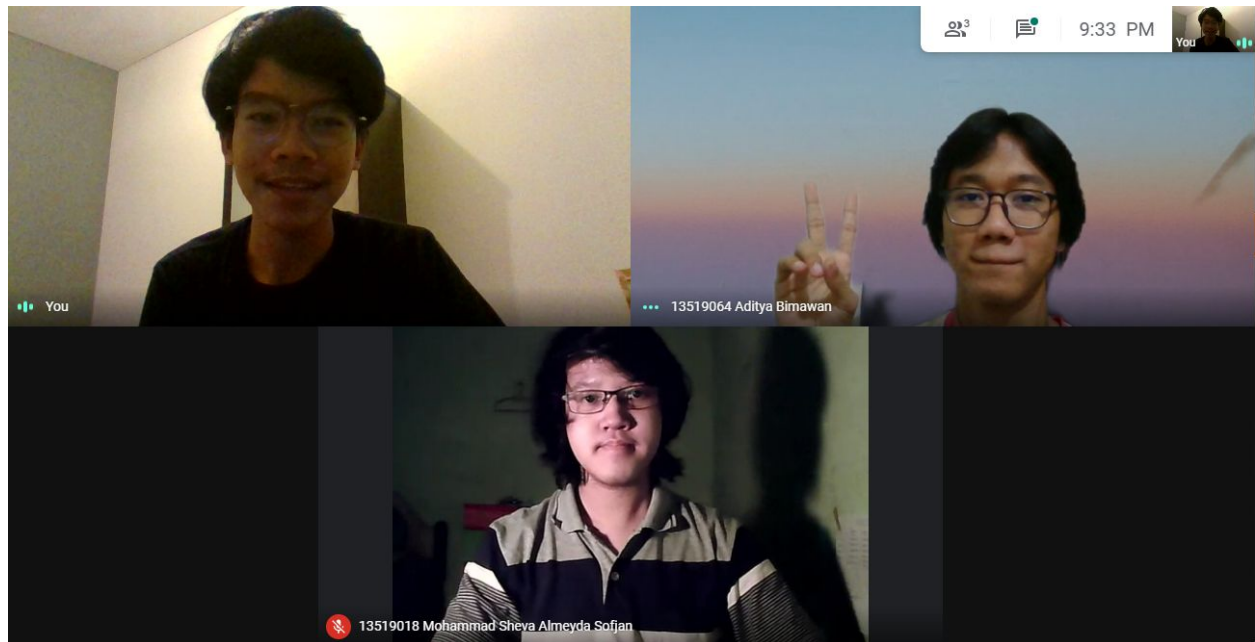


**LAPORAN TUGAS BESAR 2 IF2123**  
**Aljabar Linier dan Geometri**  
**Aplikasi Dot Product pada Sistem Temu-balik Informasi**



**DISUSUN OLEH**  
**KELOMPOK 53**

Muhammad Tito Prakasa	13519007
Mohammad Sheva Almeyda Sofjan	13519018
Aditya Bimawan	13519064

**TEKNIK INFORMATIKA**  
**SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA**  
**INSTITUT TEKNOLOGI BANDUNG**  
**2020/2021**

# **BAB I**

## **DESKRIPSI MASALAH**

Membuat sebuah program berupa *website* sederhana yang mampu menyelesaikan berbagai permasalahan yang berkaitan dengan pembelajaran pada mata kuliah Aljabar Linier dan Geometri dengan spesifikasi sebagai berikut :

1. Program mampu menerima *search query* berupa kata dasar maupun berimbuhan.
2. Program mampu menerima dokumen (*upload* maupun *scraping* dari situs web lain), minimal terdapat 15 dokumen berbeda.
3. Program mampu menampilkan tabel kemunculan *term* pada *query* dan dokumen.
4. Program mampu mengalkulasikan nilai similaritas (*cosine similarity*) antara *query* dengan dokumen , lalu mengurutkan hasil pencarian berdasarkan nilai similaritas.
5. Hasil pencarian yang terurut berdasarkan similaritas tertinggi dari hasil teratas hingga hasil terbawah berupa judul dokumen dan kalimat pertama dari dokumen tersebut. Menampilkan juga nilai similaritas tiap dokumen.

## BAB II

### TEORI SINGKAT

#### 2.1 Dot Product dan Norma Suatu Vektor

Misalkan terdapat suatu vektor tidak nol  $u$  dan  $v$  di dalam ruang vektor  $R^n$  dengan komponen  $u = (u_1, u_2, u_3, \dots, u_n)$  dan komponen  $v = (v_1, v_2, v_3, \dots, v_n)$  maka dot product (perkalian titik atau perkalian skalar) dari  $u$  dan  $v$  (ditulis  $u \cdot v$ ) adalah

$$u \cdot v = u_1 v_1 + u_2 v_2 + u_3 v_3 + \dots + u_n v_n$$

dan norma Euclidean (panjang) vektor  $u$  adalah

$$\|u\| = \sqrt{(u_1)^2 + (u_2)^2 + (u_3)^2 + \dots + (u_n)^2}$$

#### 2.2 Pemodelan Vektor dalam Information Retrieval (IR) System

*Information Retrieval System* atau Sistem Temu Balik Informasi adalah sistem pencarian informasi yang relevan terhadap kebutuhan pengguna (*query*) dari kumpulan informasi (dokumen) secara otomatis. Dalam kasus ini akan digunakan dokumen berbasis teks sebagai informasi.

Misalkan terdapat  $n$  kata berbeda sebagai suatu *term* atau kamus kata dalam suatu dokumen maupun *query*. Term / kata tersebut akan membentuk vektor  $u$  dalam ruang  $R^n$  dengan komponen  $u = (u_1, u_2, u_3, \dots, u_n)$ , dengan  $u_i$  menyatakan bobot suatu kata di dalam dokumen atau *query*.

#### 2.3 Term Frequency (TF)

*Term Frequency* merupakan frekuensi kemunculan suatu *term* (kata) dalam suatu dokumen atau *query*. *Term Frequency* dari suatu *term*  $t$  dalam dokumen  $d$  adalah

$$tf_{t,d} = \log_{10}(\text{count}(t, d) + 1)$$

, dengan  $\text{count}(t,d)$  menyatakan banyaknya kemunculan term  $t$  dalam dokumen  $d$ .

#### 2.4 Inverse Document Frequency (IDF)

*Inverse Document Frequency* menyatakan *rarity* atau kelangkaan suatu *term* dalam semua dokumen, semakin langka suatu *term* maka semakin besar nilai IDF dari *term* tersebut. IDF dari suatu *term*  $t$  dalam himpunan/ set dokumen  $D$  adalah

$$idf(t, D) = \log \left( \frac{N}{df(t) + 1} \right)$$

, dengan  $N$  menyatakan banyaknya dokumen dan  $df(t)$  menyatakan jumlah dokumen yang mengandung *term*  $t$ .

## 2.5 Term Frequency - Inverse Document Frequency (TF - IDF)

*Term Frequency - Inverse Document Frequency* menyatakan hasil perkalian antara nilai TF dan nilai IDF suatu *term* dalam sebuah dokumen dalam susunan/set dokumen. TF-IDF dari suatu *term*  $t$ , sebuah dokumen  $d$ , dan himpunan/set dokumen  $D$  adalah

$$tfidf(t, d, D) = tf(t, d) \cdot idf(t, D)$$

, dengan syarat  $d$  anggota  $D$ .

## 2.6 Cosine Similarity

*Cosine Similarity* merupakan salah satu metode untuk menentukan similaritas antara dua vektor . Semakin tinggi nilainya semakin tinggi similaritas dua vektor tersebut. Misalkan terdapat suatu vektor tidak nol  $u$  dan  $v$  di dalam ruang vektor  $R^n$  dengan komponen  $u = (u_1, u_2, u_3, \dots, u_n)$  dan komponen  $v = (v_1, v_2, v_3, \dots, v_n)$  maka nilai *Cosine Similarity* dua vektor tersebut adalah

$$sim(u, v) = \cos\theta = \frac{u \cdot v}{\|u\| \|v\|}$$

, dengan  $\theta$  menyatakan sudut yang dibentuk antara vektor  $u$  dan  $v$ .

### BAB III

#### IMPLEMENTASI PROGRAM

- File tfidf.py

Nama	Tipe data keluaran	Keterangan
stem(string)	List	Digunakan untuk stemming input kata
setkata(arrkalimat)	set	Mereturn set berisi semua kata yang muncul
wordcount(kalimat, setkata)	dictionary	Menghitung kemunculan tiap kata pada tiap dokumen/query
wordcount_matx(arrkalimat)	List of dictionary	mengubah tiap array dalam arrkalimat menjadi dictionary dengan keytiap kata dan value jumlah kemunculannya (serupa dgn vektor)
tf(wordcount_vect)	Dictionary	output dict vector dengan nilai tiap kolom berupa nilai TF pada dokumen wordcount_vect
tf_matx(wordcount_arr)	Matrix	output matrix dengan nilai tiap kolom berupa nilai TF dari kata tersebut pada dokumen masing-masing, dengan input matrix wordcount
idfvect(wordcount_arr)	Dictionary	mereturn vector (dictionary dengan value float) dengan inputarray vektor wordcount
tf_idf(tf, idf)	Dictionary	Mereturn dictionary berupa nilai tfidf suatu vektor
cosine_sim(tfidf1, tfidf2)	Float	Mereturn nilai cosine similarity antara 2 dict tfidf

- File webscraper.py

Nama	Tipe Data Keluaran	Keterangan
isiKonten(link)	String	Fungsi berparameter suatu link dan akan mereturn isi konten dari link tersebut (artikel berita)

- File mainjson.py

Nama	Tipe data keluaran	Keterangan
getsorted(query)	Dictionary	Mereturn dictionary yang berisi artikel/dokumen yang sudah disortir berdasarkan similaritas dengan query
similarity(wcount ,tiq,idfvect)	Float	Mereturn similaritas antara tiq dan tfidf dari tf(wcount) dan idfvect
tablemaker(query )	Dataframe(pandas)	Mereturn tabel kemunculan term (sudah di sortir berdasarkan similaritas) berupa dataframe

- File maintxt.py

Nama	Tipe data keluaran	Keterangan
getsortedtxt(query)	Dictionary	Mereturn dictionary yang berisi artikel/dokumen yang sudah disortir berdasarkan similaritas dengan query
get_localcorpus()	List of dictionary	Mereturn list of dictionary berupa corpus dari file .txt pada path ./txt/
get_localcontent(doc)	String	Mereplace newline (\n) pada hasil read file dengan spasi ( ' ')
tablemakertxt(query)	Dataframe(pandas)	Mereturn tabel kemunculan term (sudah di sortir berdasarkan similaritas) berupa dataframe

- File flasksite.py

Nama	Keterangan
searchForm(form )	Display search bar dan pilihan dibawahnya
allowed_file(filen	Nantinya digunakan di upload() . Mereturn apakah file dapat

ame)	diupload (dalam ekstensi .txt) atau tidak
home()	Render home.html (homepage)
search()	Render hasil search sesuai dengan pilihan
txt(txt_name)	Mereturn file txt dari path (jika ada) dan pesan FileNotFoundError(jika tidak ada file .txt)
about()	Render about.html
upload()	Fitur upload file dengan extension yang diperbolehkan (.txt) pada path tertentu
deltxt()	Menghapus semua file .txt pada path

## BAB IV

### EKSPERIMEN

#### 4.1 Webscraping

##### 4.1.1 Percobaan 1 (Ekstrak Raw)

```
PS C:\Users\Tito Prakasa\Documents\Programming\LearnFlask> python .\webscraper.py
```

Dalam dunia kesehatan maupun pendidikan, ada pepatah kuno yang sudah tidak asing lagi di telinga masyarakat. Men sana in corpore sano, sebuah pepatah kuno yang diserap dari bahasa latin ini memiliki makna "jiwa yang sehat dalam tubuh yang sehat". Artinya, pikiran yang sehat dan tenang akan menumbuhkan raga yang sehat pula. Penyakit muncul karena berbagai faktor. Faktor paling umum adalah karena makanan yang tak sehat dan pola hidup yang tidak beraturan. Namun tahukah anda bahwa pikiran yang tidak sehat juga akan menimbulkan berbagai macam penyakit? Ada banyak macam penyakit yang justru lebih parah apabila kita memikirkan hal-hal buruk kedepannya. Pikiran yang buruk akan menyebabkan stamina tubuh mengalami penurunan. Kecemasan yang berlebihan tentunya sangat tidak baik bagi tubuh, apalagi bila sedang dalam proses pemulihan. Berpikir positif dan sehat akan membuat mood naik sehingga menambah stamina tubuh. Hasilnya, seseorang yang sedang melawan penyakitnya akan semakin semangat untuk sembuh dan penuh ambisi untuk bebas dari rasa sakit. Tahun 2020 menjadi tahun yang tidak mudah bagi dunia, terutama Indonesia. Mewabahnya Covid-19 hingga ditetapkan sebagai pandemi global membuat semua sektor kocar-kacir hingga membludaknya pasien di awal-awal masa pandemi. Seluruh tenaga kesehatan sebagai garda terdepan kewalahan menangani banyaknya pasien yang terus bertambah tiap harinya. Kurangnya tempat dan fasilitas kesehatan yang memadai membuat karantina selama 14 hari terasa mencekam dan membosankan. Pikiran-pikiran negatif juga pastinya menghantui para pasien hingga membuat imunitas menurun dan keadaan menjadi drop. Seiring berjalannya waktu, berbagai kabar baik mulai meramban media. Pasien sembuh kian bertambah tiap harinya. Hal ini memicu pemikiran para pasien dan tenaga kesehatan bahwa mereka semua mampu menghadapi dengan baik. Sekali lagi, suasana yang positif, pikiran yang positif akan menimbulkan semangat-semangat baru bagi semuanya. Dalam momentum Hari Kesehatan Nasional ini, mari mulai terapkan pola hidup yang sehat. Mari mulai menebarkan kabar baik bagi lingkungan sekitar. Tentunya dengan tetap mematuhi protokol kesehatan, mari mulai saling menyemangati. Pandemi ini bukan sebuah wabah penyakit yang hanya dengan minum obat apotek sekali teguk langsung sembuh. Semua negara di dunia bahkan belum ada yang pernah menghadapi pandemi ini. Tahun 2020 bukan tahun yang mudah. Selagi menunggu pemerintah dan para ahli kesehatan berlomba-lomba membuat vaksin yang ampuh, mari mulai semuanya dari diri sendiri. Selalu jaga kebersihan, selalu taati protokol kesehatan, dan tentunya mulai terapkan pola hidup yang sehat.

Mengambil test referensi dari website

<https://www.kompasiana.com/qoniatul15454/5faca6ec633122418f314f22/opini-pepatah-kuno-untuk-hari-kesehatan-nasional> gambar di atas merupakan contoh hasil kasar dari scraping suatu webpage berita.

##### 4.1.2 Percobaan 2 (Memisahkan judul, link, dan kalimat pertama pada dokumen)

```
TERMINAL  DEBUG CONSOLE  PROBLEMS  OUTPUT  1: powershell  +  -  ^  x
0.0
42
PS C:\Users\Tito Prakasa\Documents\Programming\LearnFlask\Algeo02-19007> python .\webscraper.py
[{'judul': 'Indonesian worker mauled to death by crocodile in Malaysia', 'link': 'https://www.thejakartapost.com/seasia/2020/11/08/indonesian-worker-mauled-to-death-by-crocodile-in-malaysia.html', 'kalimat': 'An Indonesian migrant worker from Medan, North Sumatra, was found dead by the Sarawak River in Malaysia, the result of a suspected crocodile attack.'}, {'judul': 'Embassy urges Singapore to act upon violence against Indonesian worker', 'link': 'https://www.thejakartapost.com/seasia/2020/11/04/embassy-urges-singapore-to-act-upon-violence-against-indonesian-worker.html', 'kalimat': 'The Indonesian Embassy in Singapore has urged the Singaporean government to act upon a report of alleged violence committed against an Indonesian migrant worker in the city-state.'}, {'judul': 'Ten dead, three missing as 2020's strongest typhoon slams Philippines', 'link': 'https://www.thejakartapost.com/seasia/2020/11/02/ten-dead-three-missing-as-2020s-strongest-typhoon-slams-philippines.html', 'kalimat': 'At least 10 people died and three others were missing after Typhoon Goni, the world's strongest typhoon this year, barreled through the south of the Philippines' main island of Luzon on Sunday, an initial government report showed.'}, {'judul': 'Super typhoon Goni slams into Philippines, makes two landfalls', 'link': 'https://www.thejakartapost.com/seasia/2020/11/01/super-typhoon-goni-slams-into-philippines-makes-two-landfalls.html', 'kalimat': 'A super typhoon barreled into the southern part of the Philippines' main island of Luzon on Sunday, bringing "catastrophic" violent winds and intense rains with two landfalls so far, the weather bureau said.'}, {'judul': 'UMNO calls for general election after pandemic', 'link': 'https://www.thejakartapost.com/seasia/2020/10/30/umno-calls-for-general-election-after-pandemic.html', 'kalimat': 'The biggest party in Malaysia's ruling alliance has called for an election to be held once the coronavirus pandemic is over, as Prime Minister Muhyiddin Yassin grapples with a resurgence in infections and a leadership challenge.'}, {'judul': 'Malaysian king rejects Muhyiddin's proposal for emergency declaration', 'link': 'https://www.thejakartapost.com/seasia/2020/10/26/malaysian-king-rejects-muhyiddin-proposal-for-emergency-declaration.html', 'kalimat': 'The Philippine capital Manila on Friday cancelled an annual procession of a centuries-old black wooden statue of Jesus Christ that draws millions of Roman Catholic devotees as the coronavirus pandemic continues to afflict the country.'}, {'judul': 'Suga, Vietnam's Phuc to focus on economic, defense cooperation', 'link': 'https://www.thejakartapost.com/seasia/2020/10/19/suga-vietnam-phuc-to-focus-on-economic-defense-cooperation.html', 'kalimat': 'Japanese Prime Minister Yoshihide Suga started a meeting Monday with
```

Terjadi perubahan dari dokumen berbahasa indonesia ke berbahasa inggris karena library yang digunakan untuk men-stemming kata khusus untuk bahasa inggris. Pada



percobaan kali ini referensi website yang digunakan adalah <https://www.thejakartapost.com/> dan agar penyimpanan ringkas, setiap dokumen disimpan dalam list of dict dengan index judul, link, dan kalimat.

#### 4.1.3 Percobaan 3 (Mengambil isi dokumen dari link)

```
54
55 print(isiKonten(listMainArticle[0]['link']))
56
57
```



```
, goes on trial in Paris on Tuesday for sexual assault following complaints by five men who accused him of groping them during public ceremonies."}, {'judul': 'Russian peacekeepers head to Nagorno-Karabakh after peace deal', 'link': 'https://www.thejakartapost.com/news/2020/11/10/russian-peacekeepers-head-to-nagorno-karabakh-after-peace-deal-.html', 'kalimat': 'Hundreds of Russian peacekeepers were deploying to Nagorno-Karabakh on Tuesday after Armenia and Azerbaijan agreed to a peace deal to end weeks of fierce fighting over the disputed region.'}]
PS C:\Users\Tito Prakasa\Documents\Programming\LearnFlask\Algeo02-19087> python .\webscraper.py

An Indonesian migrant worker from Medan, North Sumatra, was found dead by the Sarawak River in Malaysia, the result of a suspected crocodile attack.
A plywood factory in Miri reported on Oct. 30 that one of its workers, 22-year-old Muhammad Sopyan Nasution from Medan, had gone missing, according to The Star.
Authorities found his body in a stream near the plywood factory in the Kuala Baram industrial area on Oct. 31.

The local fire and rescue department said that one of Sopyan's friends "saw him being snatched by a crocodile [...] while walking along a water pipeline", The Star reported.
Read also: Fisherman's body retrieved from crocodile's jaws in North Sumatra
An Indonesian agency for migrant workers has confirmed the incident.
"[Sopyan] was playing around the river on his day off. Suddenly, a huge crocodile mauled him to death," said Muhammad Fuad Wahyudi, the spokesperson of the Agency for the Protection of Indonesian Migrant Workers' (BP2MI) North Sumatra office.
```

Dari list of dict di percobaan kedua, kami mengambil sample link pada elemen pertama (index ke-0) lalu dimasukkan pada fungsi isiKonten(link) dan mereturn string seperti di atas.

## 4.2 Vektorisasi dan Cosine Similarity

### 4.2.1 Percobaan Vektorisasi (query : Zara love to hate hate walk)

Dibawah ini adalah dokumen tes

```
testcorpus = [
    {
        "link": "test1.com",
        "content": "poop Sometimes it is better to just walk away from things and go back to them later when you're in a better frame of mind."
    },
    {
        "link": "test2.com",
        "content": "poop He was an introvert that extroverts seemed to love."
    },
    {
        "link": "test3.com",
        "content": "poop He didn't heed the warning and it had turned out surprisingly well."
    },
    {
        "link": "test4.com",
        "content": "poop He hated that he loved what she hated about hate."
    },
    {
        "link": "test5.com",
        "content": "Zara love to hate walk"
    }
]
```

[illegible]

'po': 2.09861228866811, 'didnt': 2.09861228866811, 'well': 2.09861228866811, 'poop': 1.1823215567939547, 'back': 2.09861228866811, 'later': 2.09861228866811, 'love': 1.405  
 55081081644, 'surpris': 2.09861228866811, 'head': 2.09861228866811, 'youre': 2.09861228866811, 'seem': 2.09861228866811, 'extrovert': 2.09861228866811, 'mimd': 2.09861228  
 6811, 'zalk': 2.09861228866811, 'turn': 2.09861228866811, 'warn': 2.09861228866811, 'sometim': 2.09861228866811, 'away': 2.09861228866811, 'tntrovert': 2.09861228866811, 'frame': 2.09861228866811, 'thing': 2.09861228866811, 'hate': 1.6931471805599454, 'walk': 1.6931471805599454  
 'go': 0.0, 'didnt': 0.0, 'well': 0.0, 'poop': 0.0, 'back': 0.0, 'later': 0.0, 'love': 0.2810930216216329, 'surpris': 0.0, 'head': 0.0, 'youre': 0.0, 'seem': 0.0, 'extrovert  
 0.0, 'mimd': 0.0, 'zalk': 0.41972245773362205, 'turn': 0.0, 'warn': 0.0, 'sometim': 0.0, 'away': 0.0, 'tntrovert': 0.0, 'frame': 0.0, 'thing': 0.0, 'hate'  
 : 0.6772588722239782, 'walk': 0.3386294361119891

Menggunakan query yang sama seperti percobaan 4.2.1 , didapat hasil :

```
0.08030843093423182, {'link': 'test3.com', 'content': 'Zara love to hate walk', 'similarity': 0.9457026152328958}, {'link': 'test4.com', 'content': 'poop He hated that he loved what she hated about hate.', 'similarity': 0.7800676252548838}, {'link': 'test2.com', 'content': 'poop He was an introvert that extroverts seemed to love.', 'similarity': 0.10657337100067675}, {'link': 'test1.com', 'content': 'poop Sometimes it is better to just walk away from things and go back to them later when you're in a better frame of mind.', 'similarity': 0.08030843093423182}, {'link': 'test3.com', 'content': 'poop He didn't heed the warning and it had turned out surprisingly well.', 'similarity': 0.0}]
```

9 | Tugas Besar 2 IF2123

## 4.3 Implementasi pada Website

### 4.3.1 Percobaan dengan dokumen hasil web scraping (query: vaccine covid-19)

ya allah Home About

ya allah

vaccine covid-19

☒ Web Scraping ☐ File Local ☐ Web Scraping Table ☐ Local File Table

Sputnik V COVID-19 vaccine trial shows 92 percent efficacy: Producer (similarity : 0.4542300716318373)  
<https://www.thejakartapost.com/news/2020/11/12/sputnik-v-covid-19-vaccine-trial-shows-92-percent-efficacy-pro...>  
The interim trial for the Sputnik V vaccine in Russia has shown encouraging results, demonstrating 92 percent effectiveness at protecting people from COVID-19, according to the vaccine's producer.

Vaccine alliance raises \$2b to buy COVID shots for poor nations (similarity : 0.3361662579918895)  
<https://www.thejakartapost.com/news/2020/11/13/vaccine-alliance-raises-2-b-to-buy-covid-shots-for-poor-natio...>  
A facility set up by the World Health Organization (WHO) and the GAVI vaccine group has exceeded an interim target of raising more than \$2 billion to buy and distribute COVID-19 shots for poorer countries, but said it still needs more.

Philippines cancels 'Black Nazarene' parade as pandemic lingers (similarity : 0.027595943099332314)  
<https://www.thejakartapost.com/seasia/2020/10/23/philippines-cancels-black-nazarene-parade-as-pandemic-ling...>  
The Philippine capital Manila on Friday cancelled an annual procession of a centuries-old black wooden statue of Jesus Christ that draws millions of Roman Catholic devotees as the coronavirus pandemic continues to afflict the country.

Republicans urge Trump to allow Biden briefings as US election challenges sputter (similarity : 0.027013383298256376)  
<https://www.thejakartapost.com/news/2020/11/13/republicans-urge-trump-to-allow-biden-briefings-as-us-electi...>  
<https://www.thejakartapost.com/news/2020/11/12/sputnik-v-covid-19-vaccine-trial-shows-92-percent-efficacy-producer.html> tion should allow Joe Biden to receive

ya allah Home About

ya allah

vaccine covid-19

☐ Web Scraping ☐ File Local ☒ Web Scraping Table ☐ Local File Table

	Q	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14	D15	D16	D17	D18	D19	D20	D21	D22	D23
vaccin	1	13	5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
covid19	1	5	3	1	2	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0

### 4.3.2 Percobaan dengan menggunakan dokumen unggahan (upload)

Akan diunggah 2 dokumen (test01.txt dan test04.txt) , lalu dimasukkan query “Today rainbow shoes”, sehingga didapat hasil :

# ya allah

Today rainbow shoes

☐ Web Scraping ☒ File Local ☐ Web Scraping Table ☐ Local File Table

file2 (similarity : 0.29451291479781705)

[/bt/test04.bt](#)

She didn't understand how changed worked. When she looked at today compared to yesterday, there was nothing that she could see that was different. Yet, when she looked at today compared to last year, ...

file1 (similarity : 0.10201729387804385)

[/bt/test01.bt](#)

The shoes had been there for as long as anyone could remember. In fact, it was difficult for anyone to come up with a date they had first appeared. It had seemed they'd always been there and yet they ...

127.0.0.1:5000/bt/test04.bt

# ya allah

Today rainbow shoes

☐ Web Scraping ☐ File Local ☐ Web Scraping Table ☒ Local File Table

	Q	D1	D2
shoe	1	0	2
rainbow	1	3	0
today	1	2	0

127.0.0.1:5000/bt/test04.bt

## **BAB V**

### **KESIMPULAN, SARAN, DAN REFLEKSI**

#### **5.1 Kesimpulan**

Kami telah berhasil membuat sebuah program berbasis situs web yang dapat menyelesaikan berbagai permasalahan aljabar linier. Berdasarkan hasil percobaan yang telah dilakukan program dapat menerima search query dan mengurutkan dokumen berdasarkan similaritas dengan query, serta menampilkan tabel kemunculan term di query dan di tiap dokumen.

#### **5.2 Saran**

Kami juga menyadari jika program yang dibuat masih memiliki banyak kekurangan dan jauh dari sempurna. Oleh karena itu, saran untuk pengembangan aplikasi ini kedepannya dapat dengan memperindah UI pada frontend sehingga lebih estetik, dan *hosting* aplikasi web sehingga memiliki domain pada Internet sehingga dapat lebih mudah diakses. Selain itu, algoritma *web-scraping* perlu dikembangkan lebih lanjut karena algoritma yang saat ini statis (jika ingin menambahkan referensi dokumen harus menambahkan sintaks secara manual alias tidak efisien).

#### **5.3 Refleksi**

Dengan diberikannya tugas ini, kami dapat lebih lanjut menjelajahi dan mendalami materi-materi terkait vektor khususnya aplikasi dot product terhadap sistem temu balik informasi yang sudah maupun belum dipelajari serta mengeksplor dan membuat situs web dengan *framework* pemrograman yang terkesan 'baru' bagi kami, yaitu Flask serta juga meningkatkan kemampuan memprogram dan bekerja dalam tim.

**Referensi :**

Anton, Howard. (2013). Elementary Linear Algebra 11th Edition.

Prof. DikLee, Vector Space Retrieval Models, Univ. of Science and Tech, Hong Kong.

Salton, G.; Buckley, C. (1988). Term-weighting approaches in automatic text retrieval

Wu, H. C.; Luk, R.W.P.; Wong, K.F.; Kwok, K.L. (2008). Interpreting TF-IDF term weights as making relevance decisions