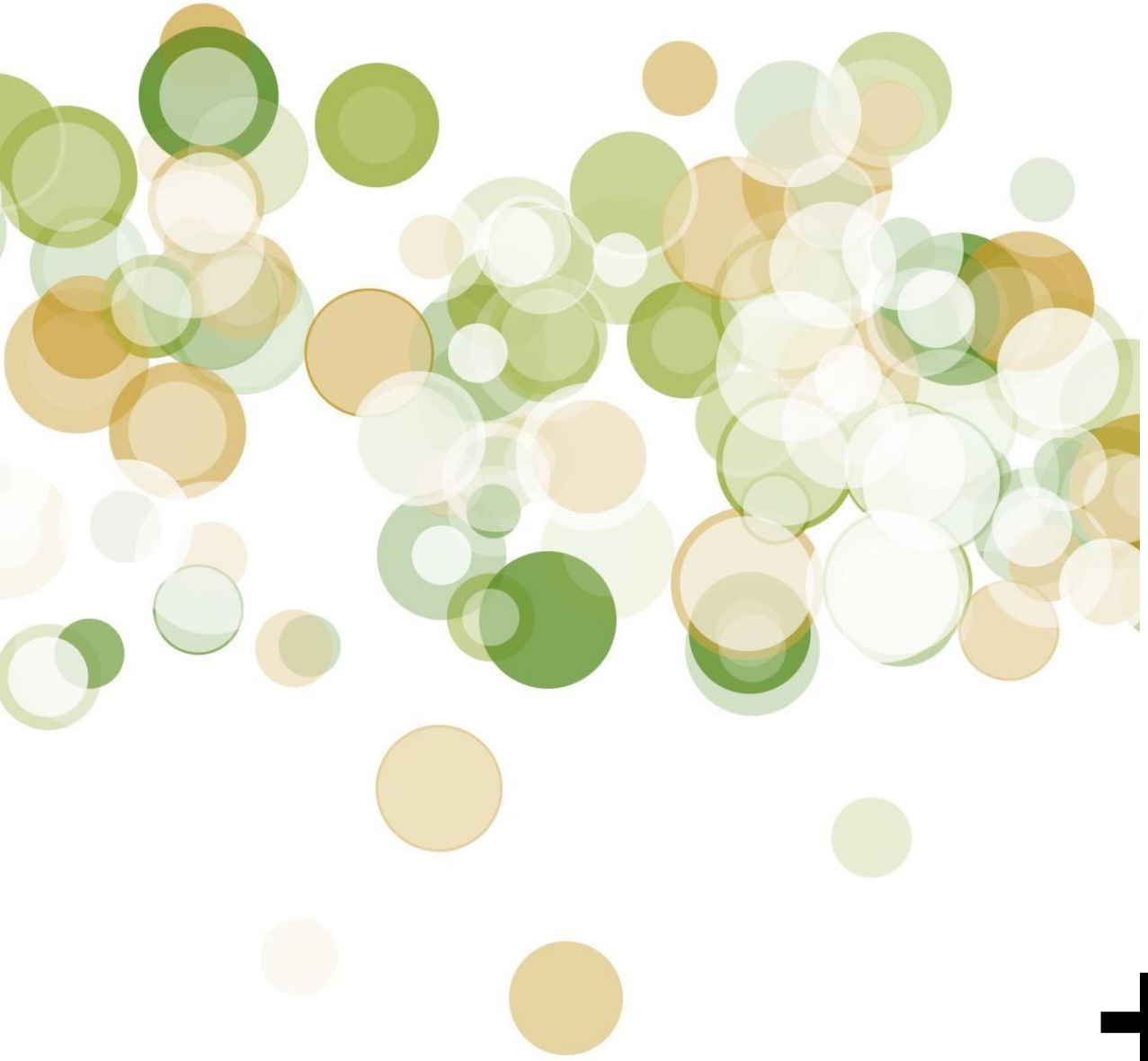


Social Network Analysis Big Data Engineering

Riccio Emanuele
Tammaro Ferdinando

a.a. 2022/23

Social Network Analysis

+

Social Network Analysis

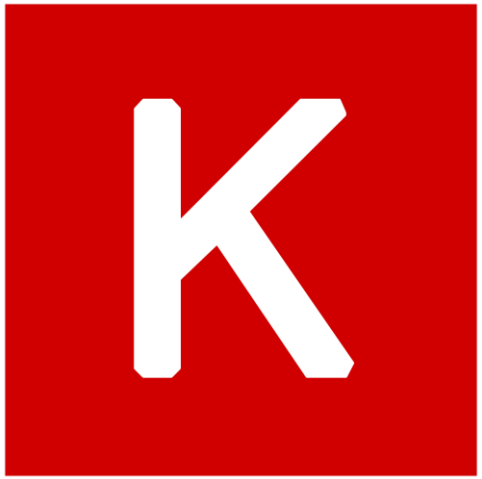
Interessi principali

Allineamento politico generale

Sentimenti espressi durante i tweet

In che modo condivide video (passivo o attivo)

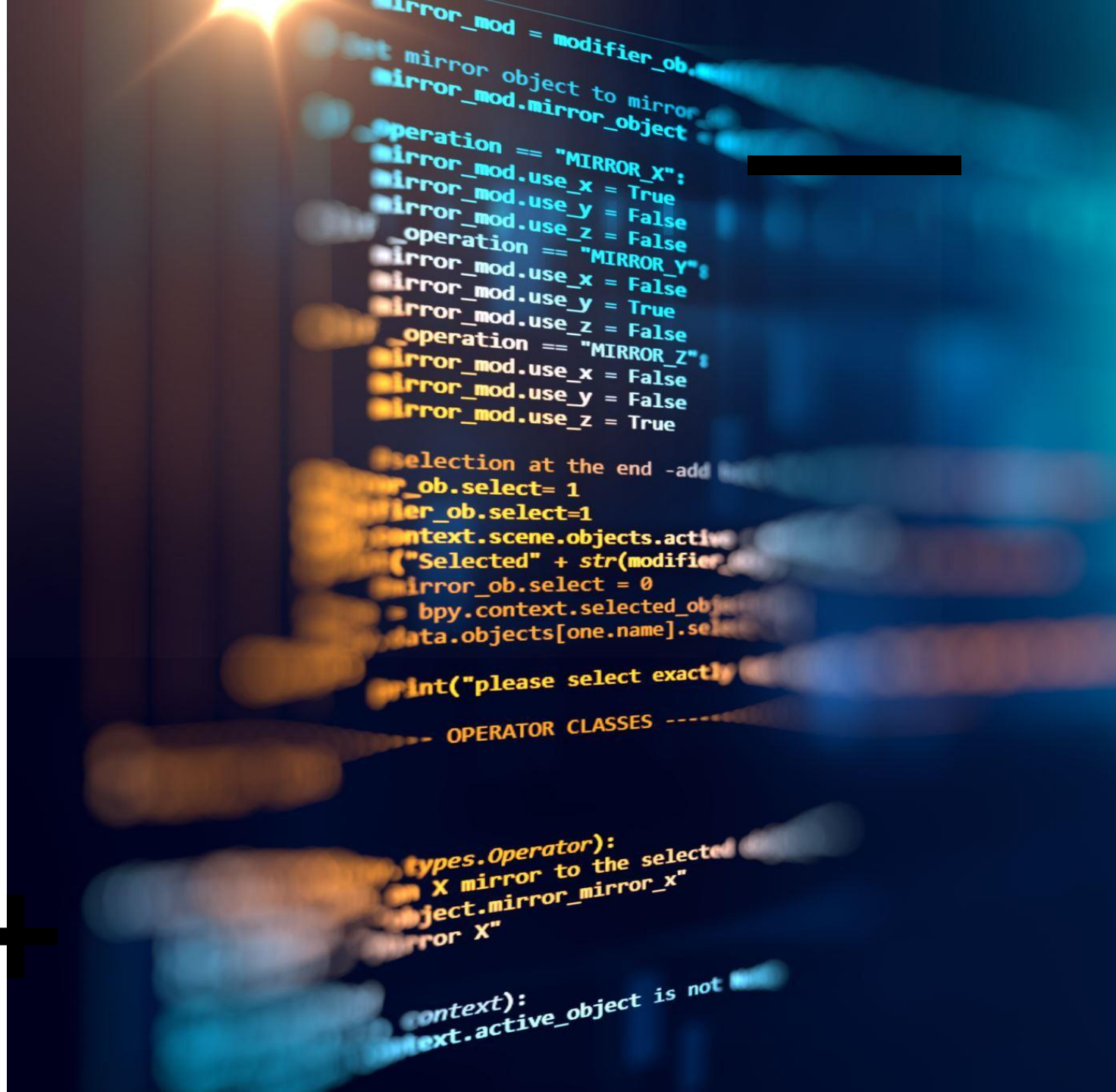
Tool utilizzati



+

Dataset

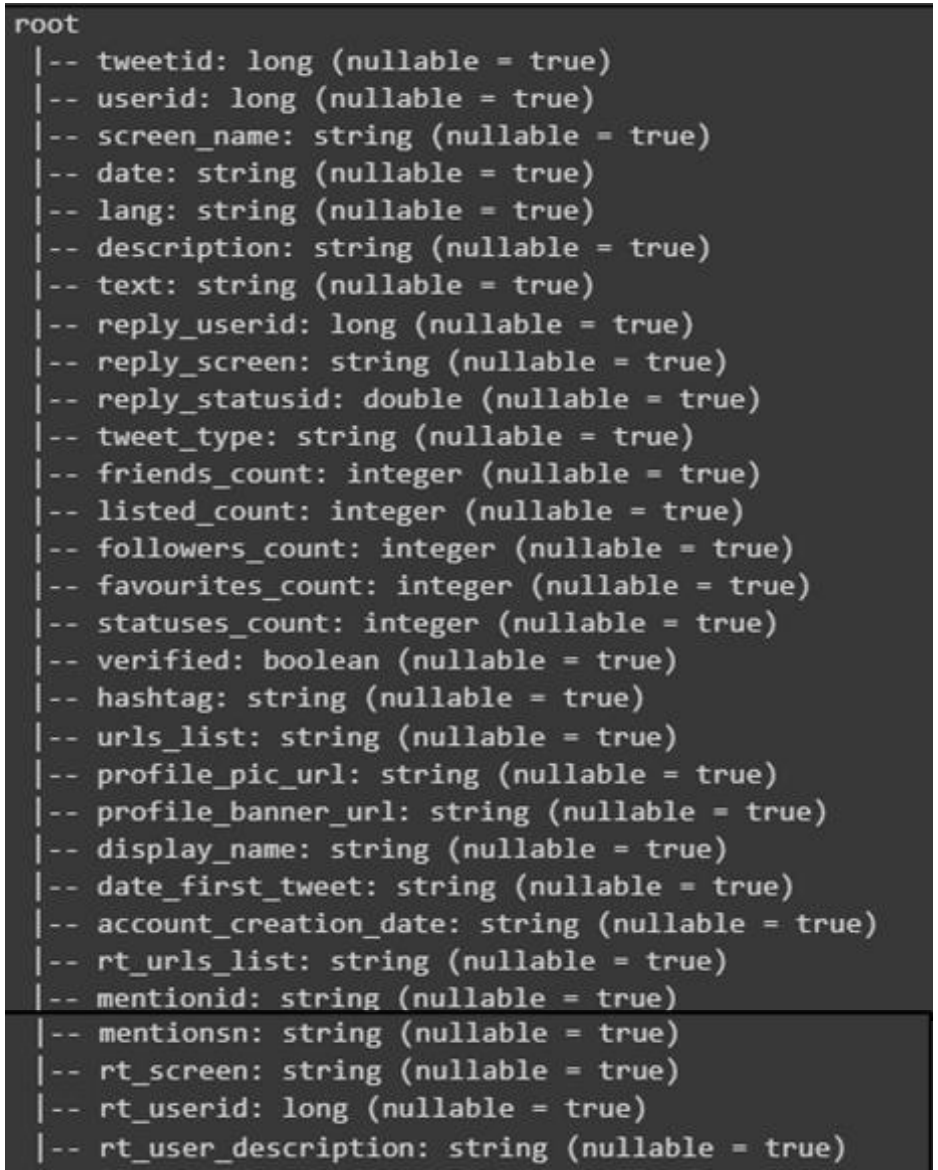
+


mirror_mod = modifier_ob.
mirror object to mirror
mirror_mod.mirror_object
operation == "MIRROR_X":
irror_mod.use_x = True
irror_mod.use_y = False
irror_mod.use_z = False
operation == "MIRROR_Y":
irror_mod.use_x = False
irror_mod.use_y = True
irror_mod.use_z = False
operation == "MIRROR_Z":
irror_mod.use_x = False
irror_mod.use_y = False
irror_mod.use_z = True
election at the end -add
_ob.select= 1
er_ob.select=1
ntext.scene.objects.activ
("Selected" + str(modifier
irror_ob.select = 0
bpy.context.selected_ob
ata.objects[one.name].se
int("please select exactl
OPERATOR CLASSES
types.Operator):
X mirror to the selected
ject.mirror_mirror_x"
ror X"
ontext):
ext.active_object is not

Dataset

- Metadati del tweet e dell’account
- Contenuto del tweet
- Tipo del tweet
- Eventuali dati di tweet correlati

+



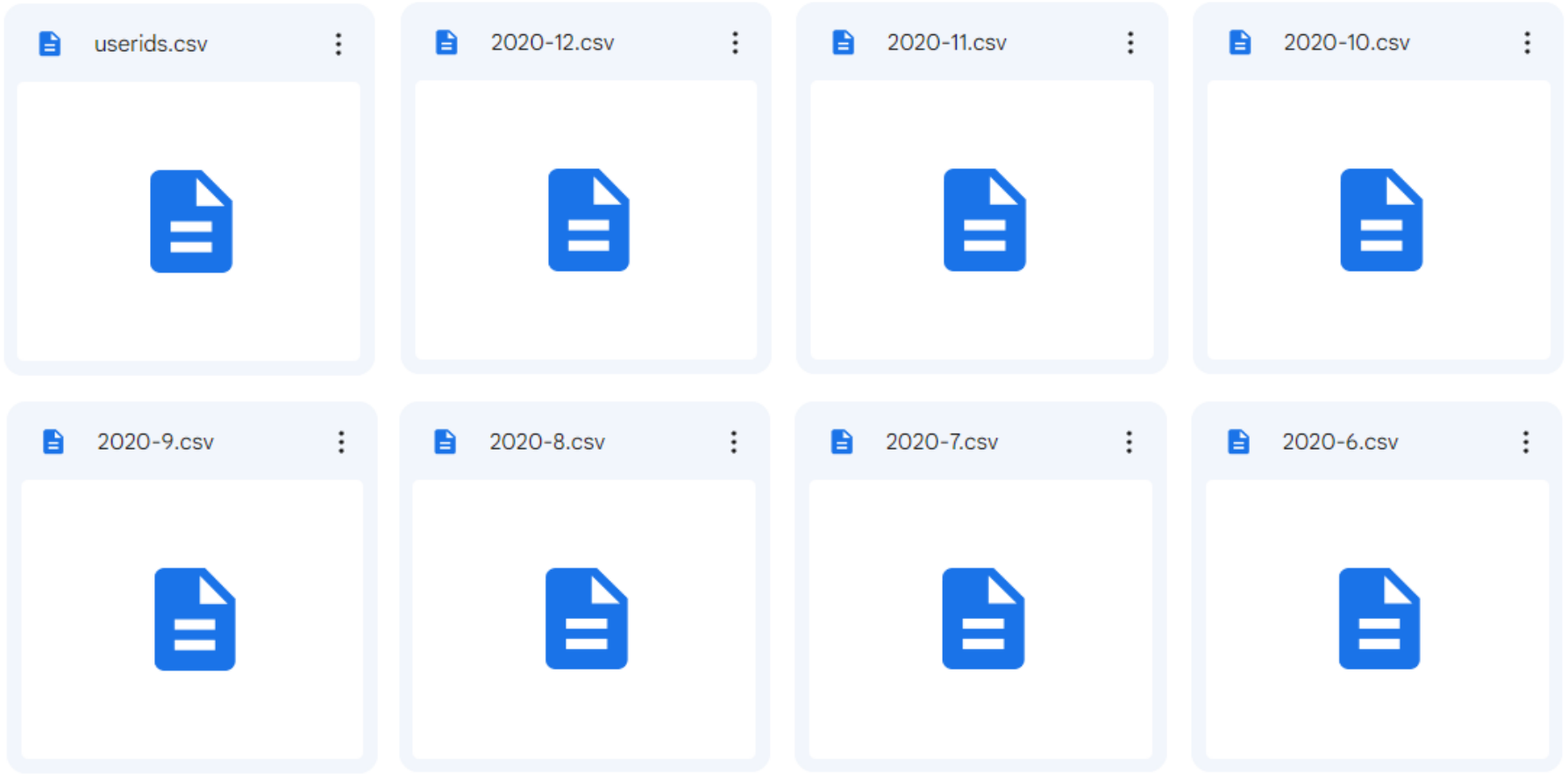
Dataset

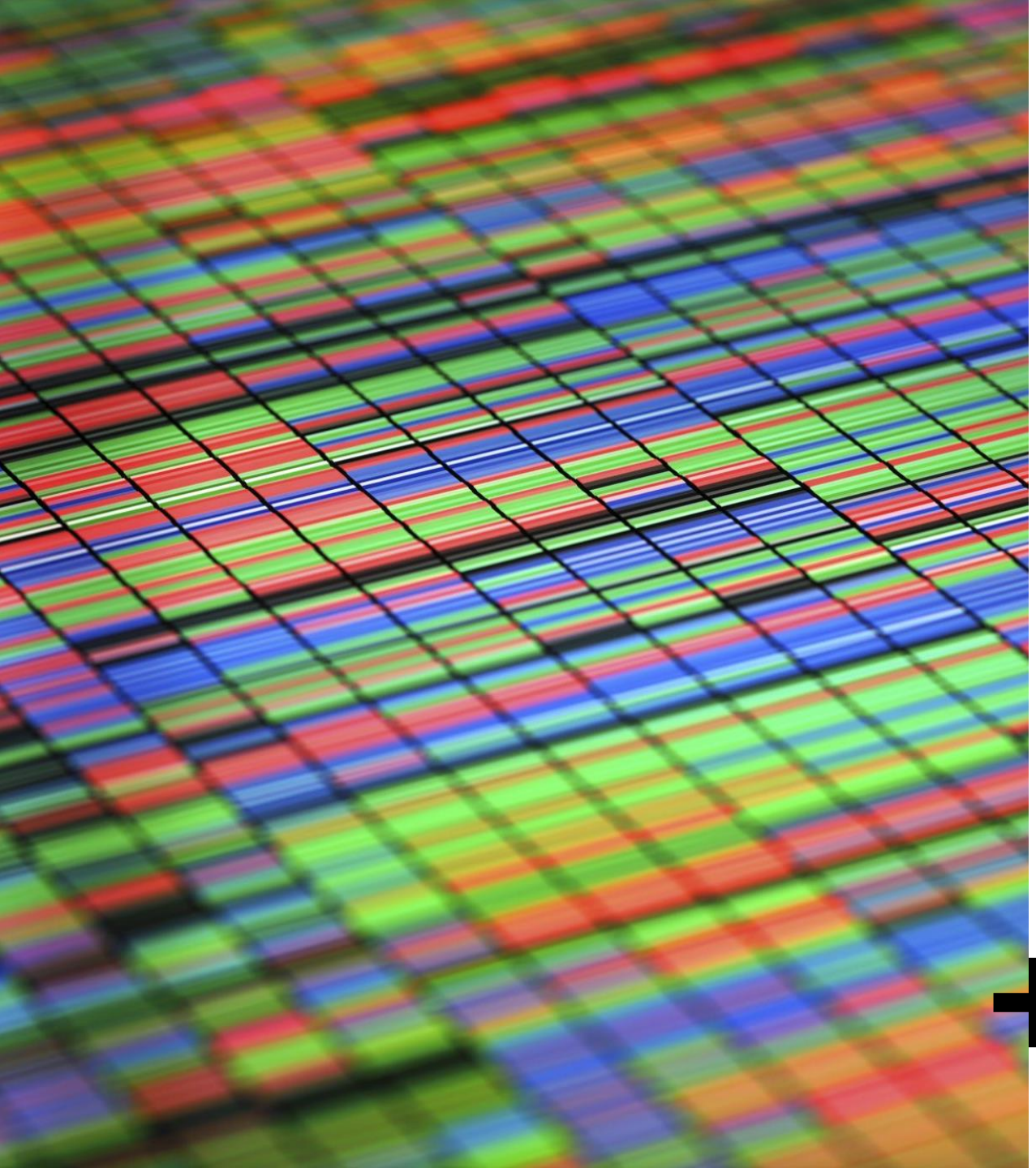
- Original
- Reply
- Retweeted without comments
- Quoted

+

```
root
 |-- tweetid: long (nullable = true)
 |-- userid: long (nullable = true)
 |-- screen_name: string (nullable = true)
 |-- date: string (nullable = true)
 |-- lang: string (nullable = true)
 |-- description: string (nullable = true)
 |-- text: string (nullable = true)
 |-- reply_userid: long (nullable = true)
 |-- reply_screen: string (nullable = true)
 |-- reply_[illegible] (nullable = true)
 |-- tweet_type: string (nullable = true)
 |-- [illegible] true)
 |-- listed_count: integer (nullable = true)
 |-- followers_count: integer (nullable = true)
 |-- favourites_count: integer (nullable = true)
 |-- statuses_count: integer (nullable = true)
 |-- verified: boolean (nullable = true)
 |-- hashtag: string (nullable = true)
 |-- urls_list: string (nullable = true)
 |-- profile_pic_url: string (nullable = true)
 |-- profile_banner_url: string (nullable = true)
 |-- display_name: string (nullable = true)
 |-- date_first_tweet: string (nullable = true)
 |-- account_creation_date: string (nullable = true)
 |-- rt_urls_list: string (nullable = true)
 |-- mentionid: string (nullable = true)
 |-- mentionsn: string (nullable = true)
 |-- rt_screen: string (nullable = true)
 |-- rt_userid: long (nullable = true)
 |-- rt_user_description: string (nullable = true)
```

Dataset


userids.csv
2020-12.csv
2020-11.csv
2020-10.csv
2020-9.csv
2020-8.csv
2020-7.csv
2020-6.csv



Elaborazione dati

APACHE
Spark
TM

+

```
1. data_paths =
   ["/content/drive/MyDrive/DS/2020-
   6.csv","/content/drive/MyDrive/DS/2020-
   7.csv","/content/drive/MyDrive/DS/2020-
   8.csv","/content/drive/MyDrive/DS/2020-
   9.csv","/content/drive/MyDrive/DS/2020-
   10.csv","/content/drive/MyDrive/DS/2020-
   11.csv","/content/drive/MyDrive/DS/2020-
   12.csv",]
2.
3. tweets = spark.read \
4.    .option("inferSchema", True) \
5.    .option("header", True) \
6.    .option("quote", "\"") \
7.    .option("escape", "\"") \
8.    .csv(data_paths, sep=',', multiLine=True)
```

```
1. name1 = "JimBonz"
2. path_query =
   f"/content/drive/MyDrive/DS/Query/{name1}/"
3. attrib = [
4. "account_creation_date","date","description
   ","followers_count","hashtag","lang","menti
   onsn","profile_pic_url","qtd_screen","qtd_u
   rls_list","reply_screen","rt_screen","rt_ur
   ls_list","screen_name","text","tweet_type",
   "tweetid","urls_list","verified"]
5. # Filtro i tweets, ottengo solo quelli
   appartenenti all'utente specificato
6. name1_dataframe =
   tweets.select(*[attrib]).where(tweets.scree
   n_name == name1)
7. print(name1_dataframe.count())
```

TweetNLP

+

```
1.
2. model_e = tweetnlp.load_model('emotion')
3. model_o = tweetnlp.load_model('offensive')
4. model_s = tweetnlp.load_model('sentiment')
5. model_i = tweetnlp.load_model('irony')
6. model_t = tweetnlp.load_model('topic_classification')
```

Keras

+

	Party	Tweet
0	Republican	Today marks the final time that Americans will...
1	Democrat	Spent the morning with Ambassador Mohib and a ...
2	Republican	RT @WMALDC: Listen now as Congressman @RobWitt...
3	Democrat	MUST READ: 7 Gutsy Women To Know For Internati...
4	Democrat	We are proud of our diversity and immigrant co...

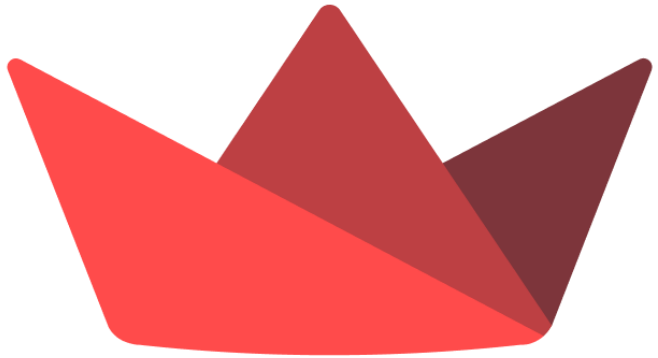
```
model = Sequential()

model.add(Dense(64, activation='relu',
input_shape=(2500,)))

model.add(Dense(1, activation='sigmoid'))
```

	Party	Tweet
0	Republican	Today marks the final time that Americans will...
1	Democrat	Spent the morning with Ambassador Mohib and a ...
2	Republican	RT @WMALDC: Listen now as Congressman @RobWitt...
3	Democrat	MUST READ: 7 Gutsy Women To Know For Internati...
4	Democrat	We are proud of our diversity and immigrant co...

```
model = Sequential()

model.add(Dense(64, activation='relu',
input_shape=(2500,)))

model.add(Dense(1, activation='sigmoid'))
```

```
▾ TfidfVectorizer

TfidfVectorizer(max_df=0.8, max_features=2500, min_df=7)
```

Streamlit



+