

```
!pip install flair
```

```

Downloading sacremoses-0.0.46-py3-none-any.whl (895 kB)
|████████████████████████████████████████| 895 kB 61.3 MB/s
Collecting tokenizers<0.11,>=0.10.1
  Downloading tokenizers-0.10.3-cp37-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.man
  |████████████████████████████████████████| 3.3 MB 33.0 MB/s
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.7/dist-packa
Requirement already satisfied: wcwidth in /usr/local/lib/python3.7/dist-packages (fro
Requirement already satisfied: PySocks!=1.5.7,>=1.5.6 in /usr/local/lib/python3.7/dis
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages (from
Building wheels for collected packages: gdown, mpld3, overrides, sqlitedict, ftfy, la
Building wheel for gdown (PEP 517) ... done
Created wheel for gdown: filename=gdown-3.12.2-py3-none-any.whl size=9704 sha256=22
Stored in directory: /root/.cache/pip/wheels/ba/e0/7e/726e872a53f7358b4b96a9975b04e
Building wheel for mpld3 (setup.py) ... done
Created wheel for mpld3: filename=mpld3-0.3-py3-none-any.whl size=116702 sha256=6fa
Stored in directory: /root/.cache/pip/wheels/26/70/6a/1c79e59951a41b4045497da187b27
Building wheel for overrides (setup.py) ... done
Created wheel for overrides: filename=overrides-3.1.0-py3-none-any.whl size=10186 s
Stored in directory: /root/.cache/pip/wheels/3a/0d/38/01a9bc6e20dcfaf0a6a7b552d0313
Building wheel for sqlitedict (setup.py) ... done
Created wheel for sqlitedict: filename=sqlitedict-1.7.0-py3-none-any.whl size=14392
Stored in directory: /root/.cache/pip/wheels/af/94/06/18c0e83e9e227da8f3582810b51f3
Building wheel for ftfy (setup.py) ... done
Created wheel for ftfy: filename=ftfy-6.0.3-py3-none-any.whl size=41933 sha256=909b
Stored in directory: /root/.cache/pip/wheels/19/f5/38/273eb3b5e76dfd850619312f69371
Building wheel for langdetect (setup.py) ... done
Created wheel for langdetect: filename=langdetect-1.0.9-py3-none-any.whl size=99324
Stored in directory: /root/.cache/pip/wheels/c5/96/8a/f90c59ed25d75e50a8c10a1b1c2d4
Building wheel for wikipedia-api (setup.py) ... done
Created wheel for wikipedia-api: filename=Wikipedia_API-0.5.4-py3-none-any.whl size
Stored in directory: /root/.cache/pip/wheels/d3/24/56/58ba93cf78be162451144e7a98896
Successfully built gdown mpld3 overrides sqlitedict ftfy langdetect wikipedia-api
Installing collected packages: requests, pyyaml, importlib-metadata, tokenizers, sent
Attempting uninstall: requests
  Found existing installation: requests 2.23.0
  Uninstalling requests-2.23.0:
    Successfully uninstalled requests-2.23.0
Attempting uninstall: pyyaml
  Found existing installation: PyYAML 3.13
  Uninstalling PyYAML-3.13:
    Successfully uninstalled PyYAML-3.13
Attempting uninstall: importlib-metadata
  Found existing installation: importlib-metadata 4.8.2
  Uninstalling importlib-metadata-4.8.2:
    Successfully uninstalled importlib-metadata-4.8.2
Attempting uninstall: more-itertools
  Found existing installation: more-itertools 8.12.0
  Uninstalling more-itertools-8.12.0:
    Successfully uninstalled more-itertools-8.12.0
Attempting uninstall: gdown
  Found existing installation: gdown 3.6.4
  Uninstalling gdown-3.6.4:
    Successfully uninstalled gdown-3.6.4

```

ERROR: pip's dependency resolver does not currently take into account all the package requirements that were pinned to 6.11.0. It is recommended to remove pip's cache and re-run pip. To do this, you can try the following commands: pip cache purge

markdown 3.3.6 requires importlib-metadata>=4.4; python_version < "3.10", but you have google-colab 1.0.0 requires requests~2.23.0, but you have requests 2.26.0 which is incompatible. datascience 0.10.6 requires folium==0.2.1, but you have folium 0.8.3 which is incompatible.

```
from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

```
from flair.data import Sentence
from flair.models import SequenceTagger
```

```
# make some sentences:
```

```
sentences_med = [
    'Satya Nadella is the CEO of Microsoft',
    'Anterior Lumbar Fusion Retro Peritoneal L-5 S-1',
    'Rigid bronchoscopy, debridement, T-tube placement.',
    'Rubella immune, hepatitis B surface antigen negative and Group B strep unknown.',
    'The pregnancy was complicated by fetal gastroschisis.',
    'The patient\'s Lasix was stopped and the patient worked with physical therapy.',
    'Back: No vertebral tenderness.',
    'No palpable masses on exam.',
    'Bring me your huddled masses',
    'I took Tylenol because my head was throbbing in pain',
    'Take asparine if you feel like it',
    'Take asparine 40mg twice a day',
    'Take Tylenol 40mg twice a day, or asparine if you feel like it',
    'Take Tylenol 40mg or asparine twice a day',
    'CURRENTLY INJECTING 2 UNITS LANTUS OD AND 8 UNITS NOVORAPID tds Cold intolerance No hyp',
    'T2 Diabeets 12 years LVF AKI',
    'T2 Diabetes25 years humulin 70/30 20,20',
    'T2 Diabetes Humulin 70/30 24 bd',
    'lung biopsy ?wegners granulomatosis due to start steroid',
    'Loss appetite nausea, Polyurea loss of conciuos ness and dizziness',
]
```

```
sentences = [
    'I love Berlin .',
    'The Effective Date is 4-8-2018',
    'The process will start in Months 3 - 5 of every yer'
]
```

```
# load the NER tagger (medical or standard)
```

```
tagger = SequenceTagger.load('/content/drive/MyDrive/Colab Notebooks/Data Sets/best-model.pt')
# tagger = SequenceTagger.load('ner-ontonotes')
```

```
# run NER over sentences
```

```
for sent in sentences_med:
```

```
    tagged_sentence = Sentence(sent)
    tagger.predict(tagged_sentence)
```

```
    print(tagged_sentence)
```

```
    entities_found = tagged_sentence.get_spans('ner')
```

```
    if entities_found:
        print(' Found entities:')
```

```
        # iterate over entities and print
        for entity in entities_found:
            print(' ' + str(entity))
```

```
    else:
        print(' (No entities found)')
```

```
print('****')
```

```
Sentence: "The pregnancy was complicated by fetal gastroschisis ." [- Tokens: 8 -
Found entities:
```

```
    Span [6,7]: "fetal gastroschisis" [- Labels: Disease (0.768)]
```

```
****
```

```
Sentence: "The patient 's Lasix was stopped and the patient worked with physical ther
(No entities found)
```

```
****
```

```
Sentence: "Back : No vertebral tenderness ." [- Tokens: 6 - Token-Labels: "Back :
Found entities:
```

```
    Span [4,5]: "vertebral tenderness" [- Labels: Disease (0.8226)]
```

```
****
```

```
Sentence: "No palpable masses on exam ." [- Tokens: 6]
(No entities found)
```

```
****
```

```
Sentence: "Bring me your huddled masses" [- Tokens: 5]
(No entities found)
```

```
****
```

```
Sentence: "I took Tylinole because my head was throbbing in pain" [- Tokens: 10 -
Found entities:
```

```
    Span [10]: "pain" [- Labels: Disease (0.9277)]
```

```
****
```

```
Sentence: "Take asparine if you feel like it" [- Tokens: 7]
(No entities found)
```

```
****
```

```
Sentence: "Take asparine 40mg twice a day" [- Tokens: 6]
(No entities found)
```

```
****
```

```
Sentence: "Take Tylenol 40mg twice a day , or asparine if you feel like it" [- Tole
(No entities found)
```

```
****
```

```
Sentence: "Take Tylenol 40mg or asparine twice a day" [- Tokens: 8]
(No entities found)
```

```
****
```

```
Sentence: "CURRENTLY INFECTING 3 UNITS LANTUS OD AND 2 UNITS NOVORAPID + 10 UNITS GLARGINE
```

```

Sentence: CURRENTLY INJECTING 2 UNITS LANTUS OD AND 8 UNITS NOVORAPID TDS COLD INTO
(No entities found)
****
Sentence: "T2 Diabeets 12 years LVF AKI" [- Tokens: 6 - Token-Labels: "T2 Diabeets
Found entities:
Span [6]: "AKI" [- Labels: Disease (0.7772)]
****
Sentence: "T2 Diabetes25 years humulin 70 / 30 20,20" [- Tokens: 8]
(No entities found)
****
Sentence: "T2 Diabetes Humulin 70 / 30 24 bd" [- Tokens: 8 - Token-Labels: "T2 Dia
Found entities:
Span [2]: "Diabetes" [- Labels: Disease (0.3926)]
****
Sentence: "lung biopsy ? wegners granulomatosis due to start steroid" [- Tokens: 9
Found entities:
Span [4,5]: "wegners granulomatosis" [- Labels: Disease (0.8675)]
****
Sentence: "Loss apetite nausea , Polyurea loss of conciuos ness and dizziness" [- T
Found entities:
Span [1,2,3]: "Loss apetite nausea" [- Labels: Disease (0.614)]
Span [5,6,7,8,9]: "Polyurea loss of conciuos ness" [- Labels: Disease (0.72)]
Span [11]: "dizziness" [- Labels: Disease (0.8078)]
****

```

```

import os, os.path
path = os.path.expanduser('../nltk_data')
if not os.path.exists(path):
    os.mkdir(path)
os.path.exists(path)
True
import nltk.data
path in nltk.data.path
True

True

```

```

import nltk, csv, codecs, pydot
import numpy as np
from nltk import sent_tokenize, word_tokenize, pos_tag
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data] /root/nltk_data...
[nltk_data] Unzipping taggers/averaged_perceptron_tagger.zip.
True

```

```

import numpy as np
import pandas as pd

```

```
data = open('/content/drive/MyDrive/Colab Notebooks/Data Sets/ICD-10-CM.txt').read()
```

```
data.read()
```

```
'<p> 2019 ICD-10-CM Codes </p>\nDiseases of the blood and blood-forming organs and cer
tain disorders involving the immune mechanism.\nSymptoms, signs and abnormal clinical
and laboratory findings, not elsewhere classified.\nD10-D36 Benign neoplasms, except b
enign neuroend.\nD37-D48 Neoplasms of uncertain behavior, polycyt...\nDiseases of the
blood and blood-forming organs and certain disorders involving the immune mechanism.\n
Q38-Q45 Other congenital malformations of the di...\nQ50-Q56 Congenital malformations
of genital orga...\nQ60-Q64 Congenital malformations of the urinary ...\nQ65-Q79 Conge
nital malformations and deformation...\nQ90-Q99 Chromosomal abnormalities, not elsewe
re...\nICD-10-CM Range R00-R99.\nSymptoms, signs and abnormal clinical and laboratory
findings. not elsewhere classified.\nR00-R09 Symtoms and signs involving the circul
```

```
from nltk import tokenize
```

```
dt = tokenize.sent_tokenize(data)
```

```
dt
```

```
'</p>\n<p> at risk for fall (history of fall) </p>\n<p> ICD-10-CM Diagnosis Code Z91
'It means "not coded here".',
'A type 1 excludes note indicates that the code excluded should never be used at the
'A type 1 excludes note is for used for when two conditions cannot occur together, s
'</p>\n<p> Toxic effect of contact with venomous animals and plants ( </p>\n<p> Y83-
'It means "not coded here".',
'A type 1 excludes note indicates that the code excluded should never be used at the
'A type 1 excludes note is for used for when two conditions cannot occur together, s
'</p>\n<p> accidental drowning and submersion due to fall into water ( </p>\n<p> A t
'A type 2 excludes note indicates that the condition excluded is not part of the con
'When a type 2 excludes note appears under a code it is acceptable to use both the c
'</p>\n<p> accidental drowning and submersion due to cataclysm ( </p>\n<p> ICD-10-CM
'It means "not coded here".',
'A type 1 excludes note indicates that the code excluded should never be used at the
'A type 1 excludes note is for used for when two conditions cannot occur together, s
'</p>\n<p> exposure to: </p>\n<p> failure in dosage of radiation or temperature duri
'It means "not coded here".',
'A type 1 excludes note indicates that the code excluded should never be used at the
'A type 1 excludes note is for used for when two conditions cannot occur together, s
'</p>\n<p> arson ( </p>\n<p> ICD-10-CM Diagnosis Code X97 </p>\n<p> Assault by smoke
'A type 2 excludes note indicates that the condition excluded is not part of the con
'When a type 2 excludes note appears under a code it is acceptable to use both the c
'</p>\n<p> explosions ( </p>\n<p> ICD-10-CM Diagnosis Code W35 </p>\n<p> Explosion a
'If no such documentation is present, code to accidental (unintentional) </p>\n<p> M
'It means "not coded here".',
'A type 1 excludes note indicates that the code excluded should never be used at the
'A type 1 excludes note is for used for when two conditions cannot occur together, s
'</p>\n<p> surgical and medical procedures as the cause of abnormal reaction of the
'A type 2 excludes note indicates that the condition excluded is not part of the con
'When a type 2 excludes note appears under a code it is acceptable to use both the c
'</p>\n<p> breakdown or malfunctioning of medical device (during procedure) (after i
'A type 2 excludes note indicates that the condition excluded is not part of the con
'When a type 2 excludes note appears under a code it is acceptable to use both the c
'</p>\n<p> later complications following use of medical devices without breakdown or
```

```
'</p>\n<p> breakdown or malfunction of medical devices (during use) (after implantat
'It means "not coded here".',
'A type 1 excludes note indicates that the code excluded should never be used at the
'A type 1 excludes note is for used for when two conditions cannot occur together, s
'</p>\n<p> misadventures to patients during surgical and medical care, classifiable
'A type 2 excludes note indicates that the condition excluded is not part of the con
'When a type 2 excludes note appears under a code it is acceptable to use both the c
'</p>\n<p> breakdown or malfunctioning of medical device (after implantation) (durin
'They are not to be used for single-condition coding.',
'</p>\n<p> V00-Y99 Supplementary factors related to causes of morbidity classified e
'They are not to be used for single-condition coding.',
'</p>\n<p> Y62-Y69 Misadventures to patients during surgica... </p>\n<p> Y70-Y82 Med
'It means "not coded here".',
'A type 1 excludes note indicates that the code excluded should never be used at the
'A type 1 excludes note is for used for when two conditions cannot occur together, s
'</p>\n<p> examinations related to pregnancy and reproduction ( </p>\n<p> ICD-10-CM
'</p>\n<p> Type 1 Excludes </p>\n<p> diagnostic examination- code to sign or symptom
'A corresponding procedure code must accompany a Z code if a procedure is performed.

"Categories Z00-Z99 </p>\n<p> ICD-10-CM Range Z00-Z99 </p>\n<p> Factors influencing
'This can arise in two main ways: </p>\n<p> (a) When a person who may or may not be
"</p>\n<p> (b) When some circumstance or problem is present which influences the per
'</p>\n<p> Persons with potential health hazards related to socioeconomic and psycho
'</p>\n<p> Type 1 Excludes </p>\n<p> examinations related to pregnancy and reproduct
```

```
from flair.data import Sentence
from flair.models import SequenceTagger

# load the NER tagger (medical or standard)
tagger = SequenceTagger.load('/content/drive/MyDrive/Colab Notebooks/Data Sets/best-model.pt')
# tagger = SequenceTagger.load('ner-ontonotes')

# run NER over sentences
for sent in dt:

    tagged_sentence = Sentence(sent)
    tagger.predict(tagged_sentence)

    print(tagged_sentence)

    entities_found = tagged_sentence.get_spans('ner')

    if entities_found:
        print(' Found entities:')

        # iterate over entities and print
        for entity in entities_found:
            print(' ' + str(entity))
    else:
        print(' (No entities found)')

    print('****')
```

```

span [10,11,12]: nonproliferative diabetic retinopathy [- Labels: Disease (0.7474)]
Span [14,15]: "macular edema" [- Labels: Disease (0.7474)]

****

Sentence: "E09.3399 Drug or chemical induced diabetes mellitus with moderate nonproli
Found entities:
Span [6,7]: "diabetes mellitus" [- Labels: Disease (0.7175)]
Span [10,11,12]: "nonproliferative diabetic retinopathy" [- Labels: Disease (0.
Span [14,15]: "macular edema" [- Labels: Disease (0.7537)]
Span [17,18]: "unspecified eye" [- Labels: Disease (0.9127)]

****

Sentence: "E09.3411 Drug or chemical induced diabetes mellitus with severe nonprolife
Found entities:
Span [6,7]: "diabetes mellitus" [- Labels: Disease (0.7339)]
Span [10,11,12]: "nonproliferative diabetic retinopathy" [- Labels: Disease (0.
Span [14,15]: "macular edema" [- Labels: Disease (0.8381)]

****

Sentence: "E09.3412 Drug or chemical induced diabetes mellitus with severe nonprolife
Found entities:
Span [6,7]: "diabetes mellitus" [- Labels: Disease (0.719)]
Span [10,11,12]: "nonproliferative diabetic retinopathy" [- Labels: Disease (0.
Span [14,15]: "macular edema" [- Labels: Disease (0.8375)]
Span [17,18]: "left eye" [- Labels: Disease (0.8173)]

****

Sentence: "E09.3413 Drug or chemical induced diabetes mellitus with severe nonprolife
Found entities:
Span [6,7]: "diabetes mellitus" [- Labels: Disease (0.6656)]
Span [10,11,12]: "nonproliferative diabetic retinopathy" [- Labels: Disease (0.
Span [14,15]: "macular edema" [- Labels: Disease (0.8112)]

****

Sentence: "E09.3419 Drug or chemical induced diabetes mellitus with severe nonprolife
Found entities:
Span [6,7]: "diabetes mellitus" [- Labels: Disease (0.7226)]
Span [10,11,12]: "nonproliferative diabetic retinopathy" [- Labels: Disease (0.
Span [14,15]: "macular edema" [- Labels: Disease (0.8187)]
Span [17,18]: "unspecified eye" [- Labels: Disease (0.9103)]

****

Sentence: "E09.3491 Drug or chemical induced diabetes mellitus with severe nonprolife
Found entities:
Span [4,5,6,7]: "chemical induced diabetes mellitus" [- Labels: Disease (0.7374)
Span [10,11,12]: "nonproliferative diabetic retinopathy" [- Labels: Disease (0.
Span [14,15]: "macular edema" [- Labels: Disease (0.7683)]

****

Sentence: "E09.3492 Drug or chemical induced diabetes mellitus with severe nonprolife
Found entities:
Span [6,7]: "diabetes mellitus" [- Labels: Disease (0.7185)]
Span [10,11,12]: "nonproliferative diabetic retinopathy" [- Labels: Disease (0.
Span [14,15]: "macular edema" [- Labels: Disease (0.7652)]
Span [17,18]: "left eye" [- Labels: Disease (0.811)]

****

Sentence: "E09.3493 Drug or chemical induced diabetes mellitus with severe nonprolife
Found entities:
Span [6,7]: "diabetes mellitus" [- Labels: Disease (0.6729)]
Span [10,11,12]: "nonproliferative diabetic retinopathy" [- Labels: Disease (0.
Span [14,15]: "macular edema" [- Labels: Disease (0.7426)]

****

Sentence: "E09.3499 Drug or chemical induced diabetes mellitus with severe nonprolife
Found entities:

```

```
Found entities:
[{"start": 5, "end": 10, "label": "PERSON", "text": "John"}, {"start": 15, "end": 20, "label": "PERSON", "text": "Doe"}, {"start": 25, "end": 30, "label": "PERSON", "text": "Smith"}]
```

▶ Executing (10m 7s) ... > pr... > for... > em... > em... > _add_embeddin... > get_repres... > for... > _call... > forw... ... ✕