Name: Trinadha Raji Muppala

Class: 7

KDM Lab3

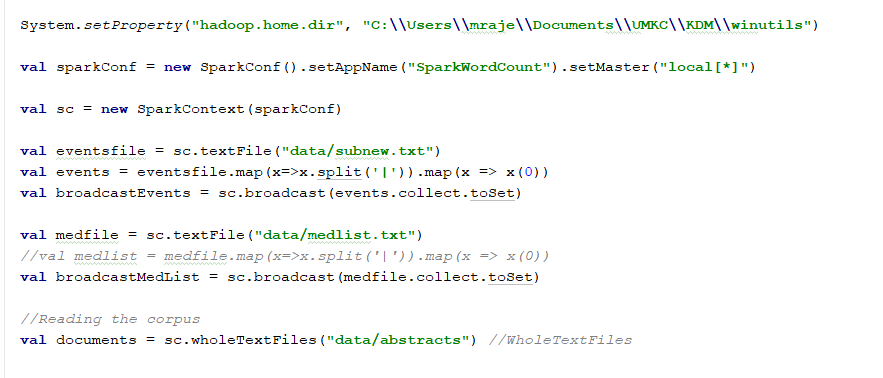
**Statistics for Parkinson's** **disease abstracts**

Part 2:

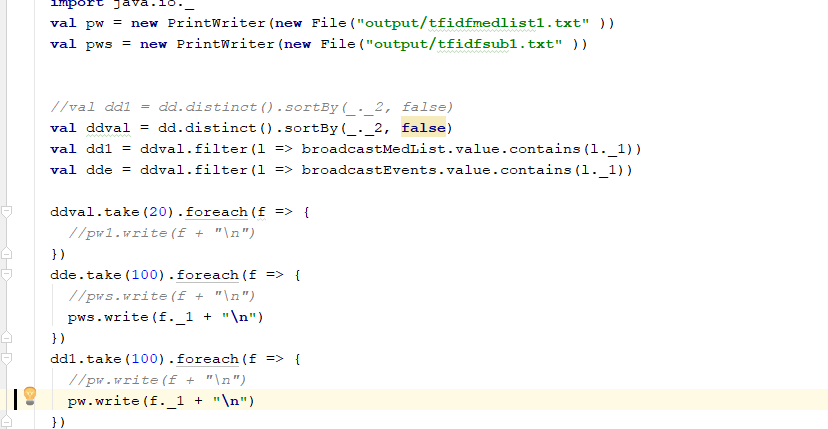
First found medical words and events which got top tfidf and then found the triplets which contains those words.

1. Find out medical words and events which got top TFIDF values

* Broad cast Medical words and Events to find out top TFIDF medical words and events

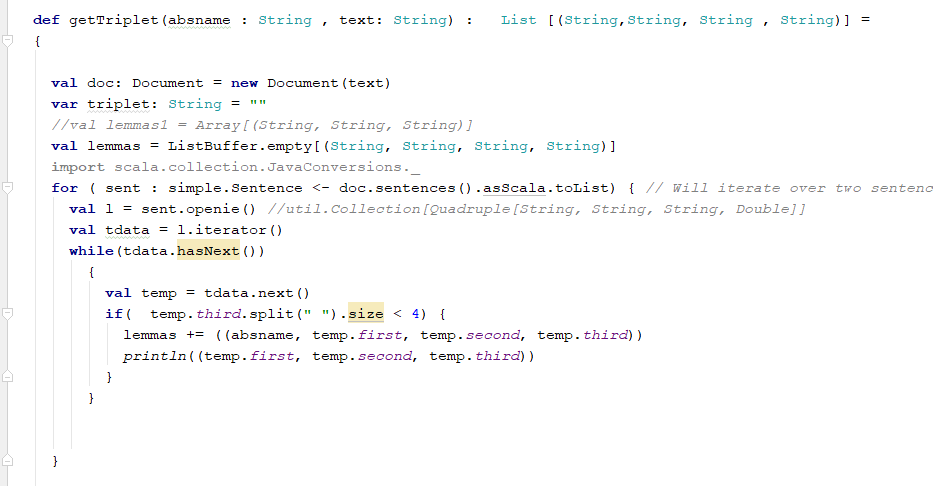


* From TFIDF list filter the list which got medical words , events. Get the top 100 words from the list and write to 2 separate files

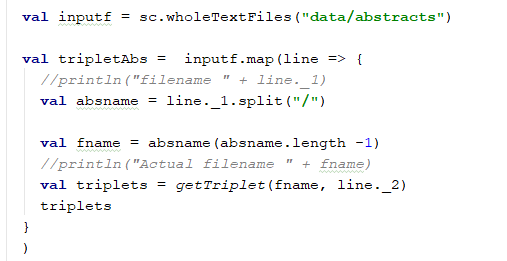


* Manually picked few medical words and top events from the top tfidf medica/events list, used that list and findout the subset of the triplets which got these words either in subject or predicate

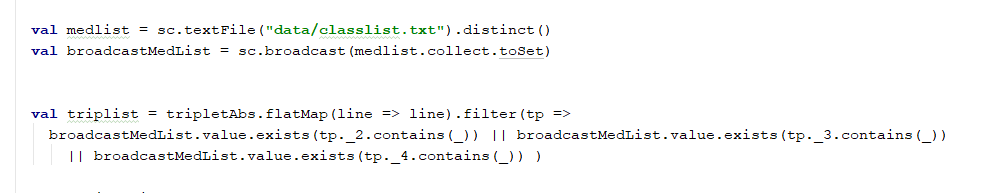
1. Function to find out triplets using coreNLP, passing abstract name also that way each triplet got which abstract it got generated from.



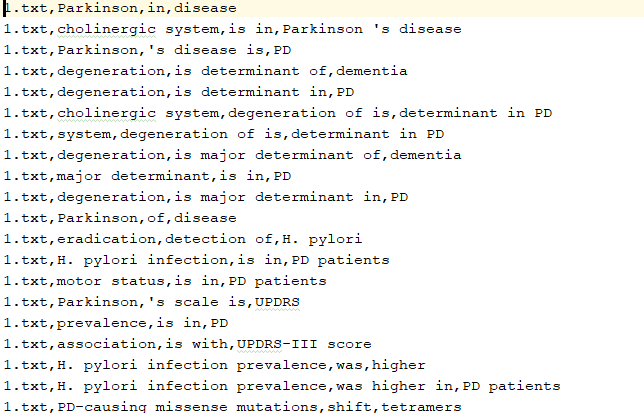
* Find all the triplets



* Filter triplets got toptfidf medical/events words

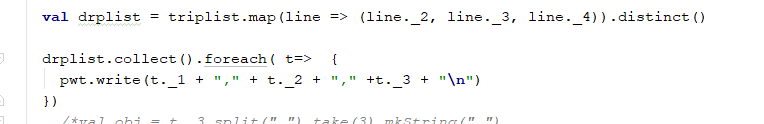


Triplets Output:

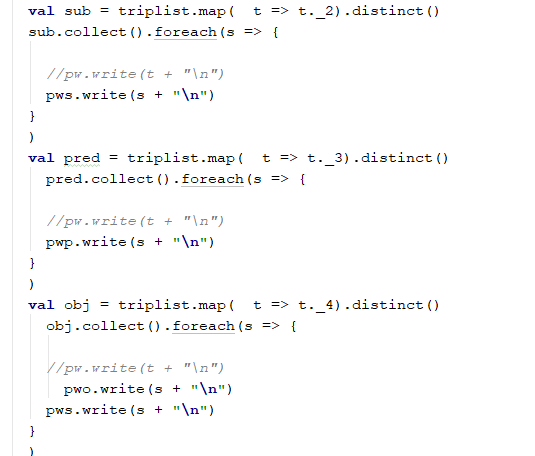


1. Find out Subjects, predicate and Object from the triplet list

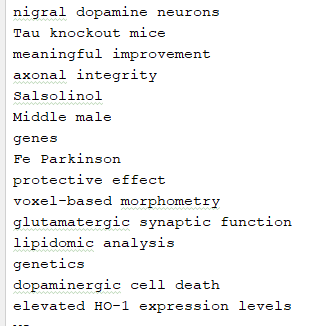
* Find distinct triplets



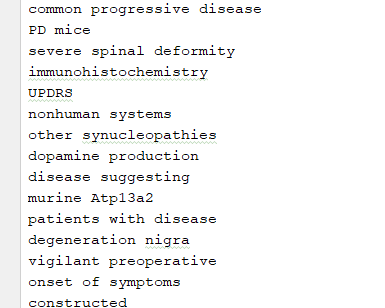
* Find out subject , object and predicate from the triplets



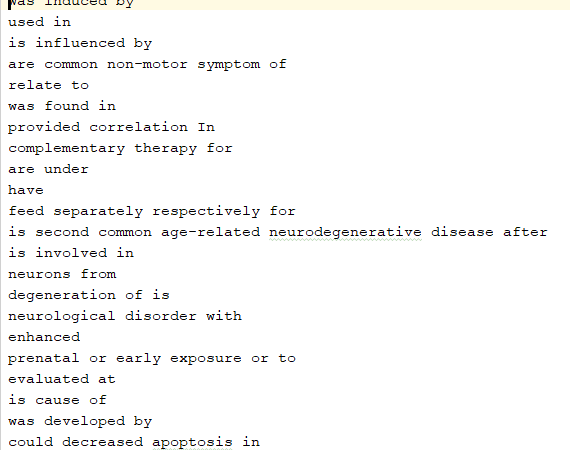
Subjects:



Objects:



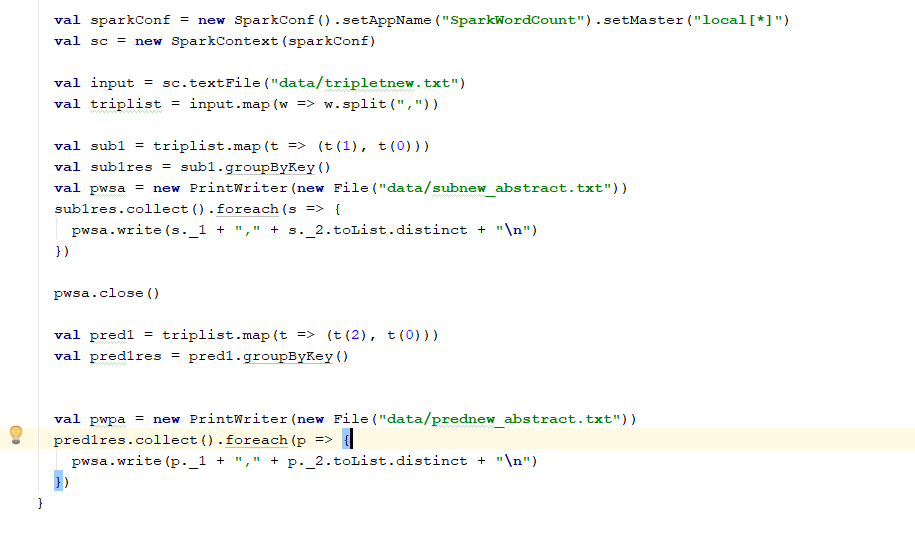
Predicates:



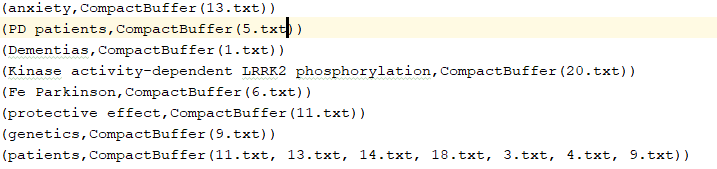
PART III

Find out the origin abstract for the subject/predicate. Above provide triplet function which has inbound parameter as abstract name and each line. Abstract name will be added to triplet . First word in the triplet is the abstract name and rest is the triplet value separated by coma.

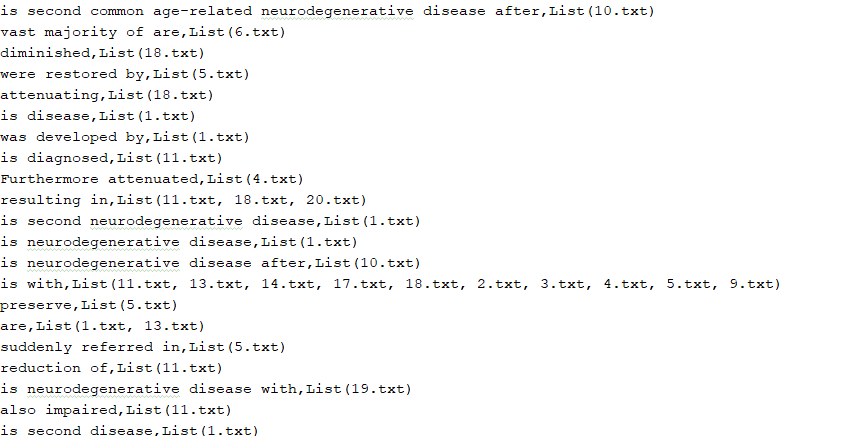
Map the triplets and change subject/predicate as key and abstract as value. Then group by subject using key



Each subject originated from the abstract:



Each predicate originated from the abstract:



**PART IV:**

Common entities:

patients,List(11.txt, 13.txt, 14.txt, 18.txt, 3.txt, 4.txt, 9.txt)

Anxiety disorders,List(13.txt,18.txt)  
Parkinson 's disease,List(1.txt, 10.txt, 11.txt, 13.txt, 15.txt, 16.txt, 17.txt, 18.txt, 19.txt, 2.txt, 20.txt, 3.txt, 4.txt, 6.txt, 7.txt, 9.txt)  
Resveratrol,List(9.txt,12.txt,13.txt)

Common Predicates:

resulting in,List(11.txt, 18.txt, 20.txt)  
is with,List(11.txt, 13.txt, 14.txt, 17.txt, 18.txt, 2.txt, 3.txt, 4.txt, 5.txt, 9.txt)

Unique Entities:

Accumulation,List(3.txt)  
acquired nonverbal expressive impairment,List(12.txt)  
manifesting,List(6.txt)  
Acupuncture,List(10.txt)  
association,List(17.txt)  
our studies,List(2.txt)

Unique Predicates:

preserve,List(5.txt)  
are,List(1.txt, 13.txt)  
suddenly referred in,List(5.txt)  
reduction of,List(11.txt)