)
  Downloading Django-4.2.18-py3-none-any.whl.metadata (4.1 kB)
Collecting appdirs>=1.4.3 (from label-studio)
  Downloading appdirs-1.4.4-py2.py3-none-any.whl.metadata (9.0 kB)
Collecting attr==0.3.1 (from label-studio)
  Downloading attr-0.3.1.tar.gz (1.7 kB)
Installing build dependencies ... done
Getting requirements to build wheel ... done
Preparing metadata (pyproject.toml) ... done
```

This installs Label Studio in the pip environment

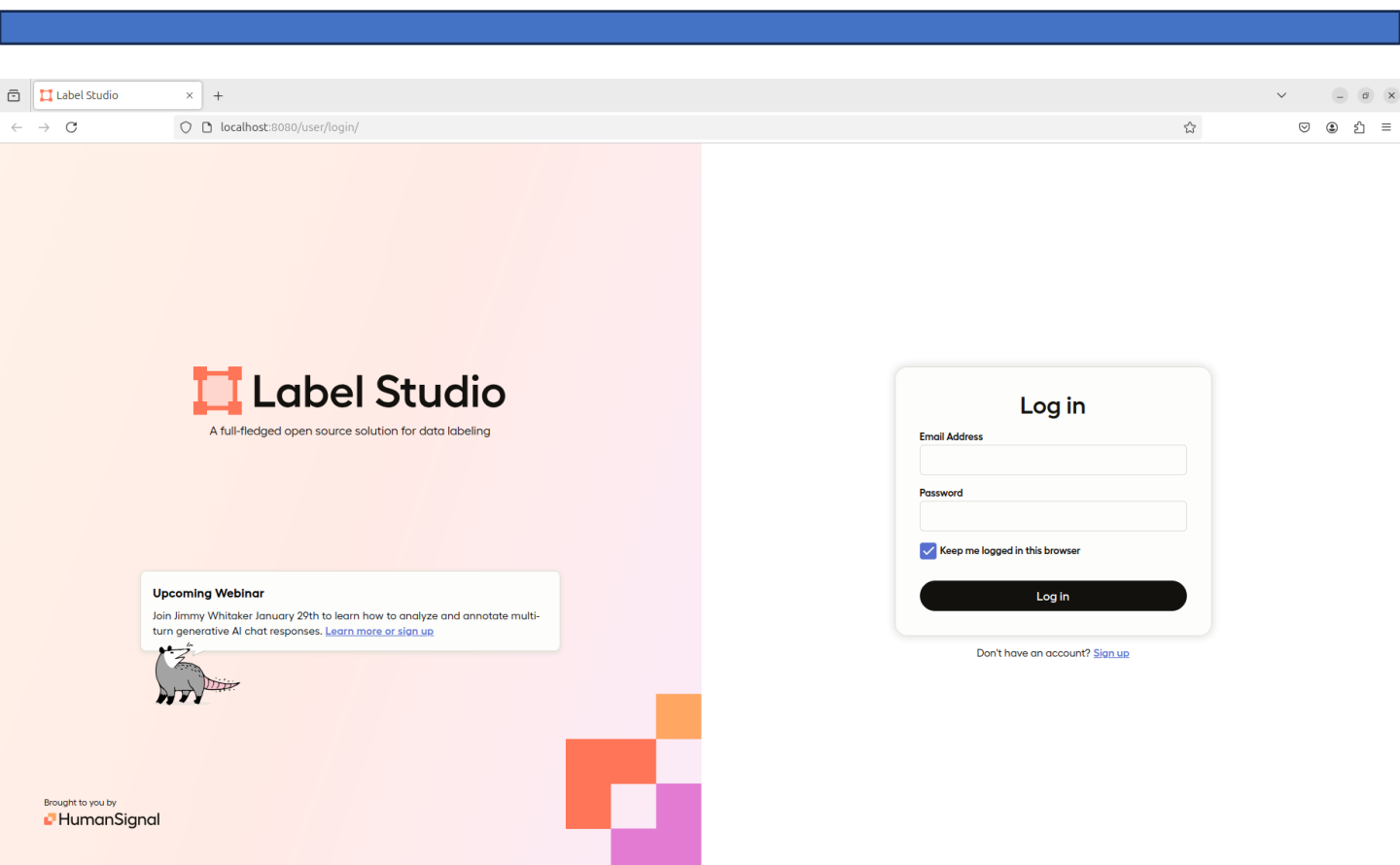
Open, Register, and Check Label Studio on the web browser.

1. Start Label Studio

```
(labelstudio_env) vansh@ubuntu:~/Python-3.10.16$ label-studio start
=> Database and media directory: /home/vansh/.local/share/label-studio
=> Static URL is set to: /static/
=> Database and media directory: /home/vansh/.local/share/label-studio
=> Static URL is set to: /static/
Read environment variables from: /home/vansh/.local/share/label-studio/.env
get 'SECRET_KEY' casted as '<class 'str'>' with default ''

Starting new HTTPS connection (1): pypi.org:443
https://pypi.org:443 "GET /pypi/label-studio/json HTTP/1.1" 200 33651
Initializing database..
January 26, 2025 - 06:40:21
Django version 4.2.18, using settings 'label_studio.core.settings.label_studio'
Starting development server at http://0.0.0.0:8080/
Quit the server with CONTROL-C.
```

This starts the Label Studio server

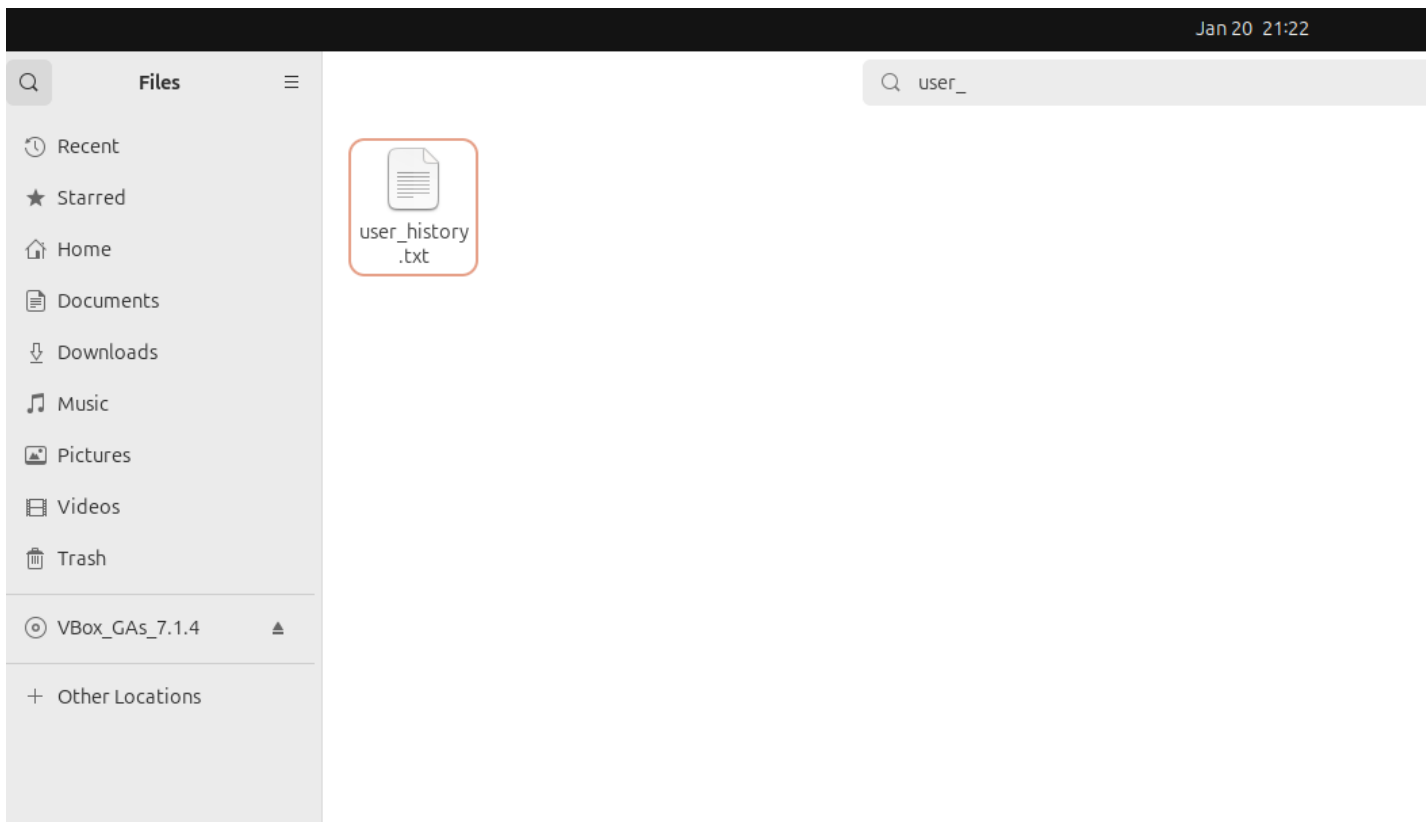


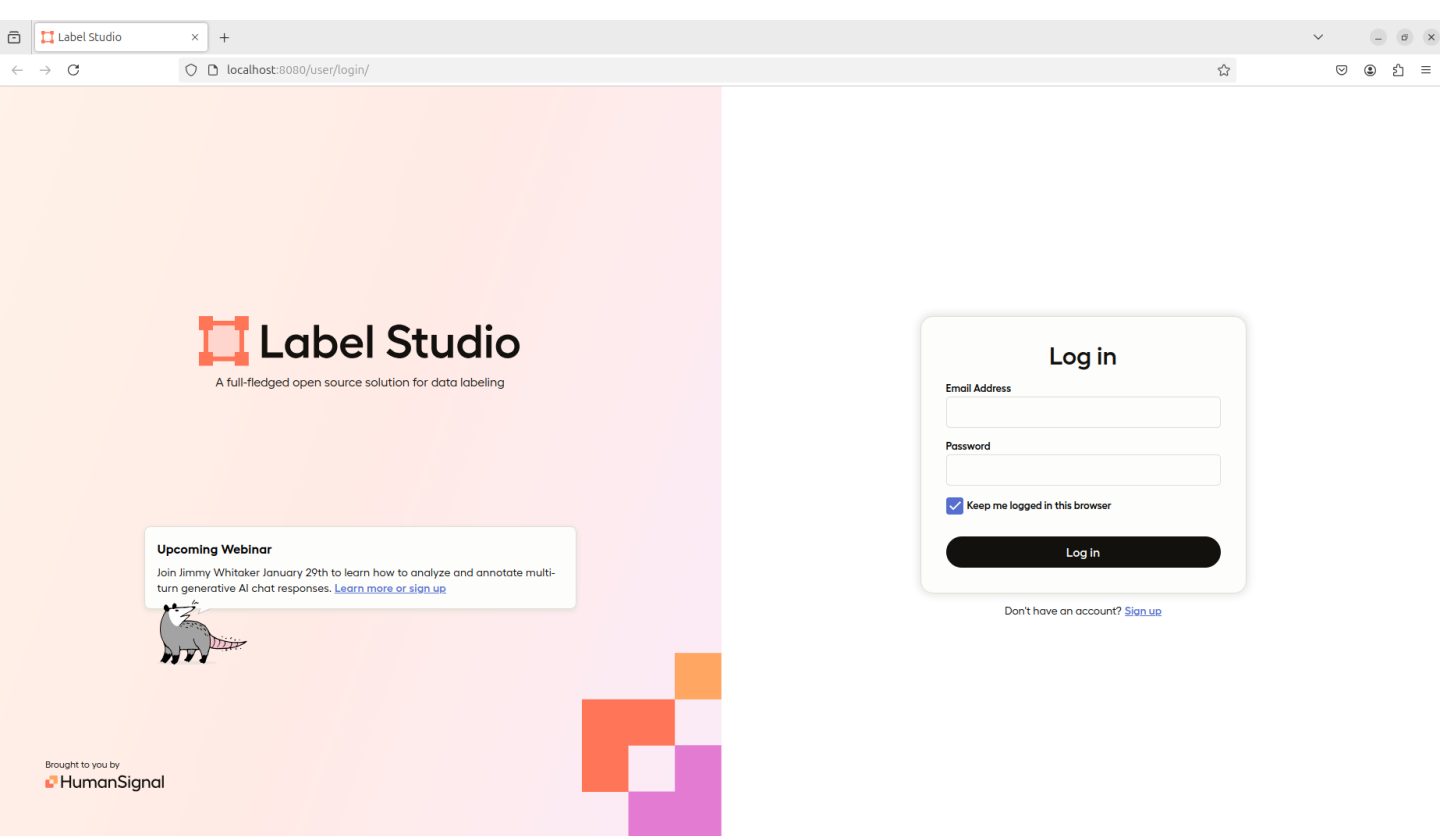
This shows Label Studio running in a web browser.

•Execute “history > user_history.txt.” Upload user_history.txt file on Git Hub.

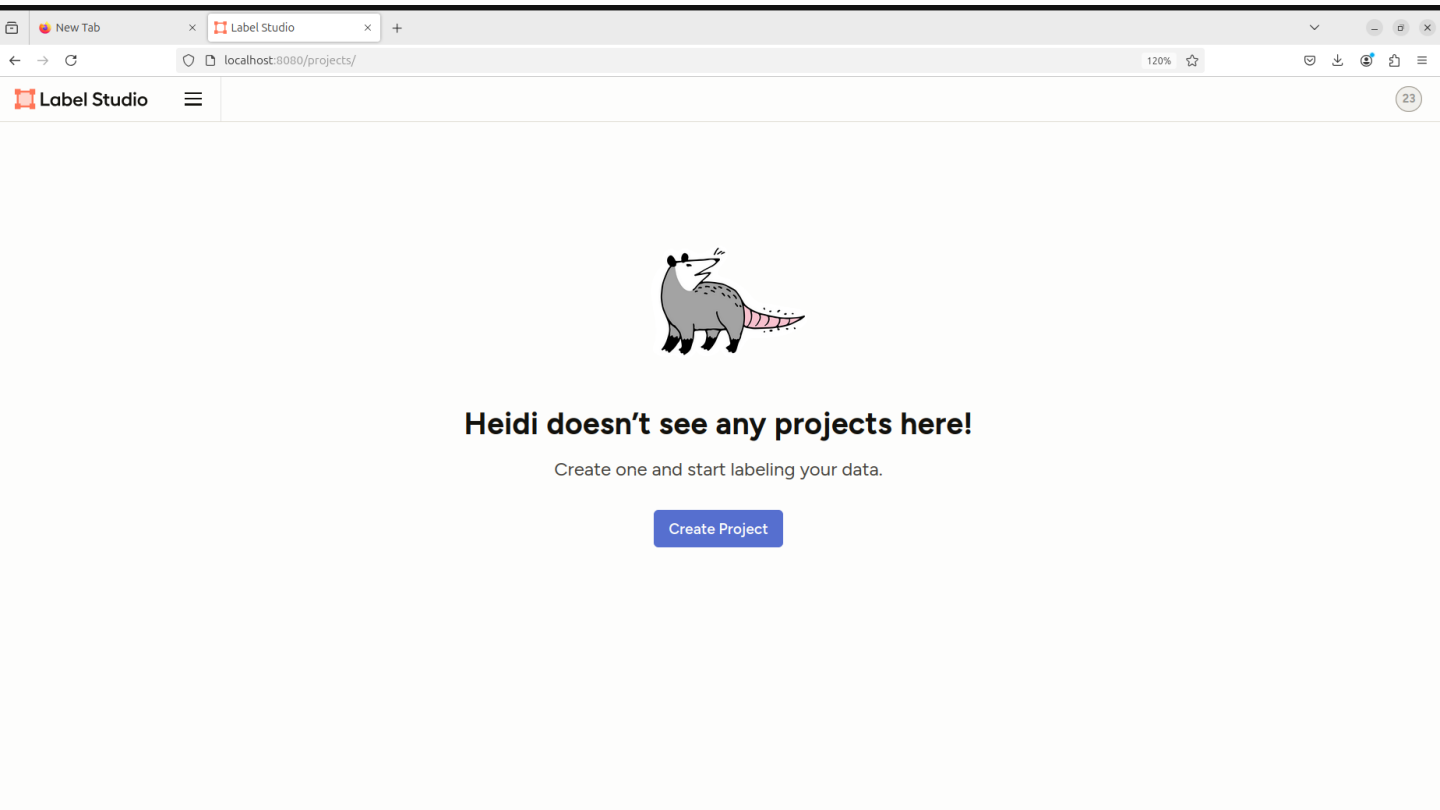
```
(labelstudio_env) vansh@ubuntu:~/Python-3.10.16$ history > user_history.txt  
(labelstudio_env) vansh@ubuntu:~/Python-3.10.16$
```

This saves the current session's command history to a file named 'user_history.txt'. It is in our GitHub Repository





This shows Label Studio running in a web browser and now we login



Task 2: Annotating a Dataset Using Label Studio

We are in group 31 so we need data points from 600 to 619 so we need to remove the data points we don't need so we do that using python

```
Open  ▾  FilterData.py  
~/Desktop/Assignment3  
  
import pandas as pd  
  
file_path = "/home/vansh/Desktop/Assignment3/NLP.csv"  
  
df = pd.read_csv(file_path)  
  
filtered_df = df.iloc[600:620]  
  
filtered_df.to_csv(file_path, index=False)  
  
print("Data points from 600 to 619 are saved.")
```

```
vansh@vansh: ~/Desktop/Assignment3  
  
vansh@vansh:~/Desktop/Assignment3$ python3 FilterData.py  
Data points from 600 to 619 are saved.  
vansh@vansh:~/Desktop/Assignment3$
```

Import NLP Dataset

Create Project

Project Name

Data Import

Labeling Setup

Delete

Save

Project Name

NLP Annotation

Description

Optional description of your project

Workspace

Enterprise

Select an option

Simplify project management by organizing projects into workspaces. [Learn more](#)

This shows the creation of a new project for annotating NLP data

Import Data

Cancel

Import

Dataset URL

Add URL

 or

Upload More Files

Treat CSV/TSV as

List of tasks

Time Series or Whole Text File

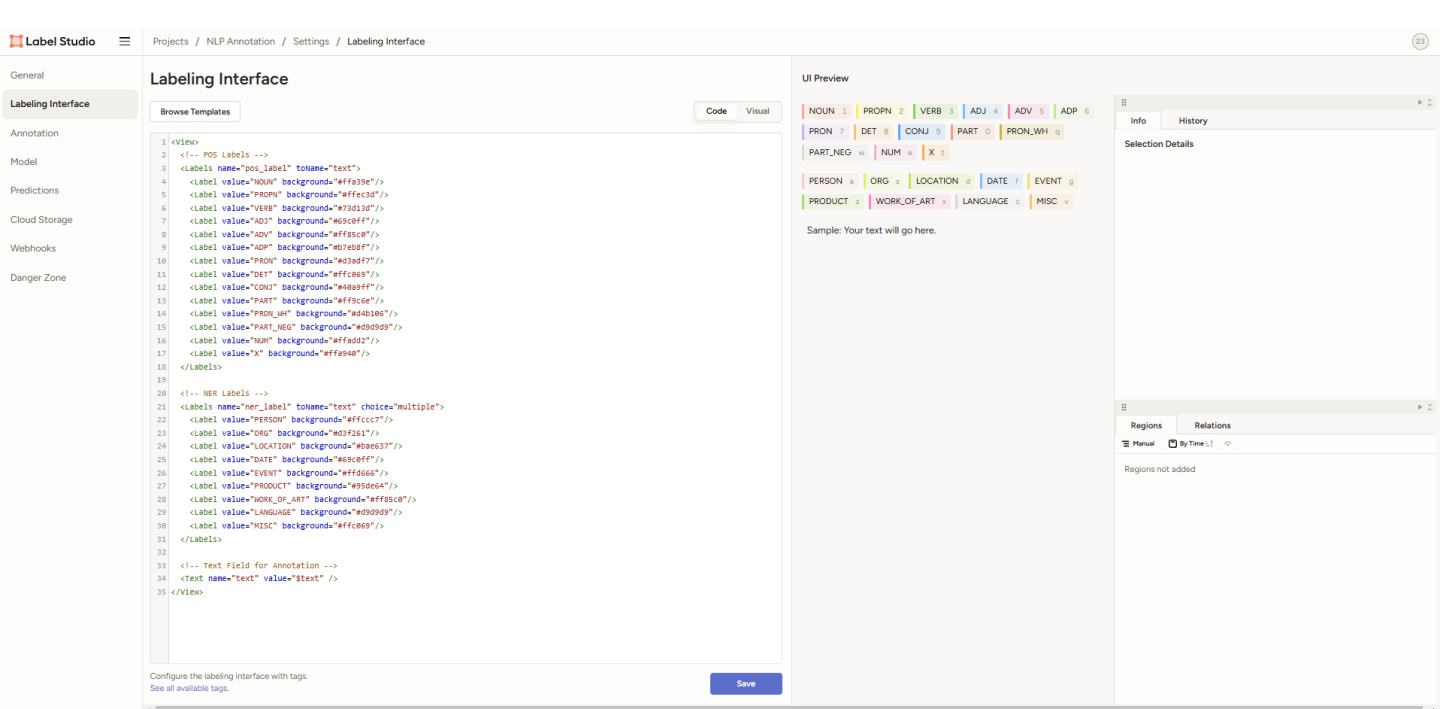
1 files uploaded

upload/1/0e86b882-NLP.csv

Importing nlp data

Default										Tasks: 20 / 20		Annotations: 0		Predictions: 0	
Actions		Columns	Filters	Order	not set		Label All Tasks					Import	Export	List	Grid
ID	Completed						Annotated by		text	str					
1	0	0	0	0			Revanth Reddy Telangana CM Oath: तेलंगाना में कांग्रेस की 6 guarantee पर रवत रेड्डी की								
2	0	0	0	0			BJP PARLIAMENTARY MEETING: दिल्ली में BJP संसदीय दल की बैठक BJP								
3	0	0	0	0			Sharmistha Mukherjee Book: क्या राष्ट्र में congress को दोबारा से खड़ा करने की योग्यता								
4	0	0	0	0			Revanth Reddy Oath Ceremony: रवत रेड्डी आज telangana के Chief Minister								
5	0	0	0	0			India Alliance Meeting Cancel : I.N.D.I.A की बैठक टली, अब टिसंबर के तीसरे हफ्ते में								
6	0	0	0	0			Rajasthan CM पर BJP में मुठों का शक्ति प्रदर्शन, Vasundhara Raje से 30 विधायक मिले								
7	0	0	0	0			MP Election Result 2023 : BJP को दो तिहाई से ज्यादा बहुमत, CM शिवराज सिंह बोले- MP के								
8	0	0	0	0			दैनिक भास्कर का रेखा उद्भव: भोपाल में वैदता समूह के संस्थापक-चेयरमैन अनिल अग्रवाल								
9	0	0	0	0			Uttarkashi Tunnel Rescue: Lucknow पहुंचे UP के मजदूर:								

Data imported



Adding POS tagging and NER functionality in the Label Setup

Default

Tasks: 20 / 20 Annotations: 20 Predictions: 0

Actions

Columns

Filters

Order

not set

Label All Tasks

Import

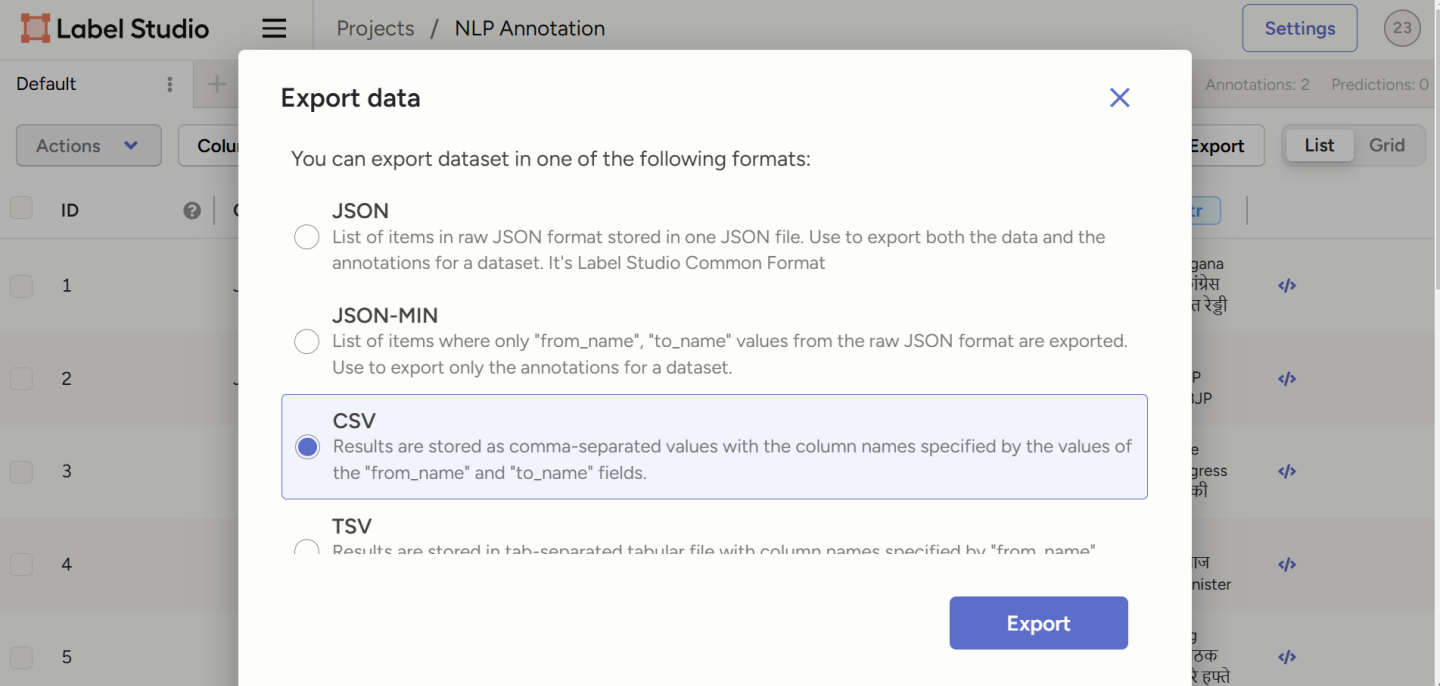
Export

List

Grid

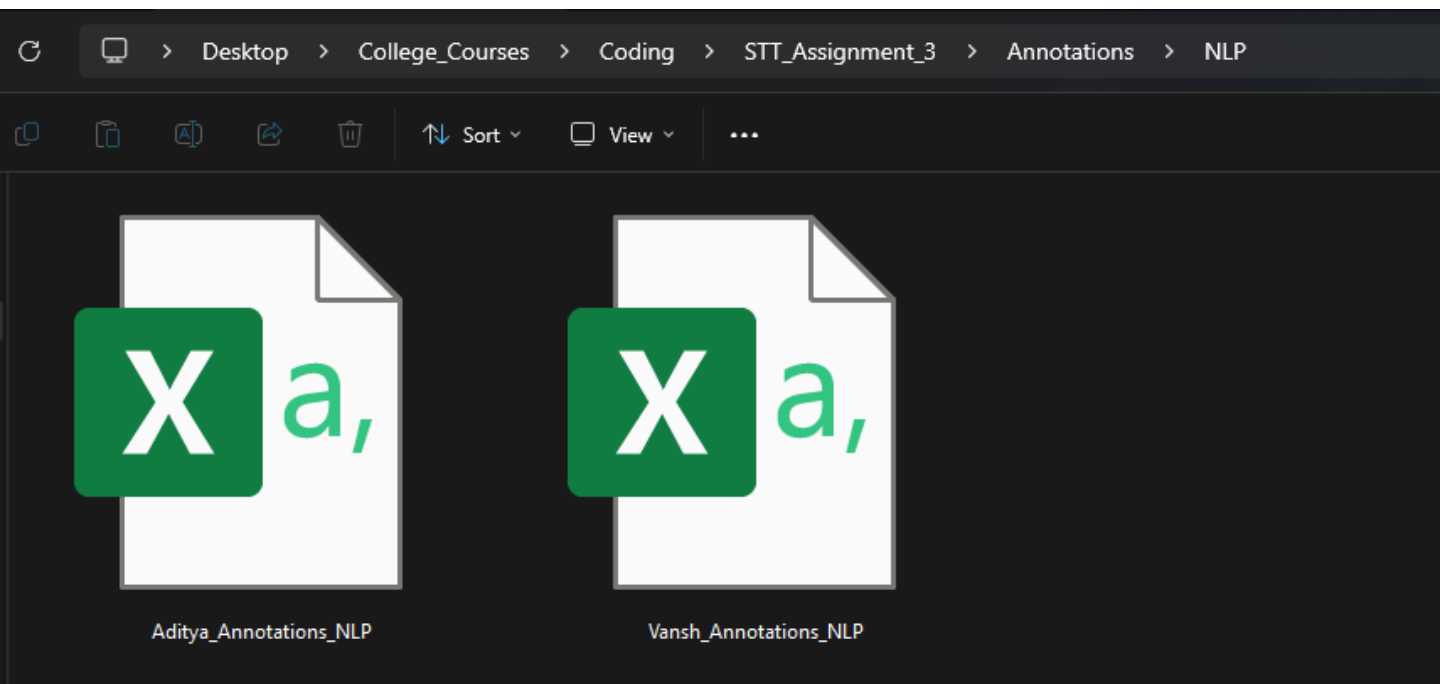
ID	Completed					Annotated by	text
1	Jan 20 2025, 22:32:19	1	0	0	23		Revanth Reddy Telangana CM Oath: तेलंगाना में कांग्रेस की 6 guarantee पर खेत रेंडू की
2	Jan 20 2025, 22:32:19	1	0	0	23		BJP PARLIAMENTARY MEETING: दिल्ली में BJP संसदीय दल की बैठक BJP
3	Jan 20 2025, 22:32:20	1	0	0	23		Sharmistha Mukherjee Book: क्या राहुल में congress को दोबारा से खड़ा करने की योग्यता
4	Jan 20 2025, 22:32:21	1	0	0	23		Revanth Reddy Oath Ceremony: खेत रेंडू आज telangana के Chief Minister
5	Jan 20 2025, 22:32:22	1	0	0	23		India Alliance Meeting Cancel : I.N.D.I.A की बैठक टली, अब दिसंबर के तीसरे हफ्ते में
6	Jan 20 2025, 22:32:23	1	0	0	23		Rajasthan CM पर BJP से गुलें का सख्ति प्रदर्शन, Vasundhara Raje से 30 रियायत मिले
7	Jan 20 2025, 22:32:23	1	0	0	23		MP Election Result 2023 : BJP को वो विहाई से ज्यादा बहुमत, CM शिवराज सिंह बोले- MP के
8	Jan 20 2025, 22:32:24	1	0	0	23		दैनिक भास्कर का प्रेरणा उत्सव: भोपाल में देवेंद्रात समूह के संस्थापक-चेयरमैन अनिल अग्रवाल
9	Jan 20 2025, 22:32:25	1	0	0	23		Uttarkashi Tunnel Rescue: Lucknow पहुंचे UP के मजदूर: CM योगी से मुलाकात, घर के लिए

All data points annotated



Exporting the annotations in the CSV File

Exporting Annotations to CSV



Now the annotations were done by both the teammates and they are saved in .Annotations/NLP Folder

The screenshot of annotations is not possible to show since each CSV line is too long. Please check the files.

Label Studio									
Projects / NLP Annotation 2									
Tasks: 20 / 20 Annotations: 0 Predictions: 0									
<div> <div>Actions</div> <div>Columns</div> <div>Filters</div> <div>Order</div> <div>not set</div> <div>Label All Tasks</div> <div>Import</div> <div>Export</div> <div>List</div> <div>Grid</div> </div>									
ID	Completed				Annotated by	data			
21		0	0	0		/data/upload/3/bd3a4129-img_600.jpg			
22		0	0	0		/data/upload/3/1e2a1258-img_601.jpg			
23		0	0	0		/data/upload/3/67a3d2f0-img_602.jpg			
24		0	0	0		/data/upload/3/03450070-img_603.jpg			
25		0	0	0		/data/upload/3/24fd1845-img_604.jpg			
26		0	0	0		/data/upload/3/7b6b6c55-img_605.jpg			
27		0	0	0		/data/upload/3/5cc05012-img_606.jpg			
28		0	0	0		/data/upload/3/6882ea04-img_607.jpg			
29		0	0	0		/data/upload/3/2e669890-img_608.jpg			

Now similarly importing the CV dataset of 20 images from img_600.jpg to img_619.jpg

Label Studio

Projects / NLP Annotation 2 / Labeling

Settings

23

image

img

0

0

0

0

0

0

0

0

0

0

#21

23110351

#RYHCl

seconds ago

☒ Trucks^[1]
☐ No Trucks^[2]

Submit

Info

History

Selection Details

Regions

Relations

Manual

By Time

Regions not added

Now Annotating the CV dataset by classifying them as Trucks or No Trucks

Label Studio Projects / NLP Annotation 2

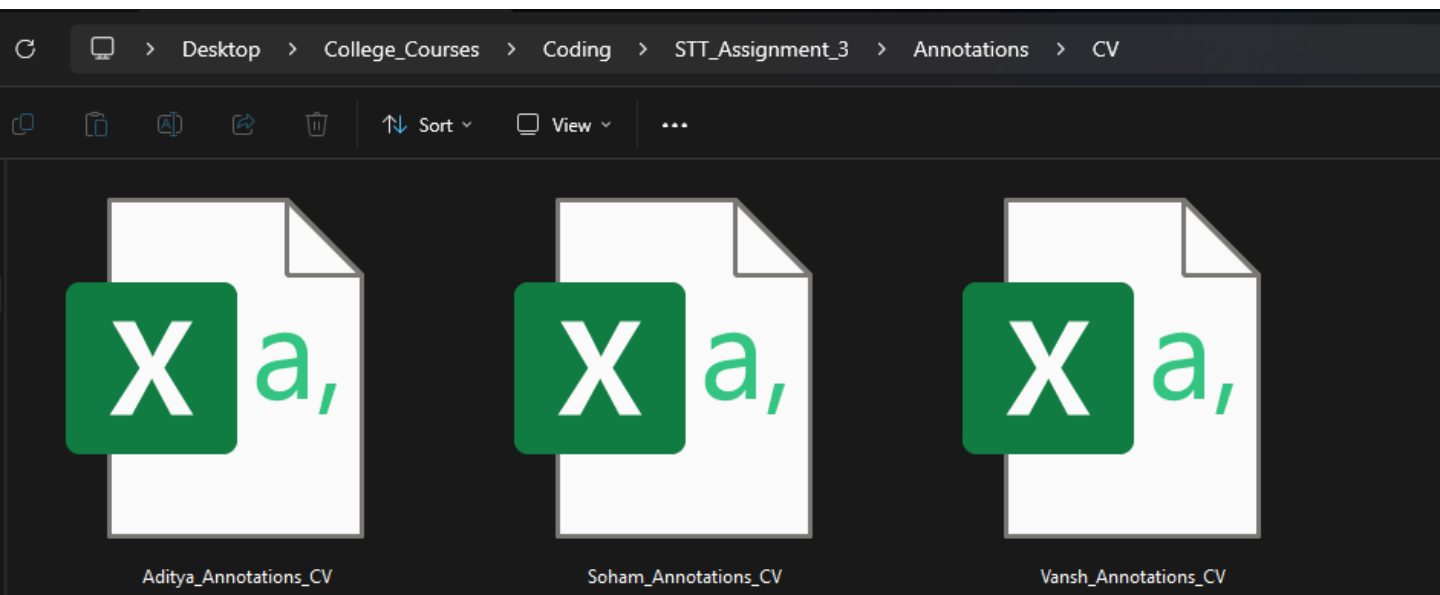
Tasks: 20 / 20 Annotations: 19 Predictions: 0

Actions Columns Filters Order not set Label All Tasks Import Export List Grid

ID	Completed				Annotated by	image	img
21	Jan 23 2025, 09:43:35	1	0	0	23		
22	Jan 23 2025, 09:43:41	1	0	0	23		
23	Jan 23 2025, 09:43:45	1	0	0	23		
24	Jan 23 2025, 09:43:50	1	0	0	23		
25	Jan 23 2025, 09:43:54	1	0	0	23		
26	Jan 23 2025, 09:43:59	1	0	0	23		
27	Jan 23 2025, 09:44:05	1	0	0	23		
28	Jan 23 2025, 09:44:08	1	0	0	23		
29	Jan 23 2025, 09:44:11	1	0	0	23		

CV Annotations done

Exporting Annotations to CSV



Now the annotations were done by both the teammates and another student from another team and files are saved in .Annotations/CV Folder

The screenshot of annotations is not possible to show since each CSV line is too long. Please check the files.

Task 3: Implementing Inter-Annotator Agreement

Please check Task_3.ipynb Jupyter file for Task 3

Thank You

CS203 Lab Assignment 3

Team 31

Vansh Kumar 23110351

Aditya Jain 23110016