Nripesh Karmacharya (21)

Umesh Shrestha (47)

Sabin Thapa (54)

AI in Health - A Mini Project for the course Artificial Intelligence

Purpose

The main purpose of this project is to extract informative and vital information from textual medical records.

Sources of data collection

1. Dhulikhel Hospital
2. Dadheldhura Hospital
3. mtsamples.com

Requirements

Working platform: Jupyter Notebook

Prerequisite: Python3, pip (`pip install pip`)

Database: MongoDB

Required libraries and installation commands

Library	Command
jupyter	*pip install jupyter*
spacy	*pip install spacy*
medspacy	*pip install medspacy*
pandas	*pip install pandas*
matplotlib	*pip install matplotlib*
pymongo	*pip install pymongo*

Required models:

a. **en_core_web_sm**: This model is used to break down the English language into tokens and helps recognize entities such as date, cardinals, people, organizations, money, etc. It is downloaded using the following code:

   spacy download en_core_web_sm

b. **en_ner_bc5cdr_md**: This model is used to recognize entities like diseases and chemicals from the text. The model is downloaded through the following code:

   pip install

   https://s3-us-west-2.amazonaws.com/ai2-s2-scispacy/releases/v0.5.1/en_ner_bc5cdr_md-0.5.1.tar.gz

c. **en_core_med7_lg**(vector based model): This model is used to recognize entities such as drugs, their dosage, frequency of dose, form of application, etc. The model is downloaded through the following code:

   pip install

   https://huggingface.co/kormilitzin/en_core_med7_lg/resolve/main/en_core_med7_lg-any-py3-none-any.whl

Working

After installing all the requirements, first we need to start mongo db. The mongo db must be running in the local host machine in port 27017. MongoDB can be started by following commands in terminal/command prompt:

a. **In linux:** *sudo systemctl start mongod*

b. **In windows:** *mongod*

If mongo db is installed successfully, then the above commands will start the mongo db at

localhost at port 27017. Now, we need to place the data file "MedicalData.txt" in the same place as the notebook "MiniProject-AI.ipynb". Then, we can run all the cells simply by clicking the command "Run all cells" under the "Cell" sub-menu. Or, individual cells can be run by the command "shift+enter". If all the cells run successfully, we can now see that the new database 'medical_records' has been initialized, and the data has been inserted into the collection named 'records'. Snapshot of the instance after the data has been inserted into the database can be seen at the results section.

Results

## a. Entities in text data recognized by the English model:

A **38-year-old** `DATE` male presented to the emergency department (ED) with complaints of fever, headache, retro-orbital pain, and myalgia of **one-week** `DATE` duration. The patient reported high-grade intermittent fever for **the past seven days** `DATE` that was associated with rigors and chills. His other complaints included reddish-colored urine with clots for **two days** `DATE` along with **two** `CARDINAL` episodes of gum bleed. The pertinent denials included sore throat, chest pain, shortness of breath, vomiting, abdominal pain, diarrhea, burning micturition, dysuria, urinary frequency, and contact with animals. On physical examination, the patient was vitally stable, oriented to person, place and time, and had mild conjunctival pallor. On skin examination, there were multiple purpura and petechiae on the left shoulder, trunk, and both legs, and **two** `CARDINAL` large ecchymotic lesions on the patient's back. The rest of the systemic examination was unremarkable.The baseline workup done in the **ED** `ORG` was normal except for a low platelet count of **20,000** `CARDINAL` x **109** `CARDINAL` /L and a low **WBC** `ORG` cell **WBC** `ORG` count of **2.4** `CARDINAL` x 109/L. The patient was admitted in the isolation ward with the suspicion of **DHF** `ORG` (**Dengue Hemorrhagic Fever** `ORG`) which was later confirmed on serological evidence with positive dengue-specific **IgM** `ORG` with a value of **1.94** `CARDINAL` on enzyme-linked immunosorbent assay (**ELISA** `ORG`). Furthermore, the tourniquet test was performed, which turned out to be positive. The patient was treated conservatively with acetaminophen **one** `CARDINAL` gram **three** `CARDINAL` times a day and **one** `CARDINAL` liter infusion of normal saline twice **daily** `DATE` . The patient continued to have reddish-colored urine and had **one** `CARDINAL` more episode of gums bleed on **his third day** `DATE` of admission with a platelet count of **6,000** `CARDINAL` x **109** `CARDINAL` /L. Infectious diseases department was consulted who recommended transfusion **six** `CARDINAL` units of platelets. Platelet count continued to deteriorate further with a count of **2,000** `CARDINAL` x **109** `CARDINAL` /L on **his sixth day** `DATE` of admission but with clinical improvement in his fever and bleeding manifestations. The patient was monitored **daily** `DATE` with blood counts and bleeding manifestations. The patient improved clinically on **the 11th day** `DATE` of his admission with no hemorrhagic manifestations and normalization of blood counts (platelet count **112** `CARDINAL` x **109** `CARDINAL` /L and **WBC** `ORG` count **6.3** `CARDINAL` x **109** `CARDINAL` /L). The patient was discharged home with a follow-up appointment **one month later** `DATE` .

b. Dependency graph created by the model:

There is **no evidence of** `NEGATED_EXISTENCE` **pneumonia** `DISEASE` .
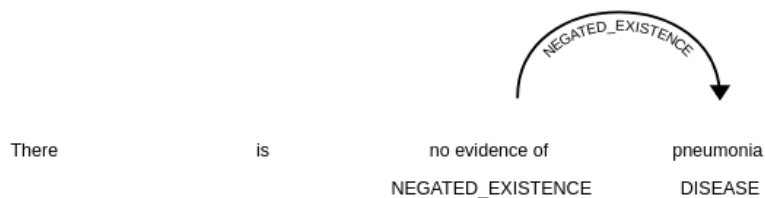
Her **mother** `FAMILY` suffered from **leukemia** `DISEASE` .
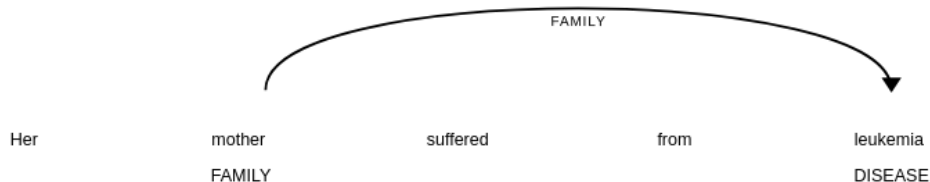
c. Output of the extracted data:

[{'Sex': 'male',

 'Age': '38-year-old',

 'Symptom Duration': 'one-week',

 'Initial Symptoms': ['complaints of fever',

 'headache',

 'myalgia'],

'Not Suffering From': [],

'Family History': [],

'Examinations': [{'physical examination': {'DISEASE': ['conjunctival pallor']}},

 {'skin examination': {'DISEASE': ['purpura',

   'petechiae',

   'ecchymotic lesions']}}],

'Vitals': {'platelet count': ['20,000x109/L',

   '2,000x109/L'],

 'WBC count': ['2.4x109/L']},

'Diseases': ['fever',

 'purpura',

 'petechiae',

 'ecchymotic lesions',

 'L. Infectious diseases',

 'bleeding'],

'Medication': [{'DRUG': 'acetaminophen',

   'STRENGTH': 'one gram',

   'FREQUENCY': 'three times a day'},

 {'DOSAGE': 'liter',

   'FORM': 'infusion',

   'DRUG': 'normal saline',

   'FREQUENCY': 'twice daily'},

 {'FORM': 'transfusion', 'DOSAGE': 'six units', 'DRUG': 'platelets'}],

'Status': 'Recovered'}]

d. Data visualization:
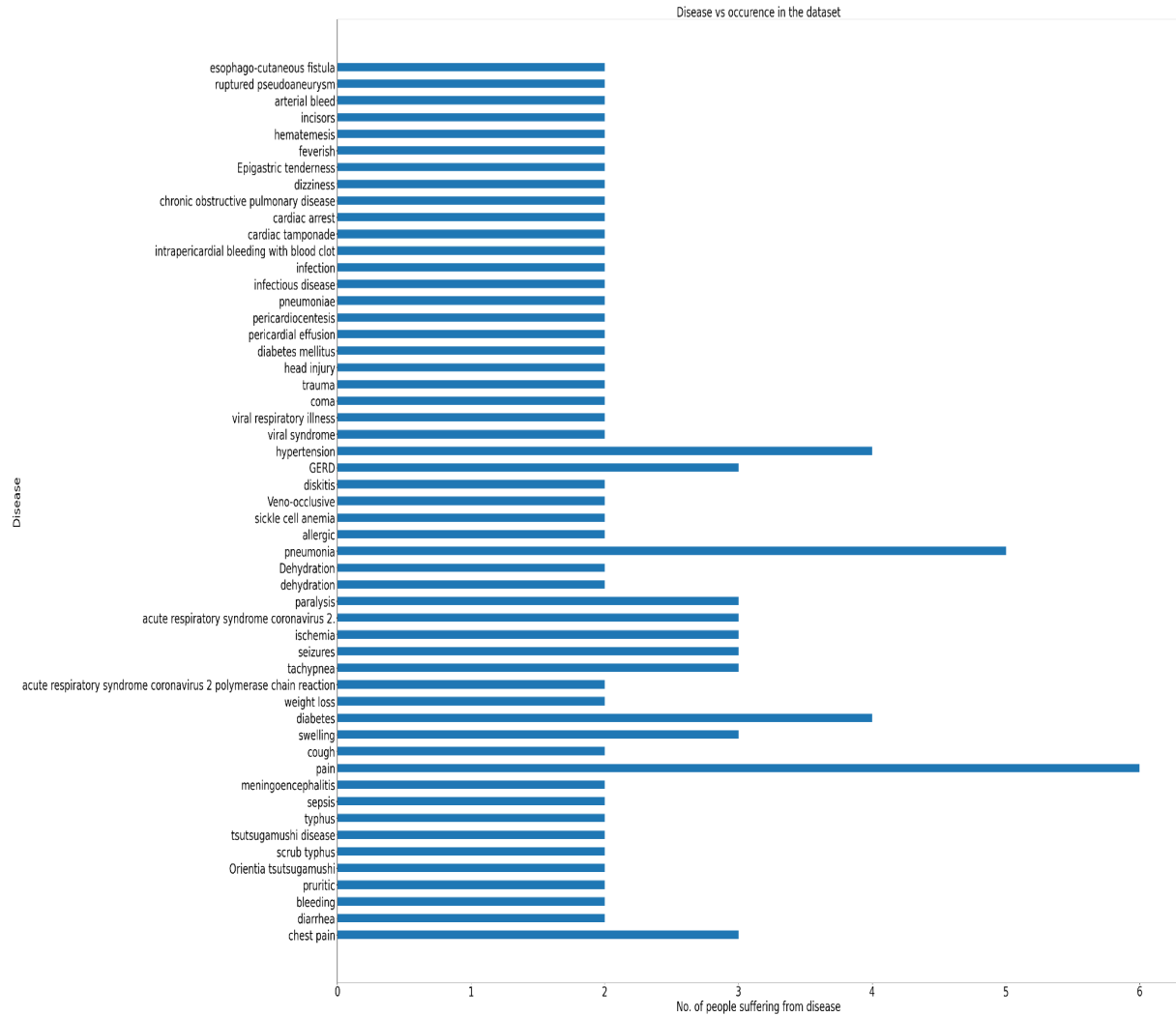
Disease vs occurence in the dataset

e. Stored in MongoDB:

medical_records.records

Documents | Aggregations | Schema | Explain Plan | Indexes | Validation

Filter ⬚ 🕐 ▼   Type a query: { field: 'value' }

⬇ ADD DATA ▼   ↗ EXPORT COLLECTION

```
_id: ObjectId('63959a756a92ee207a61e36c')
Sex: "male"
Age: "38-year-old"
Symptom Duration: "one-week"
> Initial Symptoms: Array
> Not Suffering From: Array
> Family History: Array
> Examinations: Array
> Vitals: Object
> Diseases: Array
> Medication: Array
Status: "Recovered"
```

```
_id: ObjectId('63959a756a92ee207a61e36d')
Sex: "female"
Age: "65 years old"
Symptom Duration: "4-day"
> Initial Symptoms: Array
> Not Suffering From: Array
> Family History: Array
> Examinations: Array
> Vitals: Object
> Diseases: Array
> Medication: Array
Status: "Recovered"
```

```
_id: ObjectId('63959a756a92ee207a61e36e')
Sex: "female"
Age: "27 years"
Symptom Duration: "14 days"
> Initial Symptoms: Array
> Not Suffering From: Array
> Family History: Array
> Examinations: Array
> Vitals: Object
> Diseases: Array
> Medication: Array
```

My Queries
Databases +
Search
▸ admin
▸ config
▸ local
▾ medical_records
  ■ records ...

Database Record example:

```
_id: ObjectId('63959a756a92ee207a61e36d')
Sex: "female"
Age: "65 years old"
Symptom Duration: "4-day"
∨ Initial Symptoms: Array
    0: "fever"
    1: "headache"
∨ Not Suffering From: Array
    0: "respiratory distress"
∨ Family History: Array
∨ Examinations: Array
  ∨ 0: Object
    ∨ physical examination: Object
      ∨ DISEASE: Array
          0: "pruritic"
  ∨ 1: Object
    ∨ Laboratory examination: Object
      ∨ CHEMICAL: Array
          0: "aspartate"
          1: "alanine"
  ∨ 2: Object
    ∨ Serological examination: Object
      ∨ DISEASE: Array
          0: "Orientia tsutsugamushi"
∨ Vitals: Object
    SpO2: "94"
  ∨ platelet count: Array
      0: "70000/μL"
∨ Diseases: Array
    0: "fever"
    1: "headache"
    2: "pruritic"
    3: "Orientia tsutsugamushi"
    4: "scrub typhus"
    5: "tsutsugamushi disease"
∨ Medication: Array
  ∨ 0: Object
      DRUG: "minocycline"
  Status: "Recovered"
```