20MIA1031
SANJAY.M

Predictive Analytics with Case Studies(CSE3085)-F1

Material for Assignment-2

- 1 TWITTER SCRAPPING execution
- 1.1 Health data
- 1.2 Covid data

- 2.1 GRU theory
- 2.2 EXECUTION

- 3.1 paper1

Sensors for continous monitoring of surgeons cognitive workland in the cardiac operating room

- 3.2 Paper2

Intelligent iCU for autonomous patient monitoring using pervasive sensing and deep learning

➢1 TWITTER SCRAPPING execution

➢1.1 HEALTH data

➢1.2 COVID data

➢3.1 HEALTH data twitter scrapping

PA_DA2_Twitter_scrapping_20MIA1031.ipynb

File Edit View Insert Runtime Tools Help Last edited on April 8

+ Code + Text

20MIA1031

SANJAY.M

Twitter scrapping

```
[ ] import pandas as pd
    import numpy as np
    import re
    import matplotlib.pyplot as plt
    plt.style.use('fivethirtyeight')
```

```
pip install twint
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: twint in ./src/twint (2.1.21)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.8/dist-packages (from twint) (3.8.4)
Requirement already satisfied: aiodns in /usr/local/lib/python3.8/dist-packages (from twint) (3.0.0)
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.8/dist-packages (from twint) (4.6.3)
Requirement already satisfied: cchardet in /usr/local/lib/python3.8/dist-packages (from twint) (2.1.7)
Requirement already satisfied: dataclasses in /root/.local/lib/python3.8/site-packages (from twint) (0.6)
Requirement already satisfied: elasticsearch in /usr/local/lib/python3.8/dist-packages (from twint) (8.6.2)
Requirement already satisfied: pysocks in /usr/local/lib/python3.8/dist-packages (from twint) (1.7.1)
Requirement already satisfied: pandas in /usr/local/lib/python3.8/dist-packages (from twint) (1.3.5)
Requirement already satisfied: aiohttp_socks in /usr/local/lib/python3.8/dist-packages (from twint) (0.8.0)
Requirement already satisfied: schedule in /usr/local/lib/python3.8/dist-packages (from twint) (1.1.0)
Requirement already satisfied: geopy in /usr/local/lib/python3.8/dist-packages (from twint) (1.17.0)
Requirement already satisfied: fake-useragent in /usr/local/lib/python3.8/dist-packages (from twint) (1.1.1)
Requirement already satisfied: googletransx in /usr/local/lib/python3.8/dist-packages (from twint) (2.4.2)
Requirement already satisfied: pycares>=4.0.0 in /usr/local/lib/python3.8/dist-packages (from aiodns->twint) (4.3.0)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (6.0.4)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (1.3.3)
Requirement already satisfied: async-timeout<5.0,>=4.0.0a3 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (4.0.2)
Requirement already satisfied: charset-normalizer<4.0,>=2.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (3.0.1)
Requirement already satisfied: yarl<2.0,>=1.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (1.8.2)
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (1.3.1)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (22.2.0)
Requirement already satisfied: python-socks[asyncio]<3.0.0,>=2.0.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp_socks->twint) (2.1.1)
Requirement already satisfied: elastic-transport<9,>=8 in /usr/local/lib/python3.8/dist-packages (from elasticsearch->twint) (8.4.0)
Requirement already satisfied: importlib-resources>=5.0 in /usr/local/lib/python3.8/dist-packages (from fake-useragent->twint) (5.12.0)
Requirement already satisfied: geographiclib<2,>=1.49 in /usr/local/lib/python3.8/dist-packages (from geopy->twint) (1.52)
Requirement already satisfied: requests in /usr/local/lib/python3.8/dist-packages (from googletransx->twint) (2.25.1)
Requirement already satisfied: python-dateutil>=2.7.3 in /usr/local/lib/python3.8/dist-packages (from pandas->twint) (2.8.2)
Requirement already satisfied: pytz>=2017.3 in /usr/local/lib/python3.8/dist-packages (from pandas->twint) (2022.7.1)
Requirement already satisfied: numpy>=1.17.3 in /usr/local/lib/python3.8/dist-packages (from pandas->twint) (1.22.4)
Requirement already satisfied: certifi in /usr/local/lib/python3.8/dist-packages (from elastic-transport<9,>=8->elasticsearch->twint) (2022.12.7)
Requirement already satisfied: urllib3<2,>=1.26.2 in /usr/local/lib/python3.8/dist-packages (from elastic-transport<9,>=8->elasticsearch->twint) (1.26.14)
Requirement already satisfied: zipp>=3.1.0 in /usr/local/lib/python3.8/dist-packages (from importlib-resources>=5.0->fake-useragent->twint) (3.14.0)
```

```
[ ] pip install nest_asyncio
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: nest_asyncio in /usr/local/lib/python3.8/dist-packages (1.5.6)
```

```
[ ] pip install --user --upgrade -e git+https://github.com/twintproject/twint.git@origin/master#egg=twint;
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Obtaining twint from git+https://github.com/twintproject/twint.git@origin/master#egg=twint
  Updating ./src/twint clone (to revision origin/master)
  Running command git fetch -q --tags
  WARNING: Did not find branch or tag 'origin/master', assuming revision or ref.
  Running command git reset --hard -q origin/master
  Preparing metadata (setup.py) ... done
Requirement already satisfied: aiohttp in /usr/local/lib/python3.8/dist-packages (from twint) (3.8.4)
Requirement already satisfied: aiodns in /usr/local/lib/python3.8/dist-packages (from twint) (3.0.0)
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.8/dist-packages (from twint) (4.6.3)
Requirement already satisfied: cchardet in /usr/local/lib/python3.8/dist-packages (from twint) (2.1.7)
Requirement already satisfied: dataclasses in /root/.local/lib/python3.8/site-packages (from twint) (0.6)
Requirement already satisfied: elasticsearch in /usr/local/lib/python3.8/dist-packages (from twint) (8.6.2)
Requirement already satisfied: pysocks in /usr/local/lib/python3.8/dist-packages (from twint) (1.7.1)
Requirement already satisfied: pandas in /usr/local/lib/python3.8/dist-packages (from twint) (1.3.5)
Requirement already satisfied: aiohttp_socks in /usr/local/lib/python3.8/dist-packages (from twint) (0.8.0)
Requirement already satisfied: schedule in /usr/local/lib/python3.8/dist-packages (from twint) (1.1.0)
Requirement already satisfied: geopy in /usr/local/lib/python3.8/dist-packages (from twint) (1.17.0)
Requirement already satisfied: fake-useragent in /usr/local/lib/python3.8/dist-packages (from twint) (1.1.1)
Requirement already satisfied: googletransx in /usr/local/lib/python3.8/dist-packages (from twint) (2.4.2)
Requirement already satisfied: pycares>=4.0.0 in /usr/local/lib/python3.8/dist-packages (from aiodns->twint) (4.3.0)
Requirement already satisfied: charset-normalizer<4.0,>=2.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (3.0.1)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (22.2.0)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (1.3.3)
Requirement already satisfied: async-timeout<5.0,>=4.0.0a3 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (4.0.2)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (6.0.4)
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (1.3.1)
Requirement already satisfied: yarl<2.0,>=1.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp->twint) (1.8.2)
Requirement already satisfied: python-socks[asyncio]<3.0.0,>=2.0.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp_socks->twint) (2.1.1)
```

```
[ ] import twint
    import pandas as pd
    import nest_asyncio
    nest_asyncio.apply() #necessary to allow asyncio to end event loop
    import nltk
    nltk.download('punkt')
    nltk.download('wordnet')
    from nltk import sent_tokenize, word_tokenize
    from nltk.stem.snowball import SnowballStemmer
    from nltk.stem.wordnet import WordNetLemmatizer
    from nltk.corpus import stopwords
    from textblob import TextBlob
    from wordcloud import WordCloud, STOPWORDS
```

```
[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!
```

```
[ ] s = twint.Config()
    s.Search = "health"
    s.Lang = "en"
    s.Limit = 1000          # optional
    s.Since = '2020-06-27 00:00:00'
    s.until = '2020-12-31 00:00:00'
    #Store script in csv file
    s.Store_csv = True
    s.Output = "health.csv"
    twint.run.Search(s)

    1630047826695843841 2023-02-27 03:31:01 +0000 <AshokMa81098907> Sant Shri Asharamji Bapu अपने सत्संग में बताते हैं #सर्वांत_ऋतुचर्या में उचित आहार विहार की जानकारी द्वारा Swasthya Labh लिया जा सकता है। जिन्हें ऐसे घरेलू, सस्ते व Easy Health Tips जो बड़ी से बड़ी बीमारियों में भी कार
    1630047824196014080 2023-02-27 03:31:00 +0000 <Raghav72179037> @JantaKeeAwaaz Sant Shri Asharamji Bapu बताते हैं #सर्वांत_ऋतुचर्या में Swasthya Labh के बारे में -  इस ऋतु में कफ की कुपित करने वाले पौष्टिक और गरिष्ठ पदार्थों की मात्रा धीरे-धीरे कम करते हुए बंद करके सादा सुपाच्य अ
    1630047822157586437 2023-02-27 03:31:00 +0000 <ConservativeO17> Jill Biden Gives Update On Joe's Health   https://t.co/eTOcSbIEcv
    1630047821872218112 2023-02-27 03:31:00 +0000 <HealthPolicyW> Eight Hundred Women Die Every Day During Pregnancy or Childbirth |  https://t.co/sIV5sdcMnp #ICYMI @JennyCresswell @DrTedros @ABanerjeeHMO  https://t.co/LG&srWbtfM
    1630047814297464832 2023-02-27 03:30:58 +0000 <G1Matthewin> Happineis good health and a bad  https://t.co/fuuOVYFPzI can't stay in your corner of the Forest waiting for others to come to you. You have to go to them sometimes. #APOFLYTOPARIS
    1630047809360793602 2023-02-27 03:30:57 +0000 <Jagdish38746441> Basant Ritu ki health trips #सर्वांत_ऋतुचर्या Sant Shri Asharamji Bapu Easy Health Tips Swasthya Labh
    1630047807078891520 2023-02-27 03:30:56 +0000 <Gc9d7NLVqSP3Kqv> @sahay_sandhya @YssSpeaks #सर्वांत_ऋतुचर्या Easy Health Tips by Sant Shri Asharamji Bapu - इस ऋतु के अंतर्गत आने पर यदि आपको Swasthya Labh चाहिए और बचना हो तथा, हृदय रोगों से, high ব तो bp से से तो बिन
    1630047805447237633 2023-02-27 03:30:56 +0000 <devangee_dabhi> @JantaKeeAwaaz Sant Shri Asharamji Bapu કેવે છે સૂર્યોદય થી પૂર્વ ઉઠવું, ખાલમ, પ્રાણાયમ, ચાલવું વધારે વ્યાયામથી છે #सर्वांत_ऋतुचर्या Easy Health Tips Swasthya Labh  https://t.co/CGpXcAwLPF
    1630047800992956417 2023-02-27 03:30:55 +0000 <thisisakshayds> @aditi_ambadkar Only when it takes a toll on your health and you forget the reason for being ambitious 👍
    1630047799097315330 2023-02-27 03:30:54 +0000 <KaplelHirmala> @LokKalyanSetu Swasthya Labh के लिए Sant Shri Asharamji Bapu #सर्वांत_ऋतुचर्या के Easy Health Tips बताते हुए कहते हैं  नीम की 15-20 कोंपलें 2-3 काली मिर्च के साथ चबा-चबाकर खानी चाहिए। 15-20 दिन यह प्रयोग
    1630047795716423680 2023-02-27 03:30:53 +0000 <ChohanShilpi> We comprise of different #energy levels at physical, mental, emotional and spiritual existence. Do you view your #health and #wellness from this metaphysical perspective?  https://t.co/
    1630047793665691648 2023-02-27 03:30:53 +0000 <BkVishwl11512196> Sant Shri Asharamji Bapu कहते हैं कि  #सर्वांत_ऋतुचर्या में करुणा, सहकार, मैत्री भोजन में लेना चाहिए, इन Easy Health Tips से हम  Swasthya Labh ले सकते हैं 1  https://t.co/6vtoBUc1J0
    1630047792310747138 2023-02-27 03:30:53 +0000 <MBLinton> @KellyMRosati As a family, we've had mental health concerns that weren't covered. I don't feel discriminated against because I knew it going in. I made the choice.  That may be the case, b
    1630047791878746113 2023-02-27 03:30:53 +0000 <DeconstructedAF> He doesn'treally believe that though. If he really felt that way then he would do more to make health care a basic human right. Everyone would have health insurance , health insuranc
    1630047790792421376 2023-02-27 03:30:52 +0000 <GSNAfrica> Unpacking Trauma: Exploring Its Impact on Mental Health and Well-being  1/6 In this thread, we'll be unpacking the nature of trauma, exploring its different forms and effects, and consider
    1630047788668387074 2023-02-27 03:30:52 +0000 <MageshSurwase2> @sabhiin1 पूज्य Sant Shri Asharamji Bapu #सर्वांत_ऋतुचर्या के लिए Easy Health Tips बताते हुए कहते हैं कि इस ऋतु में सिरदर्द, नेत्रे, सूजी और भारी भोजन नहीं करना चाहिए अधिक हलका,सुपाच्य भोजन करना चाहिए। 15 दिन ।
    1630047788624121858 2023-02-27 03:30:52 +0000 <MrPuppetwatcher> @NewsBFM I couldn't help but sympathies with the kid(like everyone else) because let's face it , exams pretty much make students not only feel frustated , but also sometime cause pro
    1630047786770264067 2023-02-27 03:30:51 +0000 <the16213> 10/10 kayares couple sinting, pusing aku baca nya wrwrwr. au ini bahas mental health kedua tokoh utamanya ig. must read  https://t.co/L7uA8JYXQS
    1630047784949948224 2023-02-27 03:30:51 +0000 <anildhiman52> #सर्वांत_ऋतुचर्या Sant Shri Asharamji Bapu give Easy Health Tips to get Swasthya Labh 🙏🔥 आहार : इस ऋतु में देर से पचनेवाले, शीतल पदार्थ, दिन में सोना, स्निग्ध अर्थात् घी-तेल में बने तथा अम्ल व मधुर रसप्रधान पदा
    1630047783418843140 2023-02-27 03:30:51 +0000 <Sthealth34> @Olits1olts Foote and MULTIPLE picks. Don't like this move
    1630047779153387521 2023-02-27 03:30:50 +0000 <Soni42364426> @RupaVis88060861 @JantaKeeAwaaz जी हाँ Sant Shri Asharamji Bapu सत्संग में बताते हैं की #सर्वांत_ऋतुचर्या में देर से पचनेवाले, शीतल पदार्थ, दिन में सोना, घी-तेल से बना तथा अम्ल व मधुर रसप्रधान पदार्थों का सेवन न करें क्यों
    1630047769149984770 2023-02-27 03:30:47 +0000 <Lalitvt39589453> @Itz_nandeeniii @JantaKeeAwaaz Rightly said Sant Shri Asharamji Bapu #सर्वांत_ऋतुचर्या Easy Health Tips Swasthya Labh  https://t.co/EMu6H6cEea
    1630047768663162880 2023-02-27 03:30:47 +0000 <mandoctorian> @hormonedietdoc Keto, carnivore, vegetarian, vegan, south beach, paleo… bet you can find someone saying 'x' diet helped 'x' inflammatory disease. Wish it was one size fits all.  One day
    1630047768658588163 2023-02-27 03:30:47 +0000 <airsoftkumasan1> Health Care Products  https://t.co/FMqeK4BLOP  via @amazon  12:304322  https://t.co/uBCBylDhHu

[ ] s.Store_json = True #Save in Json format
    s.Output = "health.json" #Output directory to save file

[ ] #Run
    twint.run.Search(s) #Run twint script
    #This we display all the

    1630047963719532544 2023-02-27 03:31:34 +0000 <ANOOPKU29367012> @ashramgoregaon Sant Shri Asharamji Bapu #सर्वांत_ऋतुचर्या में हल्का व सुपाच्य भोजन लेना चाहिए।कफ को अपशोषित करने वाले पदार्थ हितकारी हैं।भारी व गरिष्ठ पदार्थ नहीं लेने चाहिए। Easy Health Tips Swasthya Labh  https://t.co
    1630047962737913857 2023-02-27 03:31:33 +0000 <AnuoluwapoDs> I got global disrespect global abuse global harrd global attacks all my pregnancy eliminated constantly on site i get attacked when  I sleep o get abused cctv voices ruins my mental heath everyon
    1630047962595467264 2023-02-27 03:31:33 +0000 <nlayons> Tapi kena pulau, kena fitnah banyak le kot. Mental health terganggu. Aku ada short term memory. Tak berapa nak ingat dah macamana duduk asrama. Hshsnshsnsh
    1630047961602772485 2023-02-27 03:31:33 +0000 <Archana97430714> इस सीजन में कफ बढ़ेगा कफ दूर करने के लिए गोचरण बटी सुबहाग खाने  से कफ दूर हो जाता है। Sant Shri Asharamji Bapu #सर्वांत_ऋतुचर्या Easy Health Tips Swasthya Labh  https://t.co/qxwJ47wGL2
    1630047960515096576 2023-02-27 03:31:33 +0000 <Halnuresun117> @SKOfficial369 @JantaKeeAwaaz Swasthya Labh Easy Health Tips देवांत ऋतु में बाणमाँ खाँसी से लोग परेशान रहते हैं और फिर अंग्रेजी दवा लेना तो ज़हर लेने के बराबर है। Sant Shri Asaram Bapu Ji बताते हैं कि 10ग्राम शहद, 2 ग्राम
    1630047960112267265 2023-02-27 03:31:33 +0000 <fmlnga> @emuse I question a wife who would go on vacation while her spouse is hospitalized.  She should be the main person encouraging him to resign  in the best interest of his health and the citizens of PA.
    1630047959961468928 2023-02-27 03:31:33 +0000 <r_bhart> @balramomi @JantaKeeAwaaz हो पूज्य Sant Shri Asharamji Bapu सत्संग में #सर्वांत_ऋतुचर्या के विषय में बहुत सुंदर जानकारी दी है। Easy Health Tips Swasthya Labh  https://t.co/xJlCeuB5fd
    1630047955674882048 2023-02-27 03:31:32 +0000 <junetwentyone> I've booked my next dose of that "drug", asshole! That and following health advice from actual experts have kept me free of the disease this past three years.  Woody is a brilliant actor, but i
    1630047955381104640 2023-02-27 03:31:32 +0000 <babycandlelight> But someone deserves a person who is good for his mental health too.🌻
    1630047953434845184 2023-02-27 03:31:32 +0000 <Gc9d7NLVqSP3Kqv> @manishaParaswa2 @YssSpeaks #सर्वांत_ऋतुचर्या Easy Health Tips by Sant Shri Asharamji Bapu - इस ऋतु में देर तक जागना, देर तक बुखार सोने से मल सूखता है, नेत्र व मुख की कांति क्षीण होती हैं अंतः Swasthya Labh हेतु इन आ
    1630047952952492033 2023-02-27 03:31:31 +0000 <wFomentar> @VoidScreamerUSA @Kontana_LP War is the health of the state
    1630047950305914881 2023-02-27 03:31:30 +0000 <TeamRwB> Recovering from a service-related traumatic brain injury, Army infantry veteran Ira Brownridge Jr. found his tribe in Team RwB.  Today, he's forging a health and wellness community to support other v
    1630047946174545921 2023-02-27 03:31:29 +0000 <Kirannrajdev> मेरी Sant Shri Asharamji Bapu की ज्ञान  Easy Health Tips का ध्यान रखें और  Swasthya Labh पाएं #सर्वांत_ऋतुचर्या
    1630047943884517380 2023-02-27 03:31:29 +0000 <yuvrajharion1> जी हाँ, आपने बिल्कुल सही कहाISant Shri Asharamji Bapu  #सर्वांत_ऋतुचर्या  Easy Health Tips  Swasthya Labh
    1630047943649533953 2023-02-27 03:31:29 +0000 <KAOS_the_legend> @CryptoWendyO If you were in my area, I'd invite you to see me for some massage work.  That said, I'm a big health nerd and have had my own on going sleep troubles.  thread
    1630047943079100416 2023-02-27 03:31:29 +0000 <anhtuandn3> Limone sul Garda, Italy's village with a health 'elixir'  https://t.co/1QUOOJONTa
    1630047940449987889 2023-02-27 03:31:28 +0000 <MaheshKhade20> #सर्वांत_ऋतुचर्या  Easy Health Tips told by Sant Shri Asharamji Bapu for Swasthya Labh 👇 Do not consume heavy to digest,cold substances, made in ghee/oil, acidic &amp;amp; sweet juicy substances in sprin
    1630047940411705221 2023-02-27 03:31:28 +0000 <bk5050954> Sant Shri Asharamji Bapu ने कि वसंत ऋतु में कफज़ा उपद्रव रखने से मस नाड़ियों की कार्य शैली बेहतर होती है और आरोग्य की रक्षा होती है। Easy Health Tips Swasthya Labh #सर्वांत_ऋतुचर्या  https://t.co/np86271t2i
    1630047939078012928 2023-02-27 03:31:28 +0000 <amkitkum42575127> Sant Shri Asharamji Bapu  #सर्वांत_ऋतुचर्या -इस ऋतु में कफ को कुपित करने वाले पौष्टिक और गरिष्ठ पदार्थों की मात्रा धीरे-धीरे कम करते हुए बंद करके सादा सुपाच्य आहार लेना सुरू कर देना चाहिए। Easy Health Tips ❗  Swasthya L
    1630047937853022211 2023-02-27 03:31:27 +0000 <kayleerenee113> Protect health behavior at believe maintain.  #ᴋᴀᴛᴇ #ᴋᴀᴛᴇ ᴋ ᴋᴀᴛᴇ ᴋ #ᴋᴀᴛᴇ #ᴋᴀᴛᴇ #ᴋᴀᴛᴇ #ᴋᴀᴛᴇ #ᴋᴀᴛᴇ  https://t.co/AIdcT5uGZG
    1630047932408799234 2023-02-27 03:31:26 +0000 <HiteshDevasi14> @asharamjiashram पूज्य Sant Shri Asharamji Bapu जी विश्व मानव के कल्याण के हित में करते हैं सत्संग उनमें से #सर्वांत_ऋतुचर्या 🟦Swasthya Labh 🔥 Easy Health Tips का ये दिव्य सत्संग सुनके हमको मिलेगा ज्ञान क्या खाएं क्या ना खा
    1630047925060370434 2023-02-27 03:31:24 +0000 <LouBen81129402> Praying for strength motivation and good health 🔥
    1630047922996797444 2023-02-27 03:31:24 +0000 <RituKeswani3> @JantaKeeAwaaz Sant Shri Asharamji Bapu emphasis to follow_Vaidic_Gyan in every field of life,He says about  #सर्वांत_ऋतुचर्या that easy डाइजेस्टेबल food should be taken! Easy Health Tips give Swasthya La
    1630047915371794433 2023-02-27 03:31:22 +0000 <Nishi285harma> @8f4c5e58aa5e48e @ParchaniRekha @JantaKeeAwaaz क्या बाहिया-बाहिया तरीके बताते हैं सत्संग रहने कोSant Shri Asharamji Bapu. #सर्वांत_ऋतुचर्या Easy Health Tips में नीम के साथ तुलसी के पत्ते खाने भी बहुत लाभदायक होते हैं। Swasthy
    1630047912033136640 2023-02-27 03:31:21 +0000 <pradips39678508> @SolankiVardaram @SKOfficial369 @JantaKeeAwaaz #सर्वांत_ऋतुचर्या Easy Health Tips by Sant Shri Asharamji Bapu - इस ऋतु के अंतर्गत आने पर यदि आपको Swasthya Labh चाहिए और बचना हो तथा, हृदय रोगों से, high ব तो bp
    1630047908109904320 2023-02-27 03:31:20 +0000 <visoesdarache1> META: choose people who are good for you mental health
    1630047902763737088 2023-02-27 03:31:19 +0000 <anikshalidi> Hoarding disorder is a mental health problem indeed
    1630047902138769410 2023-02-27 03:31:19 +0000 <Yuvraj559300049> @asharamjiashram जी हाँ Sant Shri Asharamji Bapu कहते हैं कि इन दिनों जिस के तेल की मालिश करना और साफ धान्य उबटन लगाकर स्नान करना स्वास्थ्य को बढ़ाता है। #सर्वांत_ऋतुचर्या Easy Health Tips Swasthya Labh  https://t
    1630047899047408130 2023-02-27 03:31:18 +0000 <JustinoPA83030> @steEnginoov He's been going through potentially life threatening issues with his health since 2017  so unless you can find proof she abandoned him multiple times during the worst times in hi

[ ] #code to help increase the number of columns that can be displayed.
    pd.options.display.max_columns= 100 #Set to display a maximum of 100 columns

[ ] #Since the data is a json file, we will need to read data using Pandas
    stand_df = pd.read_csv("health.csv") # Directory Path to store Data

[ ]
    stand_df.shape #Shapes gives the row & columns

    (1017, 36)

[ ]
    stand_df.head() #Top 5 rows
```

	id	conversation_id	created_at	date	time	timezone	user_id	username	name	place	tweet	language	mentions	urls	photos	replies_count	retwe
0	1630047826695843841	1630047826695843841	2023-02-27 03:31:01 UTC	2023-02-27	03:31:01	0	1482655199906996225	ashokma81098907	Ashok Mahesthwari	NaN	Sant Shri Asharamji Bapu अपने सत्संग में बताते...	hi	[]	['https://youtu.be/KX8LN21QOC0']	[]	0	
1	1630047824196014080	1630038181130424320	2023-02-27 03:31:00 UTC	2023-02-27	03:31:00	0	1290292414934990848	raghav72179037	Raghavendra	NaN	@JantaKeeAwaaz Sant Shri Asharamji Bapu बताते ...	hi	[]	[]	['https://pbs.twimg.com/media/FpBX2M7WIAMKiov....	0	
2	1630047822157586437	1630047822157586437	2023-02-27 03:31:00 UTC	2023-02-27	03:31:00	0	1530047833231410178	conservativeo17	Conservative Opinion	NaN	Jill Biden Gives Update On Joe's Health http...	en	[]	['https://conservativeon.com/2022/07/25/jill-b...	[]	0	
3	1630047821872218112	1630047821872218112	2023-02-27 03:31:00 UTC	2023-02-27	03:31:00	0	930424388855435264	healthpolicyw	Health Policy Watch - Global Health News Repor...	NaN	Eight Hundred Women Die Every Day During Pregn...	en	['screen_name': 'jennycresswell', 'name': 'je...	['https://bit.ly/3XSHH0N']	['http://memory.You']	[]	0
4	1630047814297464832	1630047814297464832	2023-02-27 03:30:58 UTC	2023-02-27	03:30:58	0	1109652518257065984	g1matthewin	wmwinwin-	NaN	Happineis good health and a bad https://t.co/...	en	[]	[]	[]	0	

```
[ ] stand_df.columns # Observe the Columns header gotten using Twint

    Index(['id', 'conversation_id', 'created_at', 'date', 'time', 'timezone',
           'user_id', 'username', 'name', 'place', 'tweet', 'language', 'mentions',
           'urls', 'photos', 'replies_count', 'retweets_count', 'likes_count',
           'hashtags', 'cashtags', 'link', 'retweet', 'quote_url', 'video',
           'thumbnail', 'near', 'geo', 'source', 'user_rt_id', 'user_rt',
           'retweet_id', 'reply_to', 'retweet_date', 'translate', 'trans_src',
           'trans_dest'],
          dtype='object')

[ ] #I want to get the "Tweet" Column from the Json table.
    Tweet = stand_df.filter(['tweet'])

[ ] Tweet.head(10)
```

	tweet
0	Sant Shri Asharamji Bapu अपने सत्संग में बताते...
1	@JantaKeeAwaaz Sant Shri Asharamji Bapu बताते ...
2	Jill Biden Gives Update On Joe's Health http...
3	Eight Hundred Women Die Every Day During Pregn...
4	Happineis good health and a bad https://t.co/...
5	Basant Ritu ki health trips #सर्वांत_ऋतुचर्या San...
6	@sahay_sandhya @YssSpeaks #सर्वांत_ऋतुचर्या Easy ...
7	@JantaKeeAwaaz Sant Shri Asharamji Bapu કેવે છે...
8	@aditi_ambadkar Only when it takes a toll on y...
9	@LokKalyanSetu Swasthya Labh के लिए Sant Shri ...

```
[ ] Tweet['tweet']

    0       Sant Shri Asharamji Bapu अपने सत्संग में बताते...
    1       @JantaKeeAwaaz Sant Shri Asharamji Bapu बताते ...
    2       Jill Biden Gives Update On Joe's Health   http...
    3       Eight Hundred Women Die Every Day During Pregn...
    4       Happineis good health and a bad  https://t.co/...
                                  ...
    1012    @RiyaRan62163524 @JantaKeeAwaaz Sant Shri Asha...
    1013    @tcarpenter216 @JoeWasserman @Cucho @Cassie_Bo...
    1014    Vaccinated Latin Americans Halting Covid; Not ...
    1015    If you get health care coverage through Medica...
    1016    The Chautauqua County Health Department announ...
    Name: tweet, Length: 1017, dtype: object
```

```
[ ] #Clean the text
    import re

    #Create a function to clean the tweets
    def cleanTxt(text):
        text = re.sub(r'@[A-Za-z0-9]+', '', text) #Remove @mentions replace with blank
        text = re.sub(r'#', '', text) #Remove the '#' symbol, replace with blank
        text = re.sub(r'RT[\s]+', '', text) #Removing RT, replace with blank
        text = re.sub(r'https?:\/\/\S+', '', text) #Remove the hyperlinks

        return text

    #Cleaning the text
    Tweet['tweet']= Tweet['tweet'].apply(cleanTxt)

    #Show the clean text
    Tweet
```

	tweet
0	Sant Shri Asharamji Bapu अपने सत्संग में बताते...
1	Sant Shri Asharamji Bapu बताते हैं दसंत_ऋतुचर...
2	Jill Biden Gives Update On Joe's Health
3	Eight Hundred Women Die Every Day During Pregn...
4	Happineis good health and a bad can't stay i...
...	...
1012	Sant Shri Asharamji Bapu दसंत_ऋतुचर्या के E...
1013	_Boness For the latter, it *is* the case th...
1014	Vaccinated Latin Americans Halting Covid; Not ...
1015	If you get health care coverage through Medica...
1016	The Chautauqua County Health Department announ...

1017 rows × 1 columns

```
#Next we have to remove emoji from the Tweet data.
def remove_emoji(string):
    emoji_pattern = re.compile("["
                               u"\U0001F600-\U0001F64F"  # emoticons
                               u"\U0001F300-\U0001F5FF"  # symbols & pictographs
                               u"\U0001F680-\U0001F6FF"  # transport & map symbols
                               u"\U0001F1E0-\U0001F1FF"  # flags (iOS)
                               u"\U00002500-\U00002BEF"  # chinese char
                               u"\U00002702-\U000027B0"
                               u"\U00002702-\U000027B0"
                               u"\U000024C2-\U0001F251"
                               u"\U0001F926-\U0001F937"
                               u"\U00010000-\U0010ffff"
                               u"\u2640-\u2642"
                               u"\u2600-\u2855"
                               u"\u200d"
                               u"\u23cf"
                               u"\u23e9"
                               u"\u231a"
                               u"\ufe0f"  # dingbats
                               u"\u3030"
                               "]+", flags=re.UNICODE)
    return emoji_pattern.sub(r'', string)

#Cleaning the text
Tweet['tweet']= Tweet['tweet'].apply(remove_emoji)

#Show the clean text
Tweet
```

	tweet
0	Sant Shri Asharamji Bapu अपने सत्संग में बताते...
1	Sant Shri Asharamji Bapu बताते हैं दसंत_ऋतुचर...
2	Jill Biden Gives Update On Joe's Health
3	Eight Hundred Women Die Every Day During Pregn...
4	Happineis good health and a bad can't stay i...
...	...
1012	Sant Shri Asharamji Bapu दसंत_ऋतुचर्या के E...
1013	_Boness For the latter, it *is* the case th...
1014	Vaccinated Latin Americans Halting Covid; Not ...
1015	If you get health care coverage through Medica...
1016	The Chautauqua County Health Department announ...

1017 rows × 1 columns

```
[ ] #Create a function to get the subjectivity
    def getSubjectivity(text):
        return TextBlob(text).sentiment.subjectivity

    #Create a function to get Polarity
    def getPolarity(text):
        return TextBlob(text).sentiment.polarity

    #Now we create a new column for what we just did and add it to the Tweet_df dataframe
    Tweet['Subjectivity'] = Tweet['tweet'].apply(getSubjectivity)
    Tweet['Polarity'] = Tweet['tweet'].apply(getPolarity)

    #Now display data
    Tweet
```

	tweet	Subjectivity	Polarity
0	Sant Shri Asharamji Bapu अपने सत्संग में बताते...	0.833333	4.333333e-01
1	Sant Shri Asharamji Bapu बताते हैं दसंत_ऋतुचर...	0.833333	4.333333e-01
2	Jill Biden Gives Update On Joe's Health	0.000000	0.000000e+00
3	Eight Hundred Women Die Every Day During Pregn...	0.000000	0.000000e+00
4	Happineis good health and a bad can't stay i...	0.633333	5.551115e-17
...	...	...	...
1012	Sant Shri Asharamji Bapu दसंत_ऋतुचर्या के E...	0.833333	4.333333e-01
1013	_Boness For the latter, it *is* the case th...	0.427778	-8.333333e-02
1014	Vaccinated Latin Americans Halting Covid; Not ...	0.300000	2.000000e-01
1015	If you get health care coverage through Medica...	0.000000	0.000000e+00
1016	The Chautauqua County Health Department announ...	0.000000	0.000000e+00

1017 rows × 3 columns

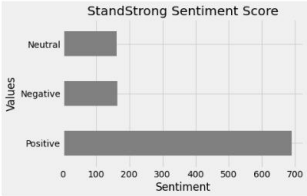
```python
#Group the range of Polarity to different categories
def getInsight(score):
    if score < 0:
        return "Negative"
    elif score == 0:
        return "Neutral"
    else:
        return "Positive"

Tweet["Insight"] = Tweet["Polarity"].apply(getInsight)

Tweet.head(50)
```

	tweet	Subjectivity	Polarity	Insight
0	Sant Shri Asharamji Bapu अपने सत्संग में बताते...	0.833333	4.333333e-01	Positive
1	Sant Shri Asharamji Bapu बताते हैं वसंत_ऋतुचर...	0.833333	4.333333e-01	Positive
2	Jill Biden Gives Update On Joe's Health	0.000000	0.000000e+00	Neutral
3	Eight Hundred Women Die Every Day During Pregn...	0.000000	0.000000e+00	Neutral
4	Happineis good health and a bad can't stay i...	0.633333	5.551115e-17	Positive
5	Basant Ritu ki health trips वसंत_ऋतुचर्या Sant...	0.833333	4.333333e-01	Positive
6	_sandhya वसंत_ऋतुचर्या Easy Health Tips by Sa...	0.686667	2.966667e-01	Positive
7	Sant Shri Asharamji Bapu કહે છે સૂર્યોદય થી ૫...	0.833333	4.333333e-01	Positive
8	_ambadkar Only when it takes a toll on your he...	0.875000	1.250000e-01	Positive
9	Swasthya Labh के लिए Sant Shri Asharamji Bapu...	0.833333	5.416667e-01	Positive
10	We comprise of different energy levels at phys...	0.345238	-2.000000e-02	Negative
11	Sant Shri Asharamji Bapu कहते हैं कि वसंत_ऋतु...	0.833333	4.333333e-01	Positive
12	As a family, we've had mental health concerns...	0.100000	-5.000000e-02	Negative
13	He doesn'treally believe that though. If he re...	0.362245	8.367347e-02	Positive
14	Unpacking Trauma: Exploring Its Impact on Ment...	0.400000	-5.000000e-02	Negative
15	पूज्य Sant Shri Asharamji Bapu वसंत_ऋतुचर्या ...	0.833333	4.333333e-01	Positive
16	I couldn't help but sympathies with the kid(I...	0.590476	1.583333e-01	Positive
17	10/10 kayares couple sinting, pusing aku baca ...	0.200000	-1.000000e-01	Negative
18	वसंत_ऋतुचर्या Sant Shri Asharamji Bapu give Ea...	0.833333	4.333333e-01	Positive

```python
Tweet["Insight"].value_counts()
```

```
Positive    689
Negative    165
Neutral     163
Name: Insight, dtype: int64
```

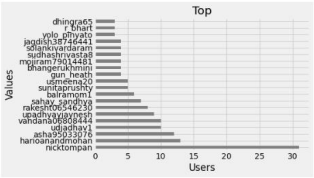
```python
import matplotlib.pyplot as plt

#Plot the values count of polarity
plt.title("StandStrong Sentiment Score")
plt.xlabel("Sentiment")
plt.ylabel("Values")
Tweet["Insight"].value_counts().plot(kind="barh", color="Gray")
plt.show()
```




```python
#Top Followers by tweet
stand_df["username"].value_counts()
```

```
nicktompan         31
harioanandmohan    13
asha95033076       12
udjadhav1          10
vandana06808444    10
                   ..
onye_emeakwari      1
coleppc1            1
gfienberg17         1
ace567ace           1
wgrz                1
Name: username, Length: 805, dtype: int64
```

```python
#Plot the values count of polarity
plt.title("Top")
plt.xlabel("Users")
plt.ylabel("Values")
stand_df["username"].value_counts()[:20].plot(kind="barh", color="Gray")
plt.show()
```

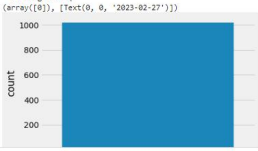



```python
import seaborn as sns
import warnings
```

```python
stand_df["date"] = pd.to_datetime(stand_df["date"]).dt.date
#Helps change the datatype from datetime to date only
```

```python
sns.countplot(stand_df["date"])
plt.xticks(rotation = 45)
```

```
/usr/local/lib/python3.8/dist-packages/seaborn/_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explici
  warnings.warn(
(array([0]), [Text(0, 0, '2023-02-27')])
```



```
[ ] stopwords = STOPWORDS
    print(stopwords)
    #Let checkout the stop words in Python
```

```
{'again', 'my', 'aren't", 'off', "you've", 'if', 'ought', 'can', 'have', 'further', 'shall', 'those', 'after', 'some', 'be', 'else', "he'd", 'is', 'hers', 'not', 'he', 'she', "i'll", "wouldn't", 'from', 'under', 'the', 'about', 'betw
```

```
[ ] #Let create a wordcloud for better representation of data
    Tweet_Word = ' '.join( [twts for twts in stand_df['tweet']] ) #To join all tweet
    wc = WordCloud(
            background_color = "black",
            stopwords = stopwords,
            height = 1000,
            width = 1500,
            random_state = 21,
            max_font_size = 120
    )
    wc.generate(Tweet_Word)

    plt.imshow(wc, interpolation = 'bilinear')
    plt.axis('off')
    plt.show()
```

➢3.2 Covid data twitter scrapping

covid3 data

```
[ ] s = twint.Config()
    s.Search = "covid3"
    s.Lang = "en"
    s.Limit = 1000          # optional
    s.Since = '2020-06-27 00:00:00'
    s.until = '2020-12-31 00:00:00'
    #Store script in csv file
    s.Store_csv = True
    s.Output = "covid3.csv"
    twint.run.Search(s)
```

```
1629681202188767235 2023-02-26 03:14:11 +0000 <giovann72> @kanfootballclub Top =Rea  flop Pantemis
1628907257990610944 2023-02-23 23:58:48 +0000 <giovann72> @Aaron_Derfel Your still at with the Covid guess the fifth sixth jab ain't working quick double mask and time for number six … one trick pony
1628514281825136641 2023-02-22 21:57:15 +0000 <cmdarwindavid> Covid³
1628489203901235203 2023-02-22 20:17:36 +0000 <MrPoleador> @casillistaaa Rico si el Covid3.0
1628226835405320192 2023-02-22 02:55:03 +0000 <giovann72> @manofbird Looks like two donkeys too me so beastality
1627285476913688578 2023-02-19 12:34:25 +0000 <jucamirand> E assim nasce o CoVID3
1627106045964300288 2023-02-19 00:41:26 +0000 <giovann72> @BBNMedias @OneFootball Thanks 🙏
1627098575762071552 2023-02-19 00:11:45 +0000 <giovann72> @BBNMedias Is the game Vs rowdies on Apple TV ?
1626869193131057152 2023-02-18 09:00:16 +0000 <Tia_Pryce> @thehill Ukraine another excuse for the Govt to shake down American Citizens for 100 Billion. If its not cash its taking their homes and businesses. #Iraq #Covid #2008Cras
1626680136408899585 2023-02-17 20:29:01 +0000 <MJnanostretch> @Ayjchan But it is a huge loss seeing that no agreement had been put in place that would've triggered proper immediate independent origin investigations in Wuhan. The
1625480720755085312 2023-02-14 13:02:58 +0000 <Karthik05488596> @HissChief01 @HrHadMeta @satishshankar1 @ikamalhaasan Neengala yawning panna mattinga la ungaluKku itha Vida periya disease varum appo theriyum namma Vikram a summa
1624501232016195585 2023-02-12 01:28:43 +0000 <TRNowry> @KEENSTAR @Cernovich An act of war that gets you nuked is blowing up a city or pelting buildings with planes.  An act of war that cripples a country but avoids getting you
1624539060223021058 2023-02-11 22:41:09 +0000 <TRNowry> @JsolomonReports If I had to bet, I'd say there's a 80% chance that millions of small balloons will cross into the US with mysteriously "empty" air canisters...a few weeks b
1622279949917851649 2023-02-05 17:04:15 +0000 <GrtVHecke> @larsroobol @ManuelSintubin Bij covid3 dagen dus moet dat nu ook kunnen niet? @TInneVdS  @egbertlachaert @JeremieVaneeckh @alexanderdecroo @SammyMahdi @GLBouchez @jmnolle
1622066667663523840 2023-02-05 02:56:44 +0000 <TRNowry> @gatewaypundit Is it a Chinese spy balloon or a PFIZER covid3.0 balloon, made in China?
1621995198640455680 2023-02-04 22:12:45 +0000 <TorukMakto2023> @Rep3effries Yes the fish in the Atlantic who have now been infected with covid3 are very thankful
1621994473524977664 2023-02-04 22:09:52 +0000 <TorukMakto2023> @TexasRepublic71 Great now all the fish have covid3
1621939446219161600 2023-02-04 18:31:12 +0000 <Red_Shark01> @bennyjohnson Covid3
1621912292072767495 2023-02-04 16:43:18 +0000 <BjuhrEva>  https://t.co/zRsVlT4xHi    Sars Covid3?
1621703857624465411 2023-02-04 02:55:04 +0000 <unbiasedhead111> @RealMattCouch Covid3 delivery!
1621680085787745484 9 2023-02-04 01:23:40 +0000 <Breannabanna80> @HoursCheckmate Haha I have read so many theories. The last of which was "it's release SARS-covid3 🤣. Lol
1621679907591028736 2023-02-04 01:19:53 +0000 <TorukMakto2023> @alx Will the covid3 bioweapon be slow released over multiple red states or just one quick dump?
1621655469264871426 2023-02-03 23:42:47 +0000 <JohnCallister14> @TexasLindsay_ Covid3
1620845635128918018 2023-02-01 18:04:47 +0000 <YomoF6Boys2> #GARLAND is a #DOJ #disease just like the #CURRENT #CORRUPT #FBI  #WHERES #PFIZER when ya need a new Covid strain to #CURE  #3Letter Agencies?  #Covid3.0  As House Repu
1620155149498081280 2023-01-30 20:21:03 +0000 <JamesOnekaka> @_MiIitia_ @TheConsultant18 Is Russia even attacking Ukraine?  Everything you see in media are false flag events as part of a nato funded civil war  When Russia speci
1618381028237901830 2023-01-25 22:51:19 +0000 <AndresVS94941115> @RosmerycastroM1 @MeDicenWelly Si yo llego a en la calle a una persona diciendo POLOPOLO te apoyamos, prometo ponerme a salvo lo más pronto posible, eso es peor que
1617828810199797760 2023-01-24 10:17:00 +0000 <Ghani_Du92> @tatiann69922625 Le covid3 et les masques c'est fini depuis très longtemps, essayer de convaincre  des convaincus, c'est déjà très con mais vouloir se donner de l'importa
1617289252571860992 2023-01-22 22:33:00 +0000 <MisterSamsaa> Im übrigen bin ich der Meinung, dass sich Gesundheitsminister Johannes Rauch ("dünne" Forschungslage über Long Covid³) umfassend über den aktuellen Stand der Wissensch
1616113529643487233 2023-01-19 16:41:06 +0000 <LuanBarboza1212> @faveladoinvest Não é essa aí que na pandemia mandava o povo ficar em casa, mas a empregada doméstica dela não deixou de trabalhar, e por ter que ir trabalhar acabo
1614598934885765122 2023-01-15 12:22:38 +0000 <VenusPeace22> @DrPaoloMezzana Lei la deve finire di dire stupidaggini.. Come mai tutti i vaccinati con 3 dosi che conosco si ammalano in continuazione e hanno fatto il covid3 volte
[!] No more data! Scraping will stop now.
found 0 deleted tweets in this search.
```

```
[ ] stand_df = pd.read_csv("covid3.csv")
    stand_df
```

```
[ ] stand_df = pd.read_csv("covid3.csv")
    stand_df
```

	id	conversation_id	created_at	date	time	timezone	user_id	username	name	place	tweet	language	mentions	urls	photos	replies_count	retweets_count	likes_count	hasht
0	1629681202188767235	1629680242301313024	2023-02-26 03:14:11 UTC	2023-02-26	03:14:11	0	323269450	giovann72	Covid3	NaN	@kanfootballclub Top =Rea flop Pantemis	en	[]	[]	[]	0	0	1	
1	1628907257990610944	1628562015290638337	2023-02-23 23:58:48 UTC	2023-02-23	23:58:48	0	323269450	giovann72	Covid3	NaN	@Aaron_Derfel Your still at with the Covid gue...	en	[]	[]	[]	0	0	0	
2	1628514281825136641	1628514281825136641	2023-02-22 21:57:15 UTC	2023-02-22	21:57:15	0	70880291	cmdarwindavid	Darwin David □	NaN	Covid³	en	[]	[]	[]	0	0	2	

```
[ ] Tweet = stand_df.filter(['geo','tweet','place','date'])
    Tweet
```

	geo	tweet	place	date
0	NaN	@kanfootballclub Top =Rea flop Pantemis	NaN	2023-02-26
1	NaN	@Aaron_Derfel Your still at with the Covid gue...	NaN	2023-02-23
2	NaN	Covid³	NaN	2023-02-22
3	NaN	@casillistaaa Rico si el Covid3.0	NaN	2023-02-22
4	NaN	@manofbird Looks like two donkeys too me so be...	NaN	2023-02-22
5	NaN	E assim nasce o CoVID3	NaN	2023-02-19
6	NaN	@BBNMedias @OneFootball Thanks 🙏	NaN	2023-02-19
7	NaN	@BBNMedias Is the game Vs rowdies on Apple TV ?	NaN	2023-02-19
8	NaN	@thehill Ukraine another excuse for the Govt t...	NaN	2023-02-18
9	NaN	@Ayjchan But it is a huge loss seeing that no ...	NaN	2023-02-17
10	NaN	@MissChief01 @MrMadMeta @satishshankar1 @kama...	NaN	2023-02-14
11	NaN	@KEEMSTAR @Cernovich An act of war that gets y...	NaN	2023-02-12
12	NaN	@jsolomonReports If I had to bet, I'd say ther...	NaN	2023-02-11
13	NaN	@larsroobol @ManuelSintubin Bij covid3 dagen d...	NaN	2023-02-05
14	NaN	@gatewaypundit Is it a Chinese spy balloon or ...	NaN	2023-02-05
15	NaN	@RepJeffries Yes the fish in the Atlantic who ...	NaN	2023-02-04
16	NaN	@TexasRepublic71 Great now all the fish have c...	NaN	2023-02-04
17	NaN	@bennyjohnson Covid3	NaN	2023-02-04
18	NaN	https://t.co/zRsVIT4xjN Sars Covid3?	NaN	2023-02-04
19	NaN	@RealMattCouch Covid3 deliveryl	NaN	2023-02-04
20	NaN	@HoursCheckmate Haha I have read so many theor...	NaN	2023-02-04
21	NaN	@alx Will the covid3 bioweapon be slow release...	NaN	2023-02-04
22	NaN	@TexasLindsay_ Covid3	NaN	2023-02-03

```
[ ] # create a function to clean the tweets
    def cleanTxt(text):
        #Remove @mentions and replace with blank
        text = re.sub(r'@[A-Za-z0-9]+', '', text)

        #Remove the '#' symbol, replace with blank
        text = re.sub(r'#', '', text)

        #Removing RT, replace with blank
        text = re.sub(r'RT[\s]+', '', text)

        #Remove the hyperlinks
        text = re.sub(r'https?:\/\/\S+', '', text)

        return text
    # clean the text
    Tweet['tweet']= Tweet['tweet'].apply(cleanTxt)
    #Show the clean text
    Tweet
```

	geo	tweet	place	date
0	NaN	Top =Rea flop Pantemis	NaN	2023-02-26
1	NaN	_Derfel Your still at with the Covid guess the...	NaN	2023-02-23
2	NaN	Covid³	NaN	2023-02-22
3	NaN	Rico si el Covid3.0	NaN	2023-02-22
4	NaN	Looks like two donkeys too me so beastality	NaN	2023-02-22
5	NaN	E assim nasce o CoVID3	NaN	2023-02-19
6	NaN	Thanks 🙏	NaN	2023-02-19
7	NaN	Is the game Vs rowdies on Apple TV ?	NaN	2023-02-19
8	NaN	Ukraine another excuse for the Govt to shake ...	NaN	2023-02-18
9	NaN	But it is a huge loss seeing that no agreemen...	NaN	2023-02-17
10	NaN	1 1 Neengala yawning panna mattinga la ungal...	NaN	2023-02-14
11	NaN	An act of war that gets you nuked is blowing...	NaN	2023-02-12
12	NaN	If I had to bet, I'd say there's a 80% chance...	NaN	2023-02-11
13	NaN	Bij covid3 dagen dus moet dat nu ook kunnen ...	NaN	2023-02-05
14	NaN	Is it a Chinese spy balloon or a PFIZER covid...	NaN	2023-02-05
15	NaN	Yes the fish in the Atlantic who have now bee...	NaN	2023-02-04

```
[ ] def remove_emoji(string):
        emoji_pattern = re.compile("["
                                   u"\U0001F600-\U0001F64F"  # emoticons
                                   u"\U0001F300-\U0001F5FF"  # symbols & pictographs
                                   u"\U0001F680-\U0001F6FF"  # transport & map symbols
                                   u"\U0001F1E0-\U0001F1FF"  # flags (iOS)
                                   u"\U00002500-\U00002BEF"  # chinese char
                                   u"\U00002702-\U00002B0"
                                   u"\U00002702-\U00002B0"
                                   u"\U000024C2-\U0001F251"
                                   u"\U0001f926-\U0001f937"
                                   u"\U00010000-\U0010ffff"
                                   u"\u2640-\u2642"
                                   u"\u2600-\u2B55"
                                   u"\u200d"
                                   u"\u23cf"
                                   u"\u23e9"
                                   u"\u231a"
                                   u"\ufe0f"  # dingbats
                                   u"\u3030"
                                   "]+", flags=re.UNICODE)
        return emoji_pattern.sub(r'', string)

    #Cleaning the text
    Tweet['tweet']= Tweet['tweet'].apply(remove_emoji)

    #Show the clean text
    Tweet
```

	geo	tweet	place	date
0	NaN	Top =Rea flop Pantemis	NaN	2023-02-26
1	NaN	_Derfel Your still at with the Covid guess the...	NaN	2023-02-23
2	NaN	Covid³	NaN	2023-02-22
3	NaN	Rico si el Covid3.0	NaN	2023-02-22
4	NaN	Looks like two donkeys too me so beastality	NaN	2023-02-22
5	NaN	E assim nasce o CoVID3	NaN	2023-02-19
6	NaN	Thanks	NaN	2023-02-19
7	NaN	Is the game Vs rowdies on Apple TV ?	NaN	2023-02-19

```
[ ] Tweet['date'] = pd.to_datetime(Tweet['date']).dt.date
```

```
[ ] Tweet['date'].unique()
```

```
array([datetime.date(2023, 2, 26), datetime.date(2023, 2, 23),
       datetime.date(2023, 2, 22), datetime.date(2023, 2, 19),
       datetime.date(2023, 2, 18), datetime.date(2023, 2, 17),
       datetime.date(2023, 2, 14), datetime.date(2023, 2, 12),
       datetime.date(2023, 2, 11), datetime.date(2023, 2, 5),
       datetime.date(2023, 2, 4), datetime.date(2023, 2, 3),
       datetime.date(2023, 2, 1), datetime.date(2023, 1, 30),
       datetime.date(2023, 1, 25), datetime.date(2023, 1, 24),
       datetime.date(2023, 1, 22), datetime.date(2023, 1, 19),
       datetime.date(2023, 1, 15)], dtype=object)
```

```
[ ]
```

➢ 2.1 GRU theory

1

20MIA1031
Sanjay.M

Hidden markov model (HMM)

→ HMM is a statistical model that can be used to describe the evolution of observable events that depend on internal factors which are not directly observable.

→ HMM are engineered to handle data which can be represented to sequence of observation over the time.

→ HMM are probabilistic frameworks where the observed data are modeled as a series of outputs generated by one of several hidden internal layers (or) states.

→ Hidden markov model are much simpler than Recurrent Neural Network. They rely on strong assumptions which may not always be true. If the assumptions are true then we may see better performance from HMM. it is less finicky to get working.

→ HMM cannot represent any dependency between the applications.

→ HMM do not explicitly capture the time in a specified state due to their markovian behaviour.

→ It doesn't consider any state sequence dominating any given state because of markovian nature.

Recurrent neural network (RNN)

→ Recurrent neural network is a type of artificial neural network which uses sequential data or time series data.

→ This comes under a deep learning technique, where as this algorithm is commonly used for ordinal or temporal problems.

ex:- natural language processing
language translation
speed recognition.

→ Generating image description
→ face detection.

like example:- Siri, google translator.

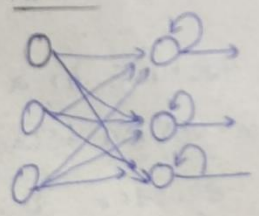
→ It is like CNN → convolutional neural network & RNN uses training data to learn.

→ They are distinguished from prior inputs to influence the current input and output.

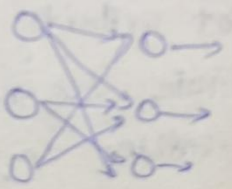
→ RNN depend on the prior elements within the sequence.

→ unidirectional recurrent neural networks cannot account for these events in their predictions.

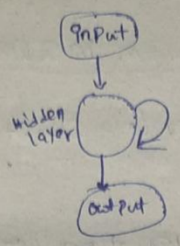
RNN



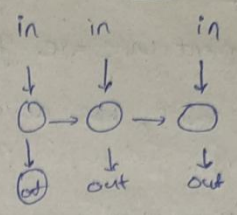
feedforward NN



Rolled RNN



Unrolled RNN

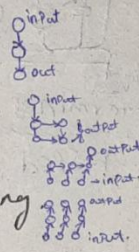


→ The weights that are adjusted through the process of backpropagation & gradient descent to facilitate Reinforcement Learning.

→ RNN leverage backpropagation through time (BPTT) algorithm to determine the gradient, which is slightly different from traditional backpropagation specifically for sequence data.

→ BPTT calculates error for input & output.

Types of RNN



classification of image. → one to one

image to sentence → one to many

semantic analysis (+ve/-ve) feelings → many to one

English to french → many to many

Applications

→ Prediction problems

→ machine translation

→ Text Summarization

→ Language modelling & generating text.

→ we use sigmoid and Relu as activation functions.

$$g(x) = \frac{1}{1+e^{-x}} \qquad g(x) = \max(0, x)$$

Variant RNN architectures.

→ Bidirectional recurrent neural networks

→ Long short term memory (LSTM)

→ Gated recurrent units.

Gated recurrent unit.

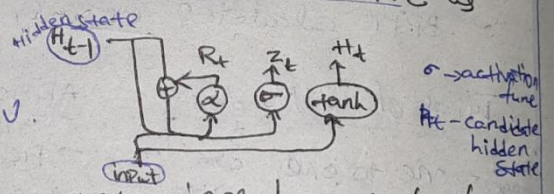
→ Gated recurrent units are gating mechanism in RNN

→ In certain cases, has advantages over long short term memory.

→ GRU is using less memory and faster than LSTM, LSTM is more accurate when using datasets with longer sequences.

→ for solving the vanishing exploding gradient problem often encountered during the operation of basic RNN many variations were developed.

→ GRU is less know but more and effective as equal to LSTM



Types of gates in GRU.

Update gate :- (Z_t) It determines how much of the past knowledge needs to be passed along into the future.

Reset gate (R_t) :- It determines how much of the past knowledge to forget. Combination of input gate and forget gate.

Current memory gate ($\tilde{h}_t$) :- It is often overlooked during a typical discussion on GRU.

→ It incorporates into the Reset gate.

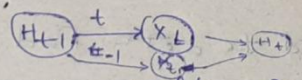
(GRU)
→ workflow is similar to the basis RNN.

→ The vanishing gradient problem occurs in machine Learning when the gradient becomes vanishingly Small, It prevents weight from changing its value.

→ applications are Hand writing detection, speech, Polyphonic music modelling.

→ They have better performance than LSTM when dealing with small datasets.

→ working process of GRU It takes time Stamp "t" at each time. It take an input "X_t" and hidden state "H_{t-1}" from before time Stamp "(t-1)". It gives a new hidden state "H_t" and process continues.



→ Using Electronic Health Records to predict future diagnosis codes with GRU.

→ Generate artificial EHR data.

→ preprocess the data.

→ basic EdA.

→ create Admission table

→ create Patient diagonisis table.

→ create a hashkey for Admission ID (unique)

-> creating mapping of Patient ID's
-> map it accordingly based on appointments, diseases etc..
-> creating Diagonosis mapped to each unique Patient and visit.
-> embed codes into visit of Patient - Admission mapping.
-> extract Patient ID's visit dates and diagnosis.
-> creating a unique diagnosis code for each visit for each unique Patient.
-> Bump data into the pickled list
-> AI model architecture to find the network to doctor.
-> model is trained on GPU enabled system.
-> Now load the data, Pad the inputs.
-> customing Layers of handling two layer GRU
-> Training model
-> find the node and go forward to next step.

Gradient Descent.

Gradient descent is an optimization algoritum for finding a local minimum of a differentiable function. Gradient descent in machine learning is simply used to find the values of a function's parameters. that minimize a cost function as far as possible. Gradient descent is an Iterative first-order optimization algoritum used to find a local minimum or maximum of a given function. This method is commonly used in machine learning and deep learning for minimizing the cost.

we have 3 types in it.

Batch gradient Descent

Stochastic gradient Descent

mini - Batch gradient Descent.

Out this 3 types mini Batch gradient Descent works faster than both gradient descent.

➢ 2.2 GRU EXECUTION

20MIA1031

SANJAY.M

Using Electronic Health Records to predict future diagnosis codes with Gated Recurrent Units

Electronic medical records(EMRs), which is sometimes interchangeably called Electronic health records(EHRs) are primarily used to electronically-store patient health data digitally.

```
import pandas as pd
import numpy as np

admission_table = {'Patient 1': {'PatientID':'A1234-B456',
                                 'Admission ID':[12,34,15],
                                 'AdmissionStartDate':['2019-01-03 9:34:55','2019-02-03 10:50:55','2019-04-03 12:34:55'],
                                 'AdmissionEndDate':['2019-01-07 8:45:43','2019-03-04 1:50:32','2019-04-03 5:38:18']},
                   'Patient 2': {'PatientID':'B1234-C456',
                                 'Admission ID':[13,34],
                                 'AdmissionStartDate':['2018-01-03 9:34:55','2018-02-03 10:50:55'],
                                 'AdmissionEndDate':['2018-01-07 8:45:43','2018-03-04 1:50:32']}}
admission_table = (pd.concat({k: pd.DataFrame(v) for k, v in admission_table.items()}).reset_index(level=1, drop=True))
admission_table = admission_table.reset_index(drop=True)
```

```
Patient_1 = {'PatientID':'A1234-B456',
             'Admission ID':[12,34,15],
             'PrimaryDiagnosisCode':[['E11.64','I25.812','I25.10'],
                                     ['E11.64','I25.812','I25.10','780.96','784.0'],
                                     ['E11.64','I25.812','I25.10','786.50','401.9','789.00']],
             'CodingSystem':['ICD-9','ICD-9','ICD-9'],
             'DiagnosisCodeDescription':[['Type 2 diabetes mellitus with hypoglycemia',
                                          'Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
                                          'Atherosclerotic heart disease of native coronary artery without angina pectoris'],
                                         ['Type 2 diabetes mellitus with hypoglycemia',
                                          'Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
                                          'Atherosclerotic heart disease of native coronary artery without angina pectoris',
                                          'Generalized Pain', 'Dizziness and giddiness'],
                                         ['Type 2 diabetes mellitus with hypoglycemia',
                                          'Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
                                          'Atherosclerotic heart disease of native coronary artery without angina pectoris',
                                          'Chest pain, unspecified','Essential hypertension, unspecified',
                                          'Abdominal pain, unspecified site']]}
Patient_2 = {'PatientID':'B1234-C456',
             'Admission ID':[13,34],
             'PrimaryDiagnosisCode':[['M05.59','Z13.85','O99.35'],['M05.59','Z13.85','O99.35','D37.0']],
             'CodingSystem':['ICD-9','ICD-9'],
             'DiagnosisCodeDescription':[['Rheumatoid polyneuropathy with rheumatoid arthritis of multiple sites',
                                          'Encounter for screening for nervous system disorders',
                                          'Diseases of the nervous system complicating pregnancy, childbirth, and the puerperium'],
                                         ['Rheumatoid polyneuropathy with rheumatoid arthritis of multiple sites',
                                          'Encounter for screening for nervous system disorders',
                                          'Diseases of the nervous system complicating pregnancy, childbirth, and the puerperium',
                                          'Neoplasm of uncertain behavior of lip, oral cavity and pharynx']]}
```

```
def process_ehr(Patient1,Patient2):
    pt_diagnosis_table = [Patient1,Patient2]
    pt_diagnosis_table = pd.concat([pd.DataFrame({k:v for k,v in d.items()}) for d in pt_diagnosis_table])

    pt_diagnosis_table = (pt_diagnosis_table.set_index(['PatientID', 'Admission ID','CodingSystem'])
            .apply(lambda x: x.apply(pd.Series).stack())
            .reset_index()
            .drop('level_3', 1))
    return pt_diagnosis_table
def hash_key(df):
    df['HashKey'] = df['PatientID'].\
    apply(lambda x: x.split('-')[0]) + '-' + df['Admission ID'].astype('str')
    cols = [df.columns[-1]] + [col for col in df if col != df.columns[-1]]
    print(cols)
    return df[cols]
diagnosis_table = process_ehr(Patient_1,Patient_2)
diagnosis_table.head()
```

```
<ipython-input-3-3e4623b65dac>:5: FutureWarning: In a future version of pandas all arguments of DataFrame.drop except for the argument 'labels' will be keyword-only.
  pt_diagnosis_table = (pt_diagnosis_table.set_index(['PatientID', 'Admission ID','CodingSystem'])
```

	PatientID	Admission ID	CodingSystem	PrimaryDiagnosisCode	DiagnosisCodeDescription
0	A1234-B456	12	ICD-9	E11.64	Type 2 diabetes mellitus with hypoglycemia
1	A1234-B456	12	ICD-9	I25.812	Atherosclerosis of bypass graft of coronary ar...
2	A1234-B456	12	ICD-9	I25.10	Atherosclerotic heart disease of native corona...
3	A1234-B456	34	ICD-9	E11.64	Type 2 diabetes mellitus with hypoglycemia
4	A1234-B456	34	ICD-9	I25.812	Atherosclerosis of bypass graft of coronary ar...

```
diagnosis_table[diagnosis_table['Admission ID']==34]
```

	PatientID	Admission ID	CodingSystem	PrimaryDiagnosisCode	DiagnosisCodeDescription
3	A1234-B456	34	ICD-9	E11.64	Type 2 diabetes mellitus with hypoglycemia
4	A1234-B456	34	ICD-9	I25.812	Atherosclerosis of bypass graft of coronary ar...
5	A1234-B456	34	ICD-9	I25.10	Atherosclerotic heart disease of native corona...
6	A1234-B456	34	ICD-9	780.96	Generalized Pain
7	A1234-B456	34	ICD-9	784.0	Dizziness and giddiness
17	B1234-C456	34	ICD-9	M05.59	Rheumatoid polyneuropathy with rheumatoid arth...
18	B1234-C456	34	ICD-9	Z13.85	Encounter for screening for nervous system dis...
19	B1234-C456	34	ICD-9	O99.35	Diseases of the nervous system complicating pr...
20	B1234-C456	34	ICD-9	D37.0	Neoplasm of uncertain behavior of lip, oral ca...

```
admission_table.head()
```

	PatientID	Admission ID	AdmissionStartDate	AdmissionEndDate
0	A1234-B456	12	2019-01-03 9:34:55	2019-01-07 8:45:43
1	A1234-B456	34	2019-02-03 10:50:55	2019-03-04 1:50:32
2	A1234-B456	15	2019-04-03 12:34:55	2019-04-03 5:38:18
3	B1234-C456	13	2018-01-03 9:34:55	2018-01-07 8:45:43
4	B1234-C456	34	2018-02-03 10:50:55	2018-03-04 1:50:32

```
admission_table[['AdmissionStartDate','AdmissionEndDate']] = admission_table[['AdmissionStartDate','AdmissionEndDate']]\
                .apply(lambda x: pd.to_datetime(x, format='%Y-%m-%d'))
admission_table.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5 entries, 0 to 4
Data columns (total 4 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   PatientID           5 non-null      object
 1   Admission ID        5 non-null      int64
 2   AdmissionStartDate  5 non-null      datetime64[ns]
 3   AdmissionEndDate    5 non-null      datetime64[ns]
dtypes: datetime64[ns](2), int64(1), object(1)
memory usage: 288.0+ bytes
```

```
diagnosis_table.head()
```

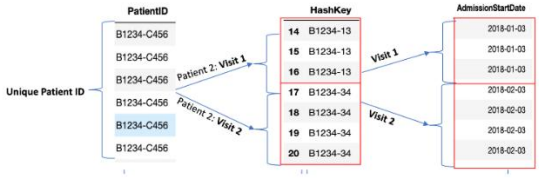
	PatientID	Admission ID	CodingSystem	PrimaryDiagnosisCode	DiagnosisCodeDescription
0	A1234-B456	12	ICD-9	E11.64	Type 2 diabetes mellitus with hypoglycemia
1	A1234-B456	12	ICD-9	I25.812	Atherosclerosis of bypass graft of coronary ar...
2	A1234-B456	12	ICD-9	I25.10	Atherosclerotic heart disease of native corona...
3	A1234-B456	34	ICD-9	E11.64	Type 2 diabetes mellitus with hypoglycemia
4	A1234-B456	34	ICD-9	I25.812	Atherosclerosis of bypass graft of coronary ar...

```
diagnosis_table.to_csv('/content/Diagnosis_Table.csv',encoding='UTF-8',index=False)
admission_table.to_csv('/content/Admissions_Table.csv',encoding='UTF-8',index=False,date_format='%Y-%m-%d')
```

```
import pandas as pd
import numpy as np
import pandas as pd
from time import time
import matplotlib.pyplot as plt
import seaborn as sns
import sys
import warnings
from datetime import datetime
import torch
import pickle
from collections import defaultdict
warnings.filterwarnings('ignore')
sns.set(style='white')
%autosave 180
```

```
Autosaving every 180 seconds
```

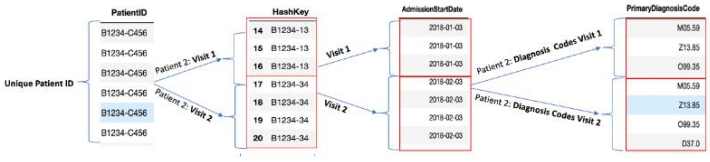
Step 1: Create mapping of patient IDs




```
print('Creating visit date mapping')
patHashMap = dict(defaultdict(list))  # this creates a dictionary with a list of values for each patient:[number of visists]
visitMap = dict(defaultdict()) # this creates a dictionary with a mapping of the patientID : visitdates

data = open('/content/Admissions_Table.csv','r')
data.readline()[1:] # read every line except the file header

for line in data:
    feature = line.strip().split(',') # split line on , and isolate columns
    visitDateID = datetime.strptime(feature[3],'%Y-%m-%d')
    patHashMap.setdefault(feature[1], []).append(feature[0]) # create a mapping for each visit for a specific PatientID
    visitMap.setdefault(feature[0], []).append(visitDateID) # create a mapping for each visit for a specific Admission Date

#Patient ID- visit mapping
patHashMap
```

```
Creating visit date mapping
{'12': ['A1234-B456'],
 '34': ['A1234-B456', 'B1234-C456'],
 '15': ['A1234-B456'],
 '13': ['B1234-C456']}
```

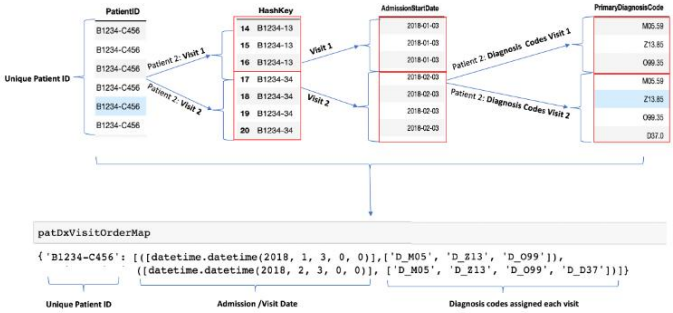
Step 2: Create Diagnosis Code Mapped to each unique patient and visit




```
[ ] # Patient Admission ID- visit date mapping
    visitMap
```

```
{'A1234-B456': [datetime.datetime(2019, 1, 7, 0, 0),
  datetime.datetime(2019, 3, 4, 0, 0),
  datetime.datetime(2019, 4, 3, 0, 0)],
 'B1234-C456': [datetime.datetime(2018, 1, 7, 0, 0),
  datetime.datetime(2018, 3, 4, 0, 0)]}
```

Step 3: Embed diagnosis codes into visit mapping Patient-Admission mapping



```
[ ] print('Creating Diagnosis-Visit mapping')
    visitDxMap = dict(defaultdict(list))

    data = open('/content/Diagnosis_Table.csv', 'r')
    data.readline()[1:]

    for line in data:
        feature = line.strip().split(',')
        visitDxMap.setdefault(feature[0], []).append('D_' + feature[4].split('.')[0]) # add a unique identifier before the

    visitDxMap # Mapping of each Admission ID to each diagnosis code assigned during that visit
```

```
Creating Diagnosis-Visit mapping
{'A1234-B456': ['D_Type 2 diabetes mellitus with hypoglycemia',
  'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
  'D_Atherosclerotic heart disease of native coronary artery without angina pectoris',
  'D_Type 2 diabetes mellitus with hypoglycemia',
  'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
  'D_Atherosclerotic heart disease of native coronary artery without angina pectoris',
  'D_Generalized Pain',
  'D_Dizziness and giddiness',
  'D_Type 2 diabetes mellitus with hypoglycemia',
  'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
  'D_Atherosclerotic heart disease of native coronary artery without angina pectoris',
  'D_"Chest pain',
  'D_"Essential hypertension',
  'D_"Abdominal pain'],
 'B1234-C456': ['D_Rheumatoid polyneuropathy with rheumatoid arthritis of multiple sites',
  'D_Encounter for screening for nervous system disorders',
  'D_"Diseases of the nervous system complicating pregnancy',
  'D_Rheumatoid polyneuropathy with rheumatoid arthritis of multiple sites',
  'D_Encounter for screening for nervous system disorders',
  'D_"Diseases of the nervous system complicating pregnancy',
  'D_"Neoplasm of uncertain behavior of lip']}
```

```
print("Sorting visit mapping")
patDxVisitOrderMap = {}
for patid, visitDates in patHashMap.items():
    sorted_list = ([(visitMap[visitDateID], visitDxMap[visitDateID]) for visitDateID in visitDates])
    patDxVisitOrderMap[patid] = sorted_list

patDxVisitOrderMap
```

```
Sorting visit mapping
{'12': [([datetime.datetime(2019, 1, 7, 0, 0),
    datetime.datetime(2019, 3, 4, 0, 0),
    datetime.datetime(2019, 4, 3, 0, 0)],
   ['D_Type 2 diabetes mellitus with hypoglycemia',
    'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
```

```
[ ] print("Extracting patient IDs, visit dates and diagnosis codes into individual lists for encoding")
    patIDs = [patid for patid, visitDate in patDxVisitOrderMap.items()]
    datesList = [[visit[0][0] for visit in visitDate] for patid, visitDate in patDxVisitOrderMap.items()]
    DxsCodesList = [[visit[1] for visit in visitDate] for patid, visitDate in patDxVisitOrderMap.items()]

    patIDs
```

```
Extracting patient IDs, visit dates and diagnosis codes into individual lists for encoding
['12', '34', '15', '13']
```

```
[ ] datesList
```

```
[[datetime.datetime(2019, 1, 7, 0, 0)],
 [datetime.datetime(2019, 1, 7, 0, 0), datetime.datetime(2018, 1, 7, 0, 0)],
 [datetime.datetime(2019, 1, 7, 0, 0)],
 [datetime.datetime(2018, 1, 7, 0, 0)]]
```

```
DxsCodesList
```

```
[[['D_Type 2 diabetes mellitus with hypoglycemia',
   'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
   'D_Atherosclerotic heart disease of native coronary artery without angina pectoris',
   'D_Type 2 diabetes mellitus with hypoglycemia',
   'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
   'D_Atherosclerotic heart disease of native coronary artery without angina pectoris',
   'D_Generalized Pain',
   'D_Dizziness and giddiness',
   'D_Type 2 diabetes mellitus with hypoglycemia',
   'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
   'D_Atherosclerotic heart disease of native coronary artery without angina pectoris',
   'D_"Chest pain',
   'D_"Essential hypertension',
   'D_"Abdominal pain']],
 [['D_Type 2 diabetes mellitus with hypoglycemia',
   'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
   'D_Atherosclerotic heart disease of native coronary artery without angina pectoris',
   'D_Type 2 diabetes mellitus with hypoglycemia',
   'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
   'D_Atherosclerotic heart disease of native coronary artery without angina pectoris',
   'D_Generalized Pain',
   'D_Dizziness and giddiness',
   'D_Type 2 diabetes mellitus with hypoglycemia',
   'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris',
   'D_Atherosclerotic heart disease of native coronary artery without angina pectoris',
   'D_"Chest pain',
   'D_"Essential hypertension',
```

```python
# ('Encoding string Dx codes to integers and mapping the encoded integer value to the ICD-10 code for interpretation')
DxCodeDictionary = {}
encodedDxs = []
for patient in DxsCodesList:
    encodedPatientDxs = []
    for visit in patient:
        encodedVisit = []
        for code in visit:
            if code in DxCodeDictionary:
                encodedVisit.append(DxCodeDictionary[code])
            else:
                DxCodeDictionary[code] = len(DxCodeDictionary)
                encodedVisit.append(DxCodeDictionary[code])
        encodedPatientDxs.append(encodedVisit)
    encodedDxs.append(encodedPatientDxs)

DxCodeDictionary # Dictionary of all unique codes in the entire dataset aka: Our Code Vocabulary
```

```
{'D_Type 2 diabetes mellitus with hypoglycemia': 0,
 'D_Atherosclerosis of bypass graft of coronary artery of transplanted heart without angina pectoris': 1,
 'D_Atherosclerotic heart disease of native coronary artery without angina pectoris': 2,
 'D_Generalized Pain': 3,
 'D_Dizziness and giddiness': 4,
 'D_"Chest pain': 5,
 'D_"Essential hypertension': 6,
 'D_"Abdominal pain': 7,
 'D_Rheumatoid polyneuropathy with rheumatoid arthritis of multiple sites': 8,
 'D_Encounter for screening for nervous system disorders': 9,
 'D_"Diseases of the nervous system complicating pregnancy': 10,
 'D_"Neoplasm of uncertain behavior of lip': 11}
```

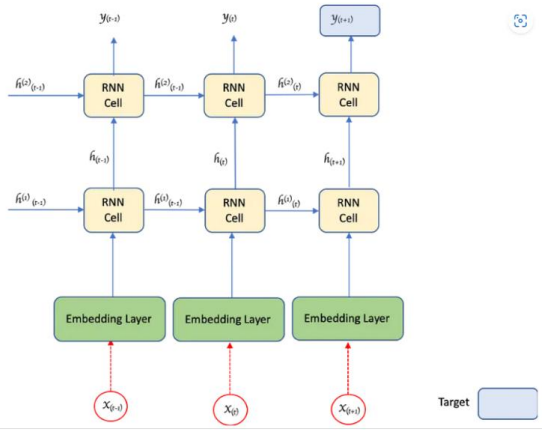
```
[ ]
encodedDxs # Converted list of list with integer converted diagnosis codes
```

```
[[[0, 1, 2, 0, 1, 2, 3, 4, 0, 1, 2, 5, 6, 7]],
 [[0, 1, 2, 0, 1, 2, 3, 4, 0, 1, 2, 5, 6, 7], [8, 9, 10, 8, 9, 10, 11]],
 [[0, 1, 2, 0, 1, 2, 3, 4, 0, 1, 2, 5, 6, 7]],
 [[8, 9, 10, 8, 9, 10, 11]]]
```

```
[ ] outFile = 'ArtificialEHR_Data'
    print('Dumping files into a pickled list')
    pickle.dump(patIDs, open(outFile+'.patIDs', 'wb'),-1)
    pickle.dump(datesList, open(outFile+'.dates', 'wb'),-1)
    pickle.dump(encodedDxs, open(outFile+'.encodedDxs', 'wb'),-1)
    pickle.dump(DxCodeDictionary, open(outFile+'.Dxdictionary', 'wb'),-1)
```

```
Dumping files into a pickled list
```

Model Architecture



➢ 3.1 paper1

Sensors for continous monitoring of surgeons cognitive workland in the cardiac operating room

paper-1

{20MIA1031
Sanjay.m}

Sensors for continous monitoring of surgeon's coggnitive workload in the cardiac operating room.

→ This study looked at the potential of collecting heart rate and near infrared spectroscopy data during aortic value surgery to measure healthcare worker's coggnitive workload.

→ The findings revealted a link between gathered signals and change in cardiovascular and brain activity during the surgical phases.

→ This paper showing the potential of improving patient safety and team performance. future study is needed, however to better understanding the link between heart rate and regional oxygen saturation, as well as to employ higher quality research-grade wireless sensors.

→ The study looked at the potential of collecting data on heart rate and near infrared spectroscopy during aortic value surgery to measure healthcare worker's cognitive workload.

→ The findings revealed a link between the signals gathered and changes in cardiovascular and brain activity during the surgical phases.

→ The study shows the gathering sensor data in a real-world operating room situation is feasible.

→ While heart rate variability and near-infrared spectroscopy are rountinely employed to quantify cognitive effort, their implementation in a real-world environment poses obstacles.

→ Data was collected during a surgical operation using a wireless heart rate sensor and a two channel cerebral/somatic oximeter.

→ The findings revealed that both heart rate and near infrared spectroscopy data were regularly distributed and correlated significantly.

→ The paper leads credence to the potential of collecting multimodal measurements of cognitive burden during surgery in order to improve patient safety and team performance.

→ However, further study is needed to better understanding in link between heart rate and regional oxygen saturation, as well as to assess the modalities using higher-quality sensors.

→ The Study looked at the Potential of collecting multi-media Psycho-physicological data during aortic valve surgery in order to measure healthcare worker's effort.

→ They collected data during operation using wireless heart rate sensor and an NIRS device and discovered a substantial association between heart rate and regional saturation of oxygen readings.

→ They need better understanding on the relationship between heartrate and rSO_2 as well as to use higher quality research grade and wireless sensors.

→ They find the gathered signals were substantially connected, intra-operative events, self reports corresponding with change in cardiovascular and brain activity during various surgical stages.

➢ 3.2 Paper2

Intelligent iCU for autonomous patient monitoring using pervasive sensing and deep learning

Paper-2

Intelligent ICU for Autonomous Patient monitoring using Pervasive Sensing and deep learning.

→ This study assess the existence of delirium in ICU patients using AI and ubiquitous sensing technologies.

→ wearable accelerometer sensors, light and sound sensors, camera (face detection), posture, extremities movement, sound & light.

→ computer vision and deep learning approaches were used to evalute data, including face Action units (AU) for facial emotion detection & 2D pose estimate model for posture categorization.

→ The findings demonstrated that the AI system may be used to monitor ICU patients and detect using change in functional status, sleep, quality & environment.

→ In the ICU, wearable accelerometer sensors, light & sound sensors, cameral were utilised to collect information about patients and their surroundings.

→ The data was analysed using computer vision and deep learning approches. (AU) and 2D pose estimate.

→ The algorithm accuracy was assessed using expert-annotated ground truth, with a positive recognition recorded if the patient was identified with a probability of 0.9 or higher.

→ The results demonstrate that the AI system may be used to monitor ICU patients and detect variations in functional status, sleep, quality, environmental exposes between delirious and non delirious patients.

→ The model employed in the study was 2D posture estimate model with part affinity field. To classify posture, this model was trained on a balanced dataset of ICU patient data and scripted data.

→ further, facial emotion identification was done during the input using openface toolkit, which used face action units to recognise 8 common expressions such as pain, pleasure, sorrow, rage, fear, disgust, contempt.

→ The data acquired from the wearable accelerometer sensors was analysed using statistical analysis as well as python tensorflow, openCV, R.

→ The study found that delirious patients had fewer visits during the date but more interruptions at night.

→ As well as greater sound pressures levels & light intensity.

→ The patient who were delirious reported a reduced overall ability to fall asleep and were more likely to find the light distraction.

→ Delirious individuals had a greater heart rate, oxygen saturation, breathing rate, low blood pressure.

→ This paper revealed that the technology potential for decribing critical skills individuals and their environment in the ICU.

Github link my files

https://github.com/Sanjay3369/Predictive/tree/DA

https://colab.research.google.com/drive/1d47zI8GmZ7FD76YteuaF_c_16g5S4Oig
https://colab.research.google.com/drive/10kni8eg4O6-obu-txtAtbcYouP9ulB10#scrollTo=NULObGy2Urva
https://colab.research.google.com/drive/198l9-2NQ0UH3MQyAsoV9mJ48Q0Dw8cXM

Video link

https://drive.google.com/drive/folders/1q9X_G0W6wTKRSP7qlbW-I46XhU38caSi?usp=sharing