Boğaziçi University

CMPE 493 - Information Retrieval

Spring 2020

Assignment III

Extractive Text Summarization with PageRank

Mahmut Uzunpostalcı

1 Assumptions

Some rows of the mapping file didn't have the abstract column. So while reading it would occur as nan type in the program. So I have simply removed those papers from the list.
I have tried different thresholds for creating edges between nodes but the 0.1 seemed fine so I have used it.
I have used nltk library as a tokenizer for both words and sentence.
I have selected topic 23 "coronavirus hypertension" as my third topic.

2 Top 10 Documents of Selected Topics

2.1 Topic 1: Coronavirus Origin

Cord ID	Score
1mjaycee	0.0528
0xhho1sh	0.0466
52kqp9yw	0.0450
awitk3se	0.0419
k2ixwz9w	0.0411
2inlyd0t	0.0395
89fol3pq	0.0391
yn8nzezq	0.0350
1qkwsh6a	0.0340
9il7coyk	0.0324

2.2 Topic 13: How Does Coronavirus Spread

Cord ID	Score
1mjaycee	0.0528
0xhho1sh	0.0466
52kqp9yw	0.0450
awitk3se	0.0419
k2ixwz9w	0.0411
2inlyd0t	0.0395
89fol3pq	0.0391
yn8nzezq	0.0350
1qkwsh6a	0.0340
9il7coyk	0.0324

2.3 Topic 23: Coronavirus Hypertension

Cord ID	Score
1mjaycee	0.0528
0xhho1sh	0.0466
52kqp9yw	0.0450
awitk3se	0.0419
k2ixwz9w	0.0411
2inlyd0t	0.0395
89fol3pq	0.0391
yn8nzezq	0.0350
1qkwsh6a	0.0340
9il7coyk	0.0324

3 Top 20 Sentences of Selected Topics

3.1 Topic 1: Coronavirus Origin

covid-19 is similar to severe acute respiratory syndrome coronavirus (sars-cov) virus in its pathogenicity, clinical spectrum, and epidemiology.
Score: 0.0359

the virus was named severe acute respiratory syndrome coronavirus 2 (sars-cov-2) by the international committee on taxonomy of viruses.
Score: 0.0266

the coronavirus disease 19 (covid-19) is a highly transmittable and pathogenic viral infection caused by severe acute respiratory syndrome coronavirus 2 (sars-cov-2), which emerged in wuhan, china and spread around the world.
Score: 0.0253

in the current review, we summarize and comparatively analyze the emergence and pathogenicity of covid-19 infection and previous human coronaviruses severe acute respiratory syndrome coronavirus (sars-cov) and middle east respiratory syndrome coronavirus (mers-cov).
Score: 0.02407

abstract the emergence of severe acute respiratory syndrome coronavirus 2 (sars-cov-2; previously provisionally named 2019 novel coronavirus or 2019-ncov) disease (covid-19) in china at the end of 2019 has caused a large global outbreak and is a major public health issue.
Score: 0.0220

abstractthe outbreak of 2019-ncov pneumonia (covid-19) in the city of wuhan, china has resulted in more than 70,000 laboratory confirmed cases, and recent studies showed that 2019-ncov (sars-cov-2) could be of bat origin but involve other potential intermediate hosts.
Score: 0.0219

the world health organization (who) has issued a warning that, although the 2019 novel coronavirus (covid-19) from wuhan city (china), is not pandemic, it should be contained to prevent the global spread.
Score: 0.0201

as of 11 february 2020, data from the world health organization (who) have shown that more than 43 000 confirmed cases have been identified in 28 countries/regions, with >99
the scientific interest on coronaviruses increased after the emergence of severe acute respiratory syndrome coronavirus (sars-cov) outbreaks in 2002-2003 followed by middle east respiratory syndrome cov (mers-cov).
Score: 0.0197

abstract in early december 2019, an outbreak of coronavirus disease 2019 (covid-19), caused by a novel severe acute respiratory syndrome coronavirus 2 (sars-cov-2), occurred in wuhan city, hubei province, china.
Score: 0.0193

world health organization has declared the ongoing outbreak of coronavirus disease 2019 (covid-19) a public health emergency of international concern.
Score: 0.0192

a comparison with sars-cov, middle east respiratory syndrome coronavirus, community-acquired human coronaviruses and other pathogenic viruses including human immunodeficiency viruses is made.
Score: 0.0191

summary an outbreak of coronavirus disease 2019 (covid-19) caused by the 2019 novel coronavirus (sars-cov-2) began in the city of wuhan in china and

has widely spread worldwide.
Score: 0.0189

in 2003, severe acute respiratory syndrome coronavirus (sars-cov) caused one of the most devastating epidemics known to the developed world.
Score: 0.0183

in 2012, a novel human coronavirus, now called middle east respiratory syndrome coronavirus (mers-cov), has emerged in the middle east to cause fatal human infections in three continents.
Score: 0.0182

covid-19 causes covid-19 disease that has similar symptoms as sars-cov.
Score: 0.0163

among patients with pneumonia caused by sars-cov-2 (novel coronavirus pneumonia or wuhan pneumonia), fever was the most common symptom, followed by cough.
Score: 0.0153

aside from ratg13, pangolin-cov is the most closely related cov to sars-cov-2.
Score: 0.0152

secondly, bats can serve as the origin and natural animal reservoir of deadly human viruses.
Score: 0.0151

this decadeã¢â‚¬â„¢s first cov, named 2019-ncov, emerged from wuhan, china, and declared as ã¢â‚¬ëœpublic health emergency of international concernã¢â‚¬â„¢ on january 30(th), 2020 by the world health organization (who).
Score: 0.0150

3.2 Topic 13: How Does Coronavirus Spread

there is a new public health crises threatening the world with the emergence and spread of 2019 novel coronavirus (2019-ncov) or the severe acute respiratory syndrome coronavirus 2 (sars-cov-2).
Score: 0.0323

coronavirus disease 2019 (covid-19) caused by severe acute respiratory syndrome coronavirus 2 (sars-cov-2) is an ongoing global health emergency.
Score: 0.0318

2019-ncov can also be transmitted through the saliva, and the fetalã¢â‚¬â€œoral routes may also be a potential person-to-person transmission route.
Score: 0.0277

here we recommend the infection control measures during dental practice to block the person-to-person transmission routes in dental clinics and hospitals.
Score: 0.0269

a novel coronavirus, severe acute respiratory syndrome coronavirus 2 (sars-cov-2), emerged in wuhan, hubei province in china in december 2019 and caused a serious type of pneumonia called coronavirus disease 2019 or covid-19.
Score: 0.0259

extensive measures to reduce person-to-person transmission of covid-19 have been implemented to control the current outbreak.
Score: 0.0230

many people are asymptomatic.
Score: 0.0226

an acute respiratory disease, caused by a novel coronavirus (sars-cov-2, previously known as 2019-ncov), the coronavirus disease 2019 (covid-19) has spread throughout china and received worldwide attention.
Score: 0.0219

in december 2019, a novel coronavirus (2019-ncov) caused an outbreak in wuhan, china, and soon spread to other parts of the world.
Score: 0.0213

the outbreak of coronavirus disease 2019 (covid-19), caused by severe acute respiratory syndrome (sars) coronavirus 2 (sars-cov-2), has thus far killed over 3,000 people and infected over 80,000 in china and elsewhere in the world, resulting in catastrophe for humans.
Score: 0.0204

since december 2019, corona virus disease 2019 (covid-19) , an emerging infection disease occurred in wuhan, has spread in the mainland china.
Score: 0.0192

since its emergence in december 2019, corona virus disease 2019 (covid-19) has impacted several countries, affecting more than 90 thousand patients and making it a global public threat.
Score: 0.0191

routes of sars-cov-2 transmission are diversified and the main routes of transmission for covid-19 are droplet transmission and close contact transmission.
Score: 0.0185

all population have susceptibility to sars-cov-2.
Score: 0.0181

the disease is mild in most people; in some (usually the elderly and those with comorbidities), it may progress to pneumonia, acute respiratory distress syndrome (ards) and multi organ dysfunction.
Score: 0.0181

prevention entails home isolation of suspected cases and those with mild illnesses and strict infection control measures at hospitals that include contact and droplet precautions.
Score: 0.0176

the virus spreads faster than its two ancestors the sars-cov and middle east respiratory syndrome coronavirus (mers-cov), but has lower fatality.
Score: 0.0175

in this review, we summarized the latest research progress of the epidemiology, pathogenesis, and clinical characteristics of covid-19, and discussed the current treatment and scientific advancements to combat the epidemic novel coronavirus.

Score: 0.0169

in this review, we highlights the symptoms, epidemiology, transmission, pathogenesis, phylogenetic analysis and future directions to control the spread of this fatal disease.
Score: 0.0168

the novel coronavirus uses the same receptor, angiotensin-converting enzyme 2 (ace2) as that for sars-cov, and mainly spreads through the respiratory tract.
Score: 0.0162

3.3 Topic 23: Coronavirus Hypertension

background: china has experienced an outbreak of a novel human coronavirus, severe acute respiratory syndrome coronavirus 2 (sars-cov-2) since december 2019, and it was announced a worldwide pandemic in march 2020. there is limited evidence on the mortality risk effect of pre-existing comorbidity for clinical disease (coronavirus disease 2019 [covid-19]), which has important implications for early treatment.
Score: 0.02506

this meta-analysis is to analyze the correlation between hypertension, diabetes, coronary heart disease and covid-19 disease severity.
Score: 0.0225

results: we found that increasing age, male gender, and angiotensin-converting enzyme 2 (ace2) associated factors (including hypertension, diabetes, and cardiovascular diseases) adversely affected the viral clearance.
Score: 0.0205

conclusion we assessed the prevalence of comorbidities in the covid-19 infection patients and found underlying disease, including hypertension, respiratory system disease and cardiovascular, may be a risk factor for severe patients compared with non-severe patients.
Score: 0.0195

prevalence of male, elevated nt-probnp and ctni, hypertension and coronary heart disease were significantly higher in critical cases care patients than in the mild cases(all $p<0.05$).
Score: 0.0193

compared with the non-severe patient, the pooled odds ratio of hypertension, respiratory system disease, cardiovascular disease in severe patients were (or 2.36, 95
both univariate and multivariate logistic regression were used to analyze the correlation of past medical history including hypertension, diabetes and coronary heart disease (chd) , as well as the levels of serum nt-probnp and ctni to the disease severity of covid-19 patients.
Score: 0.0187

interpretation: hypertension, diabetes, and coronary heart disease can affect the severity of covid-19.

Score: 0.0178

there were significant correlations between covid-19 severity and hypertension [or=2.3 [95
conclusions: in addition to older age and male sex, hypertension, diabetes, and cvd were associated in univariate analyses with severe covid-19.
Score: 0.0176

the aim of this analysis is to determine the association of cardiovascular metabolic diseases with the development of covid-19.
Score: 0.0169

aims the aim of the meta-analysis was to assess the prevalence of comorbidities in the covid-19 infection patients and the risk of underlying diseases in severe patients compared to non-severe patients.
Score: 0.0167

background: covid-19 patients with chronic diseases such as hypertension, diabetes and coronary heart diseases is more likely to worsen, but with mixed results for covid-19 severity.
Score: 0.0165

we suggest that patients with cardiac diseases, hypertension, or diabetes, who are treated with ace2-increasing drugs, are at higher risk for severe covid-19 infection and, therefore, should be monitored for ace2-modulating medications, such as ace inhibitors or arbs.
Score: 0.0165

abstract coronavirus disease 2019 (covid-19), caused by a novel betacoronavirus severe acute respiratory syndrome coronavirus 2 (sars-cov-2), was first described in a cluster of patients presenting with pneumonia symptoms in wuhan, china, in december of 2019. over the past few months, covid-19 has developed into a worldwide pandemic, with over 400,000 documented cases globally as of march 24, 2020. the sars-cov-2 virus is most likely of zoonotic origin, but has been shown to have effective human-to-human transmission.
Score: 0.0163

abstract aims to analyze the potential mechanism of cardiovascular dysfunctions induced by coronavirus disease 2019 (covid-19) and to evaluate more effective therapeutic pathways for patients with cardiovascular diseases.
Score: 0.0162

conclusions: covid-19 can significantly affect the heart function and lead to myocardial injury.
Score: 0.0157

covid-19 results in mild symptoms in the majority of infected patients, but can cause severe lung injury, cardiac injury, and death.
Score: 0.0154

objective: to evaluate the risk of the common preexisting comorbidities, including hypertension, coronary heart diseases (chd), respiratory diseases and diabetes on covid-19 mortality, and provide clinical suggestions accordingly.
Score: 0.0152

notably, the most frequent comorbidities reported in these three studies of patients with covid-19 are often treated with angiotensin-converting enzyme (ace) inhibitors; however, treatment was not assessed in either study.
Score: 0.0149

4 Screenshot

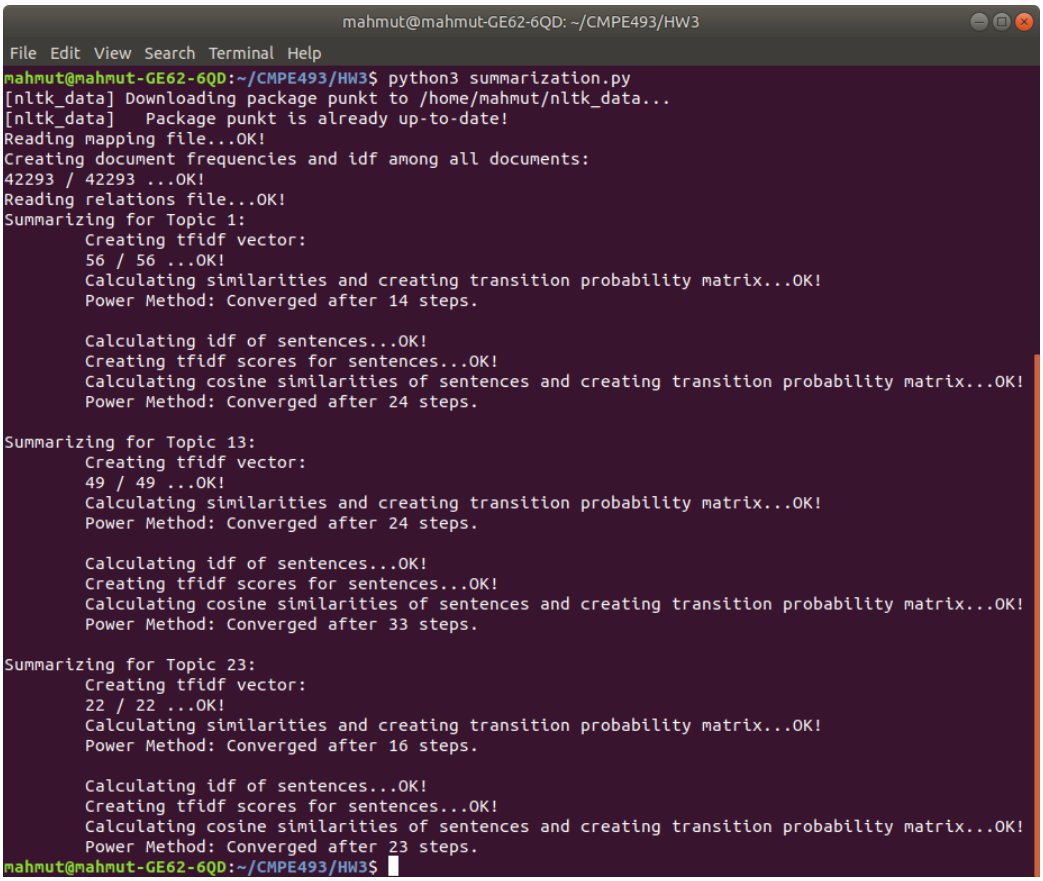

Figure 1: Screenshot of program running