

BAB III

PELAKSANAAN KERJA MAGANG

3.1 Kedudukan dan Koordinasi

Adapun kedudukan dan koordinasi pada saat kerja magang di PT Rajawali Nusantara Indonesia.

3.1.1 Kedudukan Kerja Magang

Pelaksanaan kerja magang di PT Rajawali Nusantara Indonesia dilaksanakan pada divisi Digital & Teknologi Informasi atau biasa disebut divisi TI sebagai *Data Engineer*. Mahasiswa ditempatkan di ruangan divisi SDM yang bersebelahan dengan ruangan divisi TI, penempatan tersebut dikarenakan pada ruangan divisi TI sudah penuh. Mahasiswa dibimbing oleh pembimbing lapangan kerja magang yaitu bapak Pardomuan Raja Harahap selaku *AVP* dari divisi TI. Tugas dan tanggung jawab dari mahasiswa dalam pelaksanaan kerja magang adalah dapat mengimplementasikan metode *web crawling* dengan menggunakan bahasa pemrograman *Python* dan dapat menganalisis berita pada *Google News*.

3.1.2 Koordinasi Kerja Magang

Selama pelaksanaan kerja magang pada divisi TI di PT Rajawali Nusantara Indonesia, mahasiswa melakukan koordinasi langsung dengan pembimbing lapangan kerja magang yaitu bapak Raja selaku *AVP* divisi TI dan juga melakukan koordinasi langsung dengan karyawan divisi TI yaitu mas Fariz untuk berkonsultasi perihal pengimplementasian metode *web crawling*. Penempatan mahasiswa di ruangan divisi SDM, membuat mahasiswa dapat koordinasi langsung dengan pihak divisi SDM seperti para karyawan dan petinggi divisi tersebut.

3.2 Tugas Kerja Magang

Pelaksanaan kerja magang yang dilakukan oleh mahasiswa dalam waktu kurang lebih 3 bulan, mahasiswa telah melaksanakan tugas dan tanggung jawabnya

sebagai *Data Engineer*. Hasil yang diharapkan dari tugas dan tanggung jawab adalah dapat mengimplementasikan metode *web crawling* pada *Google* menggunakan kata kunci nama perusahaan dengan menggunakan bahasa pemrograman *Python* dan melakukan analisa berita (positif, negatif, atau netral), serta dapat menyimpan hasil *crawling* ke dalam basis data. Selain itu dilakukan percobaan terlebih dahulu untuk melakukan *web crawling* pada portal media daring sebelum implementasi pada web *Google News* sebagai pelatihan dan memastikan mahasiswa dapat melakukan *web crawling*. Daftar tugas dan tanggung jawab dari mahasiswa yang sudah terlaksanakan dapat dilihat pada Tabel 3.1.

Tabel 3.1 Daftar Tugas dan Tanggung Jawab yang Sudah Terlaksanakan

No.	Jenis Pekerjaan		Mulai	Selesai
Pengenalan dan Pembelajaran Metode Web Crawling				
1	1.a	Mempelajari metode <i>web crawling</i> via Youtube.	1 Oktober 2021	5 Oktober 2021
	1.b	Riset <i>library</i> yang diperlukan dalam metode <i>web crawling</i> .	5 Oktober 2021	8 Oktober 2021
Percobaan Metode Web Crawling pada Portal Media Daring				
2	2.a	<i>Crawling</i> portal media daring (CNN Indonesia).	9 Oktober 2021	15 Oktober 2021
	2.b	<i>Crawling</i> portal media daring (Antara News).	16 Oktober 2021	22 Oktober 2021
	2.c	<i>Crawling</i> portal media daring (Bisnis.com).	23 Oktober 2021	29 Oktober 2021
	2.d	<i>Crawling</i> portal media daring (Investor.id).	30 Oktober 2021	5 November 2021
Implementasi Metode Web Crawling pada Google News				
3	3.a	<i>Crawling Google News</i> menggunakan kata kunci RNI dan Rajawali Nusantara Indonesia.	6 November 2021	3 Desember 2021
Klasifikasi Berita (Headlines) Positif, Negatif, atau Netral				
4	4.a	Riset <i>Natural Language Processing (NLP)</i>	4 Desember 2021	6 Desember 2021
	4.b	Implementasi <i>IndoBERT</i> sebagai <i>NLP</i> dalam klasifikasi berita positif, negatif, atau netral.	7 Desember 2021	10 Desember 2021
Penyimpanan Hasil Crawling dan klasifikasi pada Basis Data PostgreSQL				
5	5.a	Riset dan implementasi penyimpanan pada basis data <i>PostgreSQL</i> .	11 Desember 2021	21 Desember 2021
Finalisasi Program Google News Web Crawler dan Demo Program				
6	6.a	<i>Crawling Google News</i> menggunakan kata kunci Rajawali Nusantara Indonesia dan disimpan ke dalam basis data <i>PostgreSQL</i> .	22 Desember 2021	30 Desember 2021
	6.b	Merapikan <i>coding</i> dan memberi komen setiap bagian pada <i>coding</i> .	22 Desember 2021	30 Desember 2021
	6.c	Presentasi dan demo program, serta melampirkan hasil <i>crawling</i> dari tanggal 22 Desember 2021.	30 Desember 2021	31 Desember 2021

Berdasarkan Tabel 3.1, dapat disimpulkan bahwa terdapat 6 tahap yang sudah terlaksanakan dalam mengimplementasikan metode *web crawling*. 6 tahapan tersebut dilaksanakan dari tanggal 1 Oktober hingga 31 Desember, hal tersebut sesuai dengan perjanjian pelaksanaan kerja magang yang sudah disetujui oleh pihak PT RNI dan Universitas Multimedia Nusantara.

3.3 Uraian Kerja Magang

Tugas dan tanggung jawab yang diberikan oleh pihak divisi TI kepada mahasiswa adalah perancangan suatu program dengan bahasa pemrograman *Python* yang dapat melakukan proses *crawling* pada web khususnya portal media daring. Hasil *crawling* dilanjutkan dengan analisa berita, yaitu untuk mengklasifikasikan judul berita apakah positif, negatif, atau netral. Hasil akhir dari proses analisa dilanjutkan dengan menyimpannya ke dalam suatu basis data. Dibutuhkan alat-alat yang dapat mendukung implementasi dari tugas tersebut, mahasiswa menggunakan *Google Collaboratory* sebagai alat dalam perancangan program, *Google Collaboratory* sudah mendukung bahasa pemrograman *Python*. Untuk alat yang digunakan dalam menganalisis berita, digunakan *Natural Language Processing*. Penyimpanan juga dilakukan dengan menggunakan basis data *PostgreSQL*

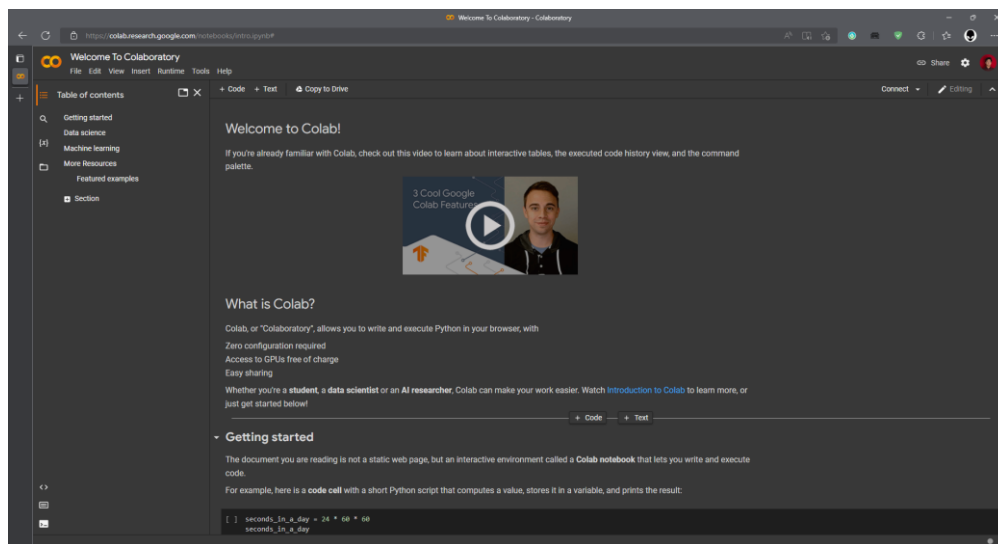
Berdasarkan Tabel 3.1, didapat 6 tahapan yang sudah terlaksanakan dalam pengimplementasian *web crawling*. Tahapan tersebut dimulai dengan tahapan pengenalan dan pembelajaran metode *web crawling* hingga tahapan terakhir yaitu presentasi dan demo program kepada pihak divisi TI khususnya pembimbing lapangan kerja magang yaitu bapak Raja dan karyawan divisi TI mas Fariz. Berikut ini adalah uraian dari pelaksanaan 6 tahapan dalam kerja magang yang sudah dilaksanakan oleh mahasiswa:

3.3.1 Pengenalan dan Pembelajaran Metode Web Crawling

Tahapan pertama yang dilakukan mahasiswa setelah diberikan tugas untuk melakukan implementasi metode *web crawling* adalah dengan melakukan riset dan pembelajaran. Diperlukan riset dalam mengidentifikasi penjelasan mengenai metode *web crawling* dan gambaran umum bagaimana metode tersebut bekerja. Setelah sudah mengetahui apa dan bagaimana

metode *web crawling* bekerja, maka dilakukan pembelajaran dengan membaca buku, riset jurnal, dan menonton video pembelajaran metode *web crawling* via Youtube. Riset dan pembelajaran dilakukan selama kurang lebih 1 minggu lamanya, hal tersebut dilakukan karena mahasiswa belum mengenal dan mengetahui metode *web crawling*. Pada saat masa kuliah mahasiswa pernah belajar mengenai metode *web scrapping* yang mana hampir mirip dengan metode *web crawling* yaitu mengambil data yang berupa informasi pada web, lalu menampilkannya yang mana berupa *output* dari metode tersebut.

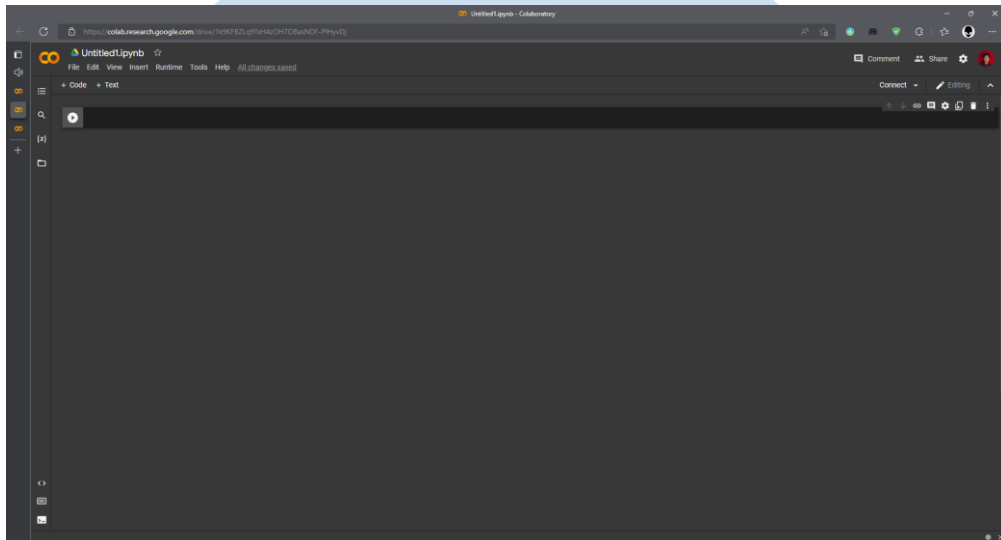
Untuk dapat mengimplementasikan metode *web crawling*, digunakan bahasa pemrograman *Python*. Terdapat banyak sekali perangkat lunak yang menyediakan sarana lingkungan pengembangan terintegrasi untuk bahasa pemrograman *Python*, namun mahasiswa menggunakan *Google Colaboratory* yang menyediakan lingkungan pengembangan berupa *notebook* berbasis *cloud*. Untuk dapat mengetahui seperti apa bentuk dari *Google Colaboratory* atau *Google Colab*, dapat dilihat pada Gambar 3.1.



Gambar 3.1 Halaman Awal dari *Google Colaboratory*

Berdasarkan Gambar 3.1 mengenai halaman awal dari *Google Colab*, untuk dapat memulai suatu proyek baru menggunakan *Google Colab*, maka pengguna dapat menekan *Ribbon* 'File' dan pilih 'New Notebook'. Setelah dipilih maka secara otomatis *Google Colab* akan membuka *new tab* pada

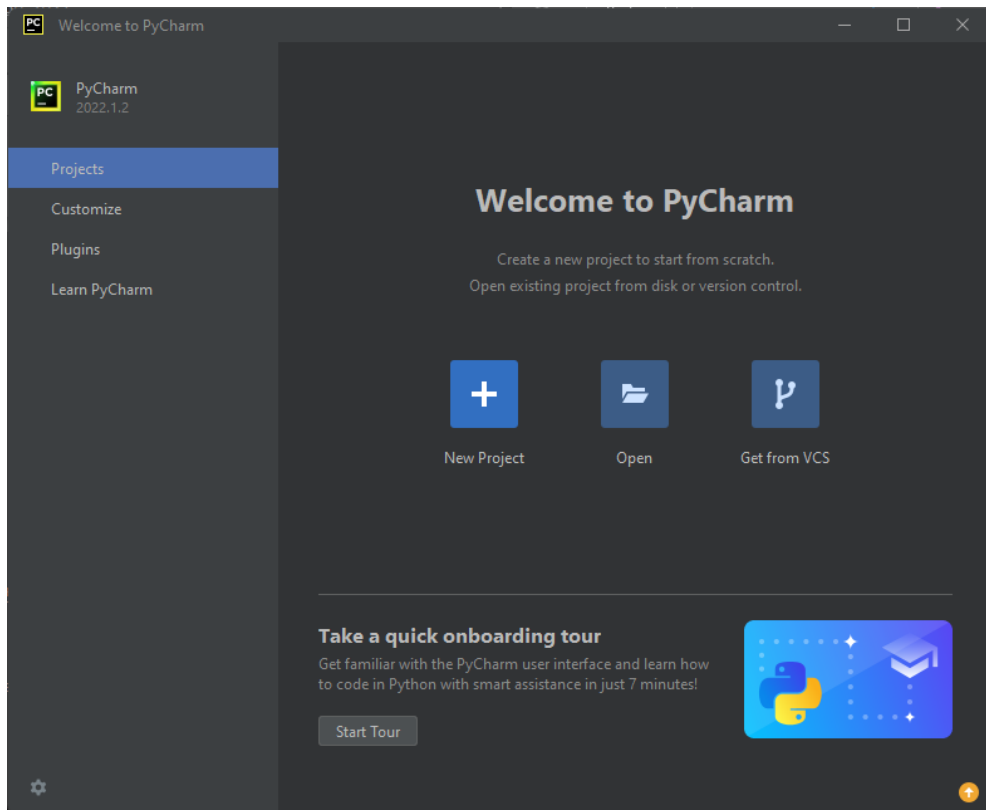
browser dan menampilkan halaman *notebook* baru yang dapat digunakan dalam mengetik dan menyusun *code* bahasa pemrograman. Tampilan dari halaman *notebook* dapat dilihat pada Gambar 3.2 berikut ini.



Gambar 3.2 Halaman *New Notebook* pada *Google Colab*

Pada Gambar 3.2, terdapat halaman kosong yang dapat digunakan sebagai *code editor*, serta dapat menambahkan *text* untuk menulis kalimat yang tidak berhubungan dengan bahasa pemrograman. Selain penggunaan *Google Colab*, mahasiswa juga menggunakan perangkat lunak dalam pengembangan menggunakan bahasa pemrograman yaitu *PyCharm*. Penggunaan *PyCharm* dilakukan untuk dapat memastikan bahwa apabila terdapat kerusakan atau permasalahan dalam penggunaan *Google Colab*, maka terdapat perangkat lunak lainnya sebagai pengganti.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A



Gambar 3.3 Perangkat Lunak *PyCharm*

PyCharm merupakan perangkat lunak yang tidak berbasis *cloud*, sehingga dijalankan pada laptop atau komputer pengguna. Tampilan awal pada perangkat lunak *PyCharm* dapat terlihat pada Gambar 3.3, untuk dapat memulai proyek baru, pengguna dapat menekan ikon dengan kalimat ‘New Project’.

3.3.2 Percobaan Metode Crawling pada Portal Media Daring

Mahasiswa pada awal pengimplementasian metode *web crawling*, ditugaskan untuk mencoba melakukan *web crawling* pada beberapa portal media daring. Maka dari itu dalam pengimplementasian metode *web crawling* portal media dari menggunakan *Google Colab*, dibutuhkan beberapa alat pendukung seperti *library*.

```
from requests_html import HTMLSession
import numpy as np
```

Gambar 3.4 *Library* dalam Implementasi *Web Crawling* pada Portal Media Daring

Berdasarkan Gambar 3.4 mengenai *library* yang diperlukan dalam implementasi *web crawling* pada portal media daring, dapat terlihat dua *library* yang diperlukan. Fungsi dari *library* 'request_html' adalah untuk membantu mahasiswa dalam meminta *Hypertext Transfer Protocol* atau *HTTP* lebih sederhana dan dapat dimengerti oleh manusia. Untuk fungsi dari *library* *numpy* adalah dalam membantu mahasiswa untuk memproses komputasi numerik.

Implementasi metode *web crawling* pada portal media daring dilakukan secara satu persatu pada web portal media, hal tersebut dilakukan untuk melakukan uji coba terlebih dahulu dalam implementasi metode *web crawling*, apabila sebagian besar berhasil mendapatkan hasil *crawling* dari beberapa portal media daring secara satu per satu dengan hasil diinginkan, maka dapat dilanjutkan ke tugas utama yaitu melakukan *web crawling* pada *Google News*. Pada percobaan pertama, mahasiswa melakukan *web crawling* pada sebuah portal media daring yaitu dengan nama portal CNBC Indonesia.

UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA


```
session = HTMLSession()
pages = np.arange(1,12,1)

newslst = []

for page in pages:
    r = session.get("https://www.cnbcindonesia.com/search?query=rni&p=" + str(page) + "&kanal=&type=&date=")
    articles = r.html.find('article')

    for item in articles:
        try:
            newsitem = item.find('a', first=True)
            title = newsitem.text
            link = newsitem.absolute_links
            newsarticle = {
                'title': title,
                'link': link
            }
            newslst.append(newsarticle)
        except:
            pass

print(len(newslst))
for i in newslst:
    print(i)
```

Gambar 3.5 Code dalam Implementasi Web Crawling pada CNBC Indonesia

Pada Gambar 3.5 mengenai proses *web crawling* pada portal media daring CNBC Indonesia, memperlihatkan proses *crawling* menggunakan kata kunci 'rni', hasil akhir yang diinginkan yaitu dapat menampilkan judul dan tautan yang ada pada halaman berita portal media daring CNBC Indonesia. Hasil keluaran dari proses *web crawling* pada CNBC Indonesia dapat dilihat pada Gambar 3.6 berikut ini.

```
121
('title': 'Dua Kawasan Rabana Akan Ditawarkan ke Investor Timur Tengah', 'link': 'https://www.cnbcindonesia.com/news/2021101111640-4-282911/dua-kawasan-rabana-akan-ditawarkan-ke-investor-timur-tengah'}
('title': 'Tekan Subsidif Pupuk BHP Pupuk Indonesia Bilin Program Reformasi', 'link': 'https://www.cnbcindonesia.com/news/2021092021652-17-28077/tekan-subsidi-pupuk-bhp-pupuk-indonesia-bilin-per'}
('title': 'Utak Atik BHM Dowl Holding Pangan, Apa Dampaknya', 'link': 'https://www.cnbcindonesia.com/news/2021092101311-4-278391/utak-atik-bhm-dowl-holding-pangan-apa-dampaknya'}
('title': 'Mendag Klair RI Bisa Ekspor Beras Harga Suka-suka, Kok Bisa?', 'link': 'https://www.cnbcindonesia.com/news/20210915162889-4-276835/mendag-klair-ri-bisa-ekspor-beras-harga-suka-suka-kok-b'}
('title': 'Mendag ke Bos Hiti: Jangan Cuma Aplikasi, Solusi 'Orang' Lain', 'link': 'https://www.cnbcindonesia.com/news/20210915163212-4-276836/mendag-ke-bos-hiti-jangan-cuma-aplikasi-solusi-orang-l'}
('title': 'Soal Produk, Erick Pak Mendag Bukan Sosis yang Semang Impor', 'link': 'https://www.cnbcindonesia.com/news/2021091517758-17-276779/soal-produk-erick-pak-mendag-bukan-sosis-yang-semang'}
('title': 'Masalah Holding BHM Pangan: Cuma 3 Perseroan yang 'Hijau' di Market', 'link': 'https://www.cnbcindonesia.com/news/20210909154915-17-275818/masalah-holding-bhm-pangan-cuma-3-perseroan-yang-hi'}
('title': 'Holding BHM Pangan Minta 'Suntikan' Rp2,1 Bt, Buat Apa? BHM', 'link': 'https://www.cnbcindonesia.com/news/20210909127988-17-269537/holding-bhm-pangan-minta-suntikan-rp2-1-bt-buat-apa'}
('title': 'Maklaha Minta 'Suntikan' Negara Rp 3 T di 2022, Buat Apa? BHM', 'link': 'https://www.cnbcindonesia.com/news/202109081181836-17-273884/maklaha-minta-suntikan-negara-rp-3-t-di-2022-buat-apa'}
('title': 'Erick Thohir Bilcaru Soal Daging & Ayam: Saya Tidak Anti Impor', 'link': 'https://www.cnbcindonesia.com/news/20210822184656-17-270399/erick-thohir-bilcaru-soal-daging-ayam-saya-tidak-an'}
('title': 'Teusyah Kebalkan Hewan untuk BHM Saat Ini? Cak Fatah Ini', 'link': 'https://www.cnbcindonesia.com/news/20210822184656-17-270399/erick-thohir-bilcaru-soal-daging-ayam-saya-tidak-an'}
('title': 'Apa Kabar BHM Holding Pangan? Ini Penjelasan Bos BHM', 'link': 'https://www.cnbcindonesia.com/news/2021080512388-4-266311/apa-kabar-bhm-holding-pangan-ini-penjelasan-bos-hm'}
('title': '09-08-2021 Food & Agriculture Summit 2021/Strategi Membangun Ekosistem Sektor Pangan yang Berdaya Saling', 'link': 'https://www.cnbcindonesia.com/news/20210805105818-8-266279/strategi-membang'}
('title': '15-04-2021 Food & Agriculture Summit 2021/Inilah Kunci Perkuat Daya Saling Sektor Pangan', 'link': 'https://www.cnbcindonesia.com/news/20210805105818-8-266279/strategi-membang'}
('title': '21-55-2021/Inilah Kunci Perkuat Daya Saling Sektor Pangan', 'link': 'https://www.cnbcindonesia.com/news/20210805105818-8-266279/strategi-membang'}
('title': 'Food & Agriculture Summit 2021/Inilah Kunci Perkuat Daya Saling Sektor Pangan', 'link': 'https://www.cnbcindonesia.com/news/20210805105818-8-266279/strategi-membang'}
('title': 'Kamus Lintas Ketahanan Pangan Dalam Food & Agriculture Summit', 'link': 'https://www.cnbcindonesia.com/news/2021080517422-4-265332/kamus-lintas-ketahanan-pangan-dalam-food-agriculture-summit'}
('title': 'Gecep! Erick Sebut PP 3 Holding BHM Rilis Agustus-September', 'link': 'https://www.cnbcindonesia.com/news/20210715158923-17-263493/gecep-erick-sebut-pp-3-holding-bhm-rilis-agustus'}
('title': 'Dilantik dengan PMA, Harga Saham BHM Ijo Ronyo-Ronyo', 'link': 'https://www.cnbcindonesia.com/news/20210715158923-17-263493/gecep-erick-sebut-pp-3-holding-bhm-rilis-agustus'}
('title': 'Erick Uban Hentikan Direksi BNI, Ini Susunan Lengkapnya', 'link': 'https://www.cnbcindonesia.com/news/20210715159221-17-260077/erick-uban-hentikan-direksi-bni-ini-susunan-lengkap'}
('title': 'Ini Hasil Lengkap Rapat Erick Thohir & DPR Soal PMA Rp 100 Triliun', 'link': 'https://www.cnbcindonesia.com/news/20210714183818-17-260818/ini-hasil-lengkap-rapat-erick-thohir-dpr-soal-pma-rp'}
('title': 'Tali Sontok PMA Saat Ini? Ditolak DPR', 'link': 'https://www.cnbcindonesia.com/news/20210714183818-17-260818/ini-hasil-lengkap-rapat-erick-thohir-dpr-soal-pma-rp'}
('title': 'DPR, Erick Sebut 12 BHM Di Sontok PMA Rp 72 Tr, Ini Listnya', 'link': 'https://www.cnbcindonesia.com/news/20210714183818-17-260818/ini-hasil-lengkap-rapat-erick-thohir-dpr-soal-pma-rp'}
('title': 'Economic Update CNBC Indonesia/Komisi Holding BHM Survei, Pangan & Pertahanan Segera Rilis', 'link': 'https://www.cnbcindonesia.com/news/20210714093314-17-260808/economic-update-cnbc-indonesia-komisi-holding-bhm-survei-pangan-pertahanan-segera-rilis'}
('title': 'Sederet Urgensi & Pentingnya Suntikan Modal Dua Bank BHM', 'link': 'https://www.cnbcindonesia.com/news/20210711112538-17-260388/sederet-urgensi-pentingnya-suntikan-modal-dua-bank-bhm'}
('title': 'BHM Megas-Megas Butuh Suntik Modal, Begini Respons Keneikan', 'link': 'https://www.cnbcindonesia.com/news/20210709205214-4-259789/bhm-megas-megas-butuh-suntik-modal-begitu-respons-kanan'}
('title': 'PMA 12 BHM Rp 72 Tr, Erick Fokus ke Pemangan Pemerintah', 'link': 'https://www.cnbcindonesia.com/news/20210708104748-17-259374/pma-12-bhm-rp-72-tr-erick-fokus-ke-pemangan-pemerintah'}
('title': 'Erick Angkat Eke Memendag Raya Krisnamurti Jadi Komis', 'link': 'https://www.cnbcindonesia.com/news/20210708104748-17-259374/pma-12-bhm-rp-72-tr-erick-fokus-ke-pemangan-pemerintah'}
('title': 'DPR Desak Erick Thohir Batalkan BHM Beli Peternakan Belgia', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Ini Jeraan Holding BHM Pangan Beritakan Erick Thohir', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Selangkah Lagi, RI Rumpi Holding BHM Pangan Tahun Ini', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Mentori Erick Percepat Pembentukan Holding BHM Pangan', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Alamat Harga Daging Sapi Masih Liar, Jalang Lebaran', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Erick Bikin Rencana BHM Beli Peternakan Sapi di Belgia', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Erick Bikin Rencana BHM Beli Peternakan Sapi di Belgia', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Gedung Impor Daging Sapi Brasil Mulai Masuk RI, Aman?', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Daging Brasil 'menjadi tabakan' lonjakan harga, Erick', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Daging Sapi Impor Brasil Akhirnya Masuk Indonesia', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Erick Mau Beli Peternakan Sapi di Belgia, BHM Ini Siap-Siap', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Seksi...!!! Risan Ton Gula Impor Mulai Hanyut, Masak Masak', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'BHM Keroyakan di Holding Pangan, Ini Skemarnya', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Erick Bikin Holding BHM Pangan, Perum Periode Jadi Perseroan', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
('title': 'Bulan Depan Sosis Bisa Ditakar, Ini Covid-19, Ini Lokasinya', 'link': 'https://www.cnbcindonesia.com/news/20210619191013-4-246811/dpr-desak-erick-thohir-batalkan-bhm-beli-peternakan-belgia'}
```

Gambar 3.6 Hasil dari Proses Web Crawling pada Portal Media Daring CNBC Indonesia

Berdasarkan Gambar 3.6 mengenai hasil proses *web crawling* pada portal media daring CNBC Indonesia, dapat disimpulkan hasil berupa judul yang berupa kalimat dan tautan. Hasil keluaran dari proses *web crawling* tersebut berjumlah 121 buah berita yang terambil dari halaman berita yang

terdapat kata kunci 'rni' pada portal media daring CNBC Indonesia. Maka dari itu dapat disimpulkan lagi bahwa hasil keluaran dari proses *web crawling* sudah berhasil dan sesuai dengan harapan, namun perlu dilakukan pembersihan dan penghilangan beberapa kata dan simbol. Maka dari itu, proses *web crawling* dilanjutkan ke portal media daring berikutnya yaitu CNN Indonesia.

Percobaan kedua dilakukan proses *web crawling* pada portal media daring CNN Indonesia. Pencarian pada halaman berita menggunakan kata kunci yang sama seperti pada percobaan pertama, yaitu menggunakan kata kunci 'rni'. Berikut ini adalah hasil penyusunan *code* yang digunakan dalam implementasi *web crawling* pada portal media daring CNN Indonesia.

```
session = HTMLSession()
pages = np.arange(1,12,1)

newslist = []

for page in pages:
    r = session.get("https://www.cnnindonesia.com/search?query=rni&p=" + str(page) + "&kanal=&tiipe=&date=")
    articles = r.html.find('article')

    for item in articles:
        try:
            newsitem = item.find('a', first=True)
            title = newsitem.text
            link = newsitem.absolute_links
            newsarticle = {
                'title': title,
                'link': link
            }
            newslist.append(newsarticle)
        except:
            pass

print(len(newslist))
for i in newslist:
    print(i)
```

Gambar 3.7 Code dalam Implementasi *Web Crawling* pada Portal Media Daring CNN Indonesia

Berdasarkan Gambar 3.7 mengenai proses implementasi *web crawling* pada portal media daring CNN Indonesia, memperlihatkan proses *web crawling* menggunakan kata kunci 'rni', hasil akhir yang diharapkan yaitu dapat menampilkan judul dan tautan yang ada pada halaman berita portal media daring CNN Indonesia yang sesuai dengan kata kunci yang digunakan. Hasil keluaran dari proses *web crawling* pada CNBC Indonesia dapat dilihat pada Gambar 3.8 berikut ini.

```
231
{
  'title': 'Sekjen PDIP Hasto Tanya Maksud Anies Undang Tukang Bakso', 'link': {'https://www.cnnindonesia.com/nasi
  'title': 'KPU: 16 Parpol Daftar Sipol, 7 Pendaatang Baru', 'link': {'https://www.cnnindonesia.com/nasional/202206
  'title': 'Buaya dengan Mulut Dilakban Ditemukan Mati di Bogor', 'link': {'https://www.cnnindonesia.com/nasional/
  'title': '9 Saran Eks Wamenlu RI ke Jokowi Damaikan Putin-Zelensky', 'link': {'https://www.cnnindonesia.com/inte
  'title': 'Rusia Kuasai Severodonetsk Sepenuhnya, Serbu Lysychansk Ukraina', 'link': {'https://www.cnnindonesia.c
  'title': '4 Jenis Kapal China yang Disebut-sebut Bisa Bikin AS Kecut', 'link': {'https://www.cnnindonesia.com/in
  'title': 'Beli Gelas Wine Set di Metro, Langsung Diskon Pakai Allo Bank', 'link': {'https://www.cnnindonesia.com
  'title': 'Siang-Siang Jajan Es Krim Wall's Viennetta, Lagi Diskon di Transmart', 'link': {'https://www.cnnindone
  'title': 'Bank DKI Dukung Puncak Jakarta Hajatan Lewat Solusi Digital JakOne Pay', 'link': {'https://www.cnnindo
  'title': 'VIDEO: Kecelakaan Miller di Kualifikasi MotoGP Belanda', 'link': {'https://www.cnnindonesia.com/olahra
  'title': 'VIDEO: Aksi Bagnaia Rebut Pole Position MotoGP Belanda', 'link': {'https://www.cnnindonesia.com/olahra
  'title': 'Cedera, Leo/Daniel Mundur dari Malaysia Open', 'link': {'https://www.cnnindonesia.com/olahraga/2022062
  'title': 'Kenali 5 Tanda WhatsApp Dibajak Dan Cara Pencegahannya', 'link': {'https://www.cnnindonesia.com/teknol
  'title': '5 Cara Mengembalikan File yang Terhapus di Flashdisk', 'link': {'https://www.cnnindonesia.com/teknolog
  'title': 'Cara Hitung Denda Telat Bayar Pajak Kendaraan', 'link': {'https://www.cnnindonesia.com/teknologi/20220
  'title': 'VIDEO: Harry Potter Rayakan 25 Tahun Sejak Rilis Philosopher's Stone', 'link': {'https://www.cnnindone
  'title': 'Marvel Kembali ke Panggung San Diego Comic-Con Bulan Depan', 'link': {'https://www.cnnindonesia.com/hi
  'title': 'Penulis Bocorkan Situasi Deadpool 3 di MCU', 'link': {'https://www.cnnindonesia.com/hiburan/2022062412
  'title': 'Viral Sidik Jari Netizen Rusak Usai Iris Bawang', 'link': {'https://www.cnnindonesia.com/gaya-hidup/20
  'title': '5 Kebiasaan Tidur yang Ampuh Usir Perut Buncit, Selamat Tinggal Lemak!', 'link': {'https://www.cnnindo
  'title': 'Gejala Omicron BA.4 & BA.5 Ringan-Sedang, Apa Saja Ciri-cirinya?', 'link': {'https://www.cnnindonesia.
  'title': 'Sekjen PDIP Hasto Tanya Maksud Anies Undang Tukang Bakso', 'link': {'https://www.cnnindonesia.com/nasi
  'title': 'KPU: 16 Parpol Daftar Sipol, 7 Pendaatang Baru', 'link': {'https://www.cnnindonesia.com/nasional/202206
  'title': 'Buaya dengan Mulut Dilakban Ditemukan Mati di Bogor', 'link': {'https://www.cnnindonesia.com/nasional/
  'title': '9 Saran Eks Wamenlu RI ke Jokowi Damaikan Putin-Zelensky', 'link': {'https://www.cnnindonesia.com/inte
  'title': 'Rusia Kuasai Severodonetsk Sepenuhnya, Serbu Lysychansk Ukraina', 'link': {'https://www.cnnindonesia.c
  'title': '4 Jenis Kapal China yang Disebut-sebut Bisa Bikin AS Kecut', 'link': {'https://www.cnnindonesia.com/in
  'title': 'Beli Gelas Wine Set di Metro, Langsung Diskon Pakai Allo Bank', 'link': {'https://www.cnnindonesia.com
  'title': 'Siang-Siang Jajan Es Krim Wall's Viennetta, Lagi Diskon di Transmart', 'link': {'https://www.cnnindone
  'title': 'Bank DKI Dukung Puncak Jakarta Hajatan Lewat Solusi Digital JakOne Pay', 'link': {'https://www.cnnindo
  'title': 'VIDEO: Kecelakaan Miller di Kualifikasi MotoGP Belanda', 'link': {'https://www.cnnindonesia.com/olahra
  'title': 'VIDEO: Aksi Bagnaia Rebut Pole Position MotoGP Belanda', 'link': {'https://www.cnnindonesia.com/olahra
  'title': 'Cedera, Leo/Daniel Mundur dari Malaysia Open', 'link': {'https://www.cnnindonesia.com/olahraga/2022062
  'title': 'Kenali 5 Tanda WhatsApp Dibajak Dan Cara Pencegahannya', 'link': {'https://www.cnnindonesia.com/teknol
  'title': '5 Cara Mengembalikan File yang Terhapus di Flashdisk', 'link': {'https://www.cnnindonesia.com/teknolog
  'title': 'Cara Hitung Denda Telat Bayar Pajak Kendaraan', 'link': {'https://www.cnnindonesia.com/teknologi/20220
  'title': 'VIDEO: Harry Potter Rayakan 25 Tahun Sejak Rilis Philosopher's Stone', 'link': {'https://www.cnnindone
  'title': 'Marvel Kembali ke Panggung San Diego Comic-Con Bulan Depan', 'link': {'https://www.cnnindonesia.com/hi
  'title': 'Penulis Bocorkan Situasi Deadpool 3 di MCU', 'link': {'https://www.cnnindonesia.com/hiburan/2022062412
  'title': 'Viral Sidik Jari Netizen Rusak Usai Iris Bawang', 'link': {'https://www.cnnindonesia.com/gaya-hidup/20
  'title': '5 Kebiasaan Tidur yang Ampuh Usir Perut Buncit, Selamat Tinggal Lemak!', 'link': {'https://www.cnnindo
  'title': 'Gejala Omicron BA.4 & BA.5 Ringan-Sedang, Apa Saja Ciri-cirinya?', 'link': {'https://www.cnnindonesia.
  'title': 'Sekjen PDIP Hasto Tanya Maksud Anies Undang Tukang Bakso', 'link': {'https://www.cnnindonesia.com/nasi
  'title': 'KPU: 16 Parpol Daftar Sipol, 7 Pendaatang Baru', 'link': {'https://www.cnnindonesia.com/nasional/202206
  'title': 'Buaya dengan Mulut Dilakban Ditemukan Mati di Bogor', 'link': {'https://www.cnnindonesia.com/nasional/
  'title': '9 Saran Eks Wamenlu RI ke Jokowi Damaikan Putin-Zelensky', 'link': {'https://www.cnnindonesia.com/inte
```

Gambar 3.8 Hasil Proses *Web Crawling* pada Portal Media Daring CNN Indonesia

Berdasarkan Gambar 3.8 perihal hasil proses *web crawling* pada portal media daring CNN Indonesia, dapat terlihat hasil keluaran berjumlah 231 dan terdapat judul dan tautan berita, namun dari judul berita tersebut setelah dilakukan evaluasi oleh mahasiswa, ditemukan bahwa berita yang terambil tersebut tidak berkaitan dengan kata kunci 'rni' yang mana seharusnya berkaitan seperti pada percobaan pertama pada portal media daring CNBC Indonesia. Maka dari itu, dapat disimpulkan bahwa dari hasil proses *web crawling* pada portal media daring CNN Indonesia tidak berhasil dikarenakan tidak dapat menghasilkan keluaran yang diharapkan, yaitu judul berita yang terkait dengan kata kunci yang digunakan yaitu 'rni'. Percobaan implementasi *web crawling* dilanjutkan ke percobaan berikutnya, yaitu pada portal media daring Investor ID.

Percobaan ketiga dilakukan implementasi *web crawling* pada portal media daring Investor ID. Pencarian pada halaman berita menggunakan kata kunci yang sama seperti pada percobaan pertama dan ke dua, yaitu menggunakan kata kunci 'rni'. Berikut ini adalah hasil penyusunan *code* yang digunakan dalam implementasi *web crawling* pada portal media daring Investor ID.

```
pages = np.arange(1,6,1)

news_all = []

for page in pages:
    time.sleep(random.randint(1,10))

    url = 'https://investor.id/search/?keyword=rni&page=' + str(page)
    results = requests.get(url)
    soup = BeautifulSoup(results.text, 'html.parser')

    news = soup.find_all('div', class_='listview-content pb10 mb20')

    for container in news:
        headline = container.div.p.a.text
        link = container.div.p.a['href']

        newsarticle = {
            'title': headline,
            'link': link
        }
        news_all.append(newsarticle)

        #summary = container.div.p.text
        #news_summary.append(summary)

print(len(news_all))
for i in news_all:
    print(i)
```

Gambar 3.9 Code dalam Implementasi *Web Crawling* pada Portal Media Daring Investor ID

Berdasarkan Gambar 3.9 mengenai proses *web crawling* pada portal media daring Investor ID, memperlihatkan proses *web crawling* menggunakan kata kunci 'rni', hasil akhir yang diharapkan yaitu dapat menampilkan judul dan tautan yang ada pada halaman berita portal media daring Investor ID yang sesuai dengan kata kunci yang digunakan. Hasil

Percobaan keempat dilakukan implementasi *web crawling* pada portal media daring Antara News. Pencarian pada halaman berita menggunakan kata kunci yang sama seperti pada percobaan pertama dan ke dua, yaitu menggunakan kata kunci 'rni'. Berikut ini adalah hasil penyusunan *code* yang digunakan dalam implementasi *web crawling* pada portal media daring Antara News.

```
pages = np.arange(1,5,1)
news_all = []
for page in pages:
    time.sleep(random.randint(1,10))
    url = 'https://www.antarane.ws.com/search/rajawali+nusantara+indonesia/' + str(page)
    results = requests.get(url)
    soup = BeautifulSoup(results.text, 'html.parser')
    news = soup.find_all('article', class_='simple-post simple-big clearfix')
    for container in news:
        headline = container.header.h3.a.text
        link = container.header.h3.a['href']
        newsarticle = {
            'title': headline,
            'link': link
        }
        news_all.append(newsarticle)
        #summary = container.div.p.text
        #news_summary.append(summary)
print(len(news_all))
for i in news_all:
    print(i)
```

Gambar 3.11 *Code* pada Implementasi *Web Crawling* pada Portal Media Daring Antara News

Berdasarkan Gambar 3.11 mengenai proses *web crawling* pada portal media daring Antara News, memperlihatkan proses *web crawling* menggunakan kata kunci 'rni', hasil akhir yang diharapkan yaitu dapat menampilkan judul dan tautan yang ada pada halaman berita portal media daring Antara News yang sesuai dengan kata kunci yang digunakan. Hasil keluaran dari proses *web crawling* pada Investor ID dapat dilihat pada Gambar 3.12 berikut ini.


```
40
{'title': 'Apresiasi kepada BUMN kembali digelar dalam Indonesia BUMN Awards 2021', 'link': 'https://www.antaranews.com/berita/241116'}
{'title': 'PT RNI ungkap harmonisasi inbreng Holding BUMN Pangan sudah dilakukan', 'link': 'https://www.antaranews.com/berita/24088'}
{'title': 'RNI: Digitalisasi sektor pangan penting turunkan biaya produksi', 'link': 'https://www.antaranews.com/berita/24088'}
{'title': 'PT RNI ajak milenial jadi petani tebu bantu transformasi industri gula', 'link': 'https://www.antaranews.com/berita/24088'}
{'title': 'KAI operasikan kembali KA Sindang Marga rute Kertapati-Lubuklinggau', 'link': 'https://www.antaranews.com/berita/24088'}
{'title': 'Anggota DPR meminta RNI serius tangani konflik pertanahan di Indramayu', 'link': 'https://www.antaranews.com/berita/24088'}
{'title': 'Anak usaha PT RNI berkomitmen jaga kemitraan dengan petani tebu', 'link': 'https://www.antaranews.com/berita/24088'}
{'title': 'PP dan RPP merger BGR ke PPI terbit, direksi sosialisasi ke karyawan', 'link': 'https://www.antaranews.com/berita/24088'}
{'title': 'Terobosan, BUMN klaster pangan ekspor perdana gurita ke AS', 'link': 'https://www.antaranews.com/berita/24088'}
{'title': 'Raih dua penghargaan, bukti RNI sukses berinovasi', 'link': 'https://www.antaranews.com/berita/2412981/raih-dua-penghargaan-bukti-rni-sukses-berinovasi'}
{'title': 'PT KAI Daop 1 turunkan tarif tes antigen di stasiun jadi Rp45 ribu', 'link': 'https://www.antaranews.com/berita/2412981/pt-kai-daop-1-turunkan-tarif-tes-antigen-di-stasiun-jadi-rp45-ribu'}
{'title': 'Tarif tes antigen di tujuh stasiun wilayah Daop Madiun turun', 'link': 'https://www.antaranews.com/berita/2412981/tarif-tes-antigen-di-tujuh-stasiun-wilayah-daop-madiun-turun'}
{'title': 'FEB: Ajang BPEA 2021 untuk bangun kinerja unggul BUMN', 'link': 'https://www.antaranews.com/berita/241116'}
{'title': 'RNI: Penggabungan BUMN pangan perkuat ekosistem pangan', 'link': 'https://www.antaranews.com/berita/24088'}
{'title': 'Peneliti: Holding BUMN pangan harus terbuka dengan kompetisi pasar', 'link': 'https://www.antaranews.com/berita/2397321/mer-erick-apresiasi-warung-pangan-dorong-bumn-dekat-dengan-rakyat'}
{'title': 'Mendag: Ekonomi digital kikis kesenjangan sosial', 'link': 'https://www.antaranews.com/berita/2397321/mer-erick-apresiasi-warung-pangan-dorong-bumn-dekat-dengan-rakyat'}
{'title': 'Erick apresiasi Warung Pangan dorong BUMN dekat dengan rakyat', 'link': 'https://www.antaranews.com/berita/238'}
{'title': 'Proses holding BUMN pangan masuk tahap persiapan merger', 'link': 'https://www.antaranews.com/berita/238'}
{'title': 'Menteri BUMN pastikan pihaknya akan dekat dengan program UMKM', 'link': 'https://www.antaranews.com/berita/238'}
{'title': 'UMKM masuk rantai pasok BUMN dengan nilai kerja sama Rp52,23 miliar', 'link': 'https://www.antaranews.com/berita/238'}
{'title': 'Teten nilai struktur ekonomi yang didominasi usaha mikro perlu diubah', 'link': 'https://www.antaranews.com/berita/238'}
{'title': 'Kementan-RNI bersinergi kembangkan cadangan komoditas hortikultura', 'link': 'https://www.antaranews.com/berita/2344462'}
{'title': 'PT Berdikari hadirkan gerai daging dan produk BeBest', 'link': 'https://www.antaranews.com/berita/2344462'}
{'title': 'RNI: "Holding" BUMN pangan ditargetkan terbentuk September 2021', 'link': 'https://www.antaranews.com/berita/2344462'}
{'title': 'PT Rajawali Nusindo siap distribusikan produk Rania Beras Premium', 'link': 'https://www.antaranews.com/berita/2322798/rni-optimalkan-program-kemitraan-petani-tebu'}
{'title': 'Inkoppas canangkan digitalisasi pasar untuk hadapi revolusi industri', 'link': 'https://www.antaranews.com/berita/2322798/rni-optimalkan-program-kemitraan-petani-tebu'}
{'title': 'RNI optimalkan program kemitraan petani tebu', 'link': 'https://www.antaranews.com/berita/2322798/rni-optimalkan-program-kemitraan-petani-tebu'}
{'title': 'Wamen BUMN teken akta pendirian PT Perikanan Indonesia (Persero)', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'Kemarin, dana Kartu Prakerja hingga angkutan oksigen pakai kereta', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'PT RNI penuhi kebutuhan alkes dan obat dukung penanganan COVID-19', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'Erick Thohir lakukan alih tugas dan perubahan nomenklatur direksi RNI', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'PT RNI produksi alkes oksigen mandiri untuk penuhi kebutuhan oksigen', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'Memastikan jalur distribusi sektor pangan aman selama PPKM darurat', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'Dukung "Kita bersama Nakes", RNI salurkan bantuan ke tenaga kesehatan', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'Erick Thohir usulkan PMN Rp72,44 triliun pada 2022, ini rinciannya', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'Erick Thohir tunjuk Bayu Krisnamurthi jadi Komisaris Utama RNI', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'Kementerian BUMN ganti jajaran direksi Sucofindo', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'PT RNI optimalkan distribusi pangan secara daring selama PPKM Darurat', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'PT RNI bukukan pendapatan hingga Rp6,9 triliun pada 2020', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
{'title': 'Perkuat bisnis ritel, RNI Group resmikan kantor dan gudang di Manado', 'link': 'https://www.antaranews.com/berita/2249662/ken-erick-thohir-lakukan-alih-tugas-dan-perubahan-nomenklatur-direksi-rni'}
```

Gambar 3.12 Hasil Proses Web Crawling pada Portal Media Daring Investor ID

Berdasarkan Gambar 3.12 perihal hasil proses *web crawling* pada portal media daring Investor ID, dapat terlihat hasil keluaran berjumlah 40 dan terdapat judul dan tautan berita, serta berita yang diambil sudah berkaitan dengan kata kunci 'rni' yang mana berkaitan seperti pada percobaan pertama pada portal media daring CNBC Indonesia. Maka dari itu, dapat disimpulkan bahwa dari hasil proses *web crawling* pada portal media daring Investor ID berhasil dikarenakan dapat menghasilkan keluaran yang diharapkan, yaitu judul berita yang terkait dengan kata kunci yang digunakan yaitu 'rni'. Percobaan implementasi *web crawling* dilanjutkan ke percobaan berikutnya, yaitu pada portal media daring Bisnis.com.

Percobaan kelima dilakukan implementasi *web crawling* pada portal media daring Bisnis.com. Pencarian pada halaman berita menggunakan kata kunci yang sama seperti pada percobaan pertama, yaitu menggunakan kata

kunci 'rni'. Berikut ini adalah hasil penyusunan *code* yang digunakan dalam implementasi *web crawling* pada portal media daring Bisnis.com.

```
pages = np.arange(1,28,1)

news_all = []

for page in pages:
    time.sleep(random.randint(1,10))

    url = 'https://search.bisnis.com/?q=rni&per_page=' + str(page)
    results = requests.get(url)
    soup = BeautifulSoup(results.text, 'html.parser')

    news = soup.find_all('div', class_='col-sm-8')

    for container in news:
        headline = container.h2.a.text
        link = container.h2.a['href']

        newsarticle = {
            'title': headline,
            'link': link
        }
        news_all.append(newsarticle)

        #summary = container.div.p.text
        #news_summary.append(summary)

print(len(news_all))
for i in news_all:
    print(i)
```

Gambar 3.13 *Code* pada Implementasi *Web Crawling* pada Portal media Daring Bisnis.com

Berdasarkan Gambar 3.13 mengenai proses *web crawling* pada portal media daring Bisnis.com, memperlihatkan proses *web crawling* menggunakan kata kunci 'rni', hasil akhir yang diharapkan yaitu dapat menampilkan judul dan tautan yang ada pada halaman berita portal media daring Bisnis.com yang sesuai dengan kata kunci yang digunakan. Hasil keluaran dari proses *web crawling* pada Bisnis.com dapat dilihat pada Gambar 3.12 berikut ini.

```
266
{'title': 'Instruksi Erick Thohir Ciptakan Industri Telur Olah Negeri', 'link': 'https://ekonomi.bisnis.c
{'title': 'Dirut RNI Dukung Merger BUMN Klaster Pangan BGR ke PPI', 'link': 'https://ekonomi.bisnis.com/read/2021
{'title': 'Kementerian BUMN Akan Bubarkan 7 BUMN, Ada Merpati Air dan Iglas', 'link': 'https://market.bisnis.com/
{'title': 'Dirut RNI Ungkap Kekuatan Holding BUMN Pangan Pasca Merger', 'link': 'https://ekonomi.bisnis.com/read/
{'title': 'Jokowi Restui Merger Holding BUMN Klaster Pangan', 'link': 'https://ekonomi.bisnis.com/read/20210916/1
{'title': 'BUMN Klaster Pangan Komitmen Terapkan Good Corporate Governance ', 'link': 'https://ekonomi.bisnis.com
{'title': 'Top 5 News Ekonomi: RNI Group Supply Oxygen Generator hingga Perpanjangan PPKM Darurat', 'link': 'http
{'title': 'RNI Group Suplai Oxygen Generator dan Tabung Oksigen untuk RS Darurat', 'link': 'https://ekonomi.bisni
{'title': 'RNI Siap Produksi Alkes Oksigen Mandiri', 'link': 'https://ekonomi.bisnis.com/read/20210711/257/141629
{'title': 'Kementerian BUMN Tetapkan Bayu Krisnamurthi Jadi Komisaris Utama RNI', 'link': 'https://market.bisnis.
{'title': 'Kebun Tebu di Indonesia Terbatas, RNI Bakal Optimalisasi Aset Lahan', 'link': 'https://ekonomi.bisnis.
{'title': 'RNI Mengeklaim Telah Distribusikan Seluruh Gula Kristal Impor', 'link': 'https://ekonomi.bisnis.com/re
{'title': 'Vaksinasi Covid-19, RNI Targetkan Produksi 100 Juta Jarum Suntik', 'link': 'https://ekonomi.bisnis.com
{'title': 'Apa Kabar Rencana Merger BUMN Pangan? Ini Kata Dirut RNI', 'link': 'https://ekonomi.bisnis.com/read/20
{'title': 'Erick Thohir Mau Beli Peternakan Sapi di Belgia, RNI: Terobosan Pangan', 'link': 'https://ekonomi.bisn
{'title': 'Lakukan Revitalisasi Pabrik, RNI Perkuat Industri Gula Nasional', 'link': 'https://ekonomi.bisnis.com/
{'title': 'Amankan Pasokan Lebaran, RNI Impor Gula Kristal Putih 75.000 Ton', 'link': 'https://ekonomi.bisnis.com
{'title': 'Lowongan Kerja BUMN PT RNI: Simak Syarat dan Cara Daftar!', 'link': 'https://kabar24.bisnis.com/read/20
{'title': 'Manfaatkan Sekam, Klaster Pangan BUMN Rambah Bisnis Energi', 'link': 'https://ekonomi.bisnis.com/read/
{'title': 'Klaster Pangan BUMN Siap Jaga Harga Daging', 'link': 'https://ekonomi.bisnis.com/read/20210122/12/1346
{'title': 'Dua Holding BUMN, PTPN III dan RNI Tidak Efektif Berdasarkan Pantauan BPK ', 'link': 'https://market.b
{'title': 'Erick Thohir Rombak Direksi RNI, Mantan Bos BUMD DKI Jadi Dirut', 'link': 'https://market.bisnis.com/r
{'title': 'Erick Thohir Tunjuk Dirut Food Station Tjipinang Jadi Bos PT RNI', 'link': 'https://jakarta.bisnis.com
{'title': 'Pembentukan Holding BUMN Pangan Kembali Dibahas', 'link': 'https://ekonomi.bisnis.com/read/2021112/12
{'title': 'Kuartal III/2020, RNI Bukukan Laba Bersih Rp68 Miliar', 'link': 'https://ekonomi.bisnis.com/read/20201
{'title': 'RNI Catatkan Laba Rp68 Miliar pada Kuartal III/2020', 'link': 'https://surabaya.bisnis.com/read/202011
{'title': 'Sinergi Generasi Milenial dan BUMN', 'link': 'https://kabar24.bisnis.com/read/20201031/79/1311823/sine
{'title': 'Perkuat Ekspansi ke Pasar Ritel, RNI Hadirkan Raja Gula Kemasan Baru', 'link': 'https://ekonomi.bisnis
{'title': 'RNI Lunasi MTN dan Sukuk Rp665 Miliar', 'link': 'https://market.bisnis.com/read/20200929/192/1298090/r
{'title': 'Sempat Merugi 2019, Laba RNI Naik 124 Persen di Semester I/2020', 'link': 'https://ekonomi.bisnis.com/
{'title': 'Mitra Kerinci Ekspor 24,5 Ton Teh Ke Taiwan dan Malaysia', 'link': 'https://ekonomi.bisnis.com/read/20
{'title': 'RNI Dorong Warga di Kaki Gunung Kerinci Menanam Teh', 'link': 'https://ekonomi.bisnis.com/read/2020062
{'title': 'Bersiap New Normal, RNI Gelar Tes Massal untuk Karyawan', 'link': 'https://ekonomi.bisnis.com/read/202
{'title': 'Mendag: Gula Impor Siap Konsumsi Akan Masuk Pertengahan Mei', 'link': 'https://ekonomi.bisnis.com/read
{'title': 'RNI Tunggu Izin Kemenkes untuk Impor Alat Pendeteksi Corona', 'link': 'https://ekonomi.bisnis.com/read
{'title': 'RNI Produksi 2 Juta Masker dan Impor 500.000 Alat Tes Virus Corona', 'link': 'https://market.bisnis.com/
{'title': 'Masker Kain Layak Dipakai Sampai Dua Minggu', 'link': 'https://lifestyle.bisnis.com/read/20200307/106/
{'title': 'Bahan Baku Kosong, RNI Setop Ekspor Masker', 'link': 'https://ekonomi.bisnis.com/read/20200306/98/1210
{'title': 'Dirut RNI: Penyaluran Stok Masker Bakal Diatur Presiden', 'link': 'https://kabar24.bisnis.com/read/202
{'title': 'Genjot Produksi Masker, RNI Ingin Gandeng Jepang ', 'link': 'https://kabar24.bisnis.com/read/20200306/
{'title': 'RNI Salurkan Bantuan ke 6 Kabupaten/Kota Terkena Banjir', 'link': 'https://kabar24.bisnis.com/read/202
{'title': 'Kementerian BUMN Geser Dirut Berdikari ke RNI', 'link': 'https://market.bisnis.com/read/20200104/192/1
{'title': 'Bangun Pabrik Baru, RNI Buka Peluang Ekspor Produk Alkes', 'link': 'https://semarang.bisnis.com/read/2
{'title': 'Kerja Sama dengan Wika di Subang, PTPN VIII Optimalkan Aset Lahan', 'link': 'https://ekonomi.bisnis.co
{'title': 'RNI Pasang Target Produksi Gula 293.000 Ton', 'link': 'https://ekonomi.bisnis.com/read/20190704/99/111
{'title': 'Kawasan Industri Subang, Kemenperin Sebut 3 Sektor Industri', 'link': 'https://ekonomi.bisnis.com/read
{'title': 'Kawasan Industri Subang, HKI : Potensinya Sangat Jelas', 'link': 'https://ekonomi.bisnis.com/read/2019
{'title': 'Kembangkan Kawasan Industri, Wika Bakal Bentuk JV dengan PTPN VIII dan RNI', 'link': 'https://market.b
{'title': 'Lahan Tebu Menyusut, RNI Tunda Revitalisasi Pabrik di Jabar', 'link': 'https://bandung.bisnis.com/read
{'title': 'AP II Realisasikan Kepemilikan Bandara Kertajati, RNI Bidik Aerocity', 'link': 'https://ekonomi.bisnis
{'title': 'RNI Genjot Revitalisasi Pabrik Gula di Jatim, Tunda Jabar', 'link': 'https://ekonomi.bisnis.com/read/2
{'title': 'Kimia Farma (KAEF) Sebut Nilai Akuisisi Phapros Rp1 triliun-Rp1,5 triliun', 'link': 'https://marke
```

Gambar 3.14 Hasil Proses *Web Crawling* pada Portal Media Daring Bisnis.com

Berdasarkan Gambar 3.12 perihal hasil proses *web crawling* pada portal media daring Bisnis.com, dapat terlihat hasil keluaran berjumlah 266 dan terdapat judul dan tautan berita, serta berita yang diambil sudah berkaitan dengan kata kunci ‘rni’ yang mana berkaitan seperti pada percobaan pertama pada portal media daring CNBC Indonesia. Maka dari itu, dapat disimpulkan bahwa dari hasil proses *web crawling* pada portal media daring Bisnis.com berhasil dikarenakan dapat menghasilkan keluaran yang diharapkan, yaitu judul berita yang terkait dengan kata kunci yang digunakan yaitu ‘rni’. Percobaan implementasi *web crawling* dapat terselesaikan, yaitu berhenti pada percobaan kelima yaitu pada percobaan implementasi proses *web crawling* pada portal media daring Bisnis.com.

Setelah dilakukan lima percobaan dalam melakukan implementasi proses *web crawling* pada lima portal media daring, maka dapat disusun sebuah tabel kesimpulan terhadap portal media daring yang dapat dilakukan *crawling* dan portal media daring yang tidak dapat dilakukan *crawling*, berikut ini adalah Tabel 3.1 perihal kesimpulan tersebut.

Tabel 3.2 Hasil Evaluasi dari Proses *Web Crawling* pada lima Portal Media Daring

No.	Portal Media Daring	Hasil Web Crawling Sesuai (Ya/Tidak)	Keterangan
1	CNBC Indonesia	Ya	Dapat dilakukan <i>web crawling</i> dan hasil keluaran sudah sesuai dengan harapan.
2	CNN Indonesia	Tidak	Dapat dilakukan <i>web crawling</i> , namun hasil keluaran tidak sesuai dengan harapan.
3	Investor ID	Ya	Dapat dilakukan <i>web crawling</i> dan hasil keluaran sudah sesuai dengan harapan.
4	Antara News	Ya	Dapat dilakukan <i>web crawling</i> dan hasil keluaran sudah sesuai dengan harapan.
5	Bisnis.com	Ya	Dapat dilakukan <i>web crawling</i> dan hasil keluaran sudah sesuai dengan harapan.

Berdasarkan Tabel 3.2 mengenai hasil evaluasi terhadap proses *web crawling* yang sudah dilaksanakan pada lima portal media daring, dapat disimpulkan bahwa empat dari lima portal media daring mendapatkan hasil keluaran yang sesuai dengan target. Hasil *web crawling* dari portal CNN Indonesia tidak sesuai, dikarenakan pada hasil keluaran yang ada tidak terdapat kata kunci 'rni' pada judul dan *headlines* berita. Pada hasil keluaran *web crawling* dari portal CNN Indonesia menampilkan judul yang melenceng dari kata kunci 'rni'.

Maka dari itu, *code* dalam menjalankan metode *web crawling* dapat dijalankan dengan baik. Pengimplementasian metode *web crawling* dilanjutkan ke tugas berikutnya yaitu tugas utama dari kerja magang.

3.3.3 Implementasi Metode Crawling pada Google News

Keberhasilan implementasi metode *web crawling* pada beberapa portal media daring, membuat mahasiswa dapat melanjutkan ke tugas utama yang diberikan oleh pembimbing lapangan kerja magang yaitu dengan

melakukan implementasi metode *web crawling* pada *Google News*. Seperti halnya pada implementasi *web crawling* pada portal media daring, implementasi pada *Google News* juga diperlukan beberapa *library* dalam mendukung proses *web crawling*.

```
[ ] import requests
    from requests import get
    from bs4 import BeautifulSoup
    import numpy as np
    import time
    import random
```

Gambar 3.15 *Library* pada Implementasi *Web Crawling* pada Portal Media Daring

Berdasarkan Gambar 3.15, dapat terlihat bahwa *library* yang digunakan adalah *requests*, *bs4*, *NumPy*, *time*, dan *random*. Fungsi *library request* adalah untuk dapat meminta *Hypertext Transfer Protocol* atau *HTTP* lebih sederhana dan dapat dimengerti oleh manusia. Fungsi *library bs4* adalah dalam membantu mahasiswa untuk menarik data dari sebuah berkas *Hypertext Markup Language (HTML)* dan *eXtensible Markup Language (XML)*. Untuk fungsi *library NumPy* adalah dalam membantu mahasiswa untuk memproses komputasi numerik. Fungsi *library time* dan *random* adalah membantu mahasiswa dalam menggunakan fungsi yang berkaitan dengan waktu yang tersedia pada *library* tersebut, serta untuk *library random* digunakan dalam menghasilkan angka acak untuk keperluan penelitian. Setelah sudah dapat menentukan dan memasang *library* yang diperlukan, maka dapat dilanjutkan dalam menyusun *code* pengimplementasian metode *web crawling*.

Implementasi metode *web crawling* pada *Google News* dilakukan untuk dapat memenuhi tugas utama dari kerja magang. Implementasi metode *web crawling* yang akan dilakukan pada web *Google News* sedikit berbeda jika dibandingkan dengan dilakukannya pada tiap portal media daring, maka dari itu mahasiswa melakukan riset dan uji coba terlebih dahulu dalam memastikan agar *code* yang digunakan akan dapat dijalankan

dan menghasilkan keluaran yang diharapkan. Uji coba dilakukan selama 2 minggu lamanya, pada akhirnya dapat diselesaikan dan dapat dipastikan penggunaan *code* yang dapat menghasilkan keluaran yang diinginkan adalah pada gambar berikut ini.

```
def jelajah(link):
    upperframe = []
    # Mengambil tanggal hari ini
    tanggal = datetime.datetime.now()
    root = 'https://www.google.com/'
    # Meminta HTML
    req = Request(link, headers={'User-Agent': 'Edge/95.0.1020.40'})
    webpage = urlopen(req).read()
    with requests.Session() as c:
        # Baca HTML5 menggunakan Library BeautifulSoup
        soup = BeautifulSoup(webpage, 'html5lib')
        frame = []
        # Mencari tiap Classes yang mengandung judul dan headline berita, serta link portal beritanya
        for item in soup.find_all('div', attrs={'class': 'Gx5Zad fP1Qef xpd EtOod pkph0e'}): #html secara keseluruhan dari inspect google news
            # Variabel link portal berita
            raw_links = (item.find('a', href=True)['href'])
            links = (raw_links.split('/url?q=')[1]).split('&sa=U&')[0]
            # Variabel judul berita
            titles = (item.find('div', attrs={'class': 'BNeawe vvjw3b AP7Wnd'}).get_text())
            # Variabel nama website dari portal berita
            portal = (item.find('div', attrs={'class': 'BNeawe UPmit AP7Wnd'}).get_text())
            # Variabel headline berita
            desc = (item.find('div', attrs={'class': 'BNeawe s3v9rd AP7Wnd'}).get_text())
            #descs = (descs.split(' ')[1])
            # Menggabungkan seluruh hasil variabel ke dalam bentuk Pandas dataframe
            frame.append((portal, titles, desc, links, tanggal.strftime('%d-%m-%Y')))
        upperframe.extend(frame)
    df_klasifikasi = pd.DataFrame(upperframe, columns=['Portal Berita', 'Judul Berita', 'Deskripsi Berita', 'Tautan', 'Tanggal'])
    # Print dataframe
    df_klasifikasi
    # Print dataframe to CSV
    df_klasifikasi.to_csv('h' + str(tanggal.strftime('%d-%m-%Y')) + '.csv', index=False)
```

Gambar 3.16 Code dalam Implementasi Metode *Web Crawling* pada *Google News*

Berdasarkan Gambar 3.16 mengenai *code* yang digunakan dalam melakukan proses *web crawling* pada web *Google News*, diketahui bahwa penggunaan *library* yang digunakan adalah 'bs4' khususnya '*beautifulsoup*', berbeda dengan *library* yang digunakan dalam *web crawling* pada tiap portal media daring yang menggunakan *library* '*request_html*'. *Code* dari Gambar 3.16 hanya membuat sebuah fungsi yang harus dipanggil jika ingin menjalankan dan melihat keluarannya, maka dari itu berikut ini adalah pemanggilan terhadap fungsi untuk menjelajah (*crawling*).

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A


```

# per minggu
#link = 'https://www.google.com/search?q=rajawali+nusantara+indonesia&tbm=news&source=Int&tbs=qdr:w'

# per hari
#link = 'https://www.google.com/search?q=rajawali+nusantara+indonesia&tbm=news&source=Int&tbs=qdr:d'

# User input berupa katakunci
kata_kunci = input('Masukkan kata kunci yang ingin dijelajah: ')
kata_kunci = str(kata_kunci)
kata_kunci = kata_kunci.replace(' ', '+')
link = 'https://www.google.com/search?q=' + kata_kunci + '&tbm=news&source=Int&tbs=qdr:d'

# Memanggil function jelajah/crawler
jelajah(link)

Masukkan kata kunci yang ingin dijelajah: rni

```

Gambar 3.17 Code dalam Implementasi Metode *Web Crawling* pada *Google News* (Lanjutan)

Pada Gambar 3.17 mengenai *code* lanjutan yang ada pada Gambar 3.16 sebelumnya, dapat diketahui bahwa ketika memanggil fungsi untuk menjalankan proses *crawling* pada *Google News*, pengguna diwajibkan memasukkan kata kunci yang ingin dicari. Pada proses *crawling* saat ini kata kunci yang ingin dicari adalah 'rni'. Berikut ini adalah gambar mengenai hasil dari ketika fungsi *jelajah* dipanggil dan dijalankan.

Portal Berita	Judul Berita	Deskripsi Berita	Tautan	Tanggal
0	Nikkei Asia	Jokowi to urge Putin, Zelenskyy to open room f...	Indonesia has condemned the war and expressed ...	https://asia.nikkei.com/Politics/International... 26-06-2022
1	ANTARA News	Jokowi to discuss investment cooperation with UAE	Jakarta (ANTARA) - Presiden Joko Widodo (Joko...	https://en.antaranews.com/news/236241/jokowi... 26-06-2022
2	The Jakarta Post	President Jokowi to visit Ukraine, Russia on p...	President Joko "Jokowi" Widodo said on Sunday ...	https://www.thejakartapost.com/world/2022/06/2... 26-06-2022
3	South China Morning Post	'War has to be stopped', Jokowi says he will a...	Ukraine war: Indonesia's Jokowi says he will u...	https://www.scmp.com/news/asia/southeast-asia/... 26-06-2022
4	South China Morning Post	Can Jokowi's new trade minister fix Indonesia's...	Can Jokowi's new trade minister Zulkifli Hasan...	https://www.scmp.com/week-asia/opinion/article... 26-06-2022
5	CNBC Indonesia	Melihat Lagi Pengamanan Jokowi Selama di Ukral...	Jakarta, CNBC Indonesia - Hari ini, Minggu (26...	https://www.cnbcindonesia.com/news/20220626135... 26-06-2022
6	Suara.com	Analisis: Tak Terbayangkan Bagaimana Situasi Pan...	Begitu terpilih sebagai Presiden, Jokowi mengg...	https://www.suara.com/bisnis/2022/06/26/171443... 26-06-2022
7	detikNews	Pakar Bicara Rencana Jokowi ke Ukraina Mirip M...	Lawatan Jokowi ke Ukraina ini dinilai mirip se...	https://news.detik.com/berita/d-6147503/pakar... 26-06-2022
8	SINDOnews	Jalankan Politik Bebas Aktif, Jokowi Dinilai M...	JAKARTA - Presiden Joko Widodo (Jokowi) dinila...	https://nasional.sindonews.com/read/809481/12/... 26-06-2022
9	Detikcom	Sapi Korban Presiden Jokowi untuk Sulsel Tahun...	Terus dari Bone sapinya," ungkap dia. Sapi kur...	https://www.detik.com/sulsel/berita/d-6147736/... 26-06-2022

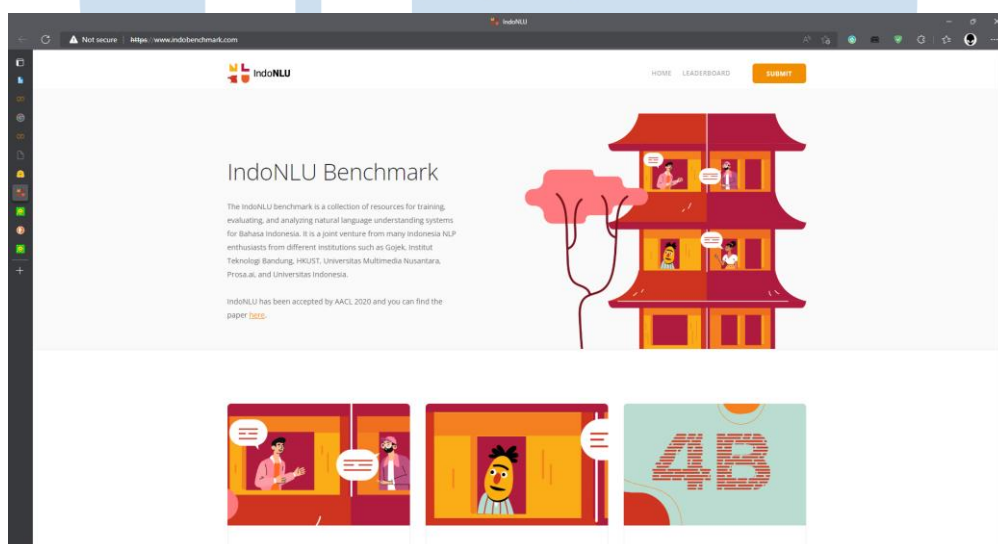
Gambar 3.18 Hasil Proses *Web Crawling* pada *Google News*

Berdasarkan gambar 3.18 mengenai hasil proses *web crawling* pada *Google News*, dapat diketahui terdapat lima kolom yaitu nama portal berita, judul berita, *headline*, tautan dan tanggal berita diambil. Dapat disimpulkan bahwa dari hasil yang ada pada Gambar 3.18 bahwa pengimplementasian proses *web crawling* pada *Google News* telah berjalan dan menghasilkan keluaran yang diinginkan.

3.3.4 Klasifikasi Berita Headlines

Hasil dari implementasi *web crawling* pada web *Google News* sudah didapatkan, tugas berikutnya yaitu mahasiswa ditugaskan untuk melakukan analisa terhadap *headlines* dari berita yang didapatkan, apakah terdapat

makna positif, negatif, atau netral. Maka dari itu, mahasiswa memerlukan *Natural Language Processing* atau biasa disingkat *NLP* dalam memproses klasifikasi terhadap kalimat. *NLP* yang digunakan oleh mahasiswa adalah IndoBERT. IndoBERT di buat oleh IndoNLU yang mana merupakan suatu model yang digunakan dalam menerjemahkan Bahasa Indonesia ke dalam bahasa yang dapat dimengerti oleh komputer atau membuat komputer mengerti dan paham arti dan makna dari kalimat yang berbahasa Indonesia [9]. Pada implementasi yang akan dilakukan oleh mahasiswa, digunakan salah satu model IndoBERT yang tersedia yaitu model ‘indobert-base-p1’.



Gambar 3.19 Halaman Utama Web IndoNLU

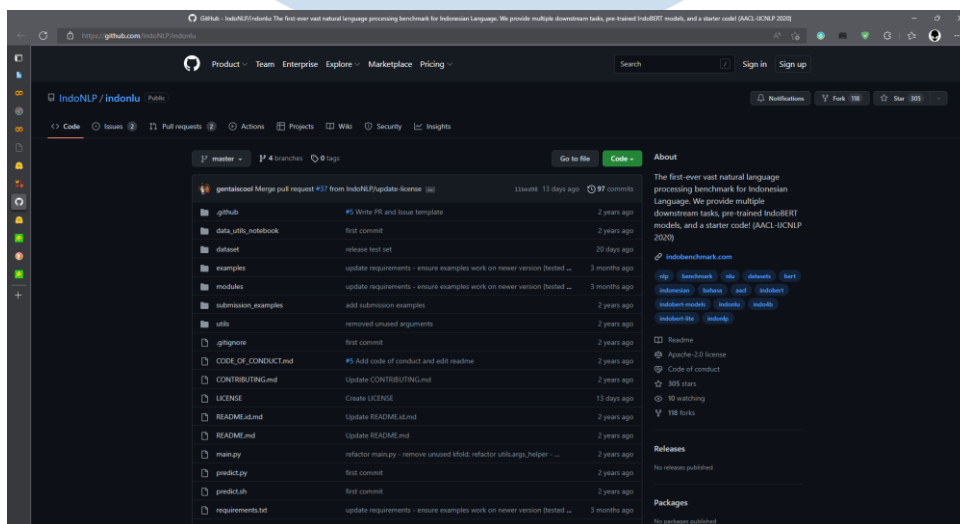
Gambar 3.19 menunjukkan halaman dari web IndoNLU yang mana menyediakan model IndoBERT untuk digunakan bebas tanpa biaya, tentunya dengan menambahkan sitase. Model ‘indobert-base-p1’ dapat dilihat pada gambar berikut ini.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Model	#params	Arch.	Training data
indobenchmark/indobert-base-p1	124.5M	Base	Indo4B (23.43 GB of text)
indobenchmark/indobert-base-p2	124.5M	Base	Indo4B (23.43 GB of text)
indobenchmark/indobert-large-p1	335.2M	Large	Indo4B (23.43 GB of text)
indobenchmark/indobert-large-p2	335.2M	Large	Indo4B (23.43 GB of text)
indobenchmark/indobert-lite-base-p1	11.7M	Base	Indo4B (23.43 GB of text)
indobenchmark/indobert-lite-base-p2	11.7M	Base	Indo4B (23.43 GB of text)
indobenchmark/indobert-lite-large-p1	17.7M	Large	Indo4B (23.43 GB of text)
indobenchmark/indobert-lite-large-p2	17.7M	Large	Indo4B (23.43 GB of text)

Gambar 3.20 Model-Model IndoBERT

Berdasarkan Gambar 3.20 mengenai model-model IndoBERT yang tersedia, dapat terlihat bahwa terdapat 8 model, namun yang akan digunakan oleh mahasiswa yaitu model pertama. Model yang tersedia tentunya diperlukan pelatihan dan data yang mendukung pelatihan tersebut, namun dari IndoNLU sudah menyediakannya. IndoNLU menyediakannya di dalam penyimpanan pada web *GitHub* yang terlihat pada gambar berikut ini.



Gambar 3.21 Halaman Pengunduhan *Dataset* dan *Training Model*

Berdasarkan Gambar 3.21 mengenai halaman pengunduhan *dataset* dan *training model*, terdapat beberapa *file* dan *folder* yang harus diunduh untuk dapat melakukan pelatihan terhadap model IndoBERT. Pelaksanaan pelatihan dimulai dengan melakukan pemasangan *library*, persiapan *dataset*, model pelatihan pada *Google Colab*. Berikut ini adalah pelatihan

dari model IndoBERT, pertama pemasangan *library* 'transformers' yang dapat terlihat pada Gambar 3.22.

```
[ ] !pip install -q transformers
```

	3.4 MB	5.4 MB/s
	3.3 MB	28.3 MB/s
	61 kB	466 kB/s
	596 kB	50.4 MB/s
	895 kB	48.6 MB/s

Gambar 3.22 Pemasangan Transformers Python Library

Setelah dilakukan pemasangan *library*, maka dapat dilanjutkan dengan melakukan *clone* dari penyimpanan atau *repository* IndoNLU pada *GitHub* yang dapat dilihat pada Gambar 3.23

```
!git clone https://github.com/indobenchmark/indonlu
```

```
Cloning into 'indonlu'...
remote: Enumerating objects: 427, done.
remote: Counting objects: 100% (33/33), done.
remote: Compressing objects: 100% (19/19), done.
remote: Total 427 (delta 14), reused 30 (delta 14), pack-reused 394
Receiving objects: 100% (427/427), 4.83 MiB | 13.11 MiB/s, done.
Resolving deltas: 100% (209/209), done.
```

Gambar 3.23 Cloning Repository

Dilanjutkan dengan mengimpor beberapa *library* yang dibutuhkan yaitu seperti *random library*, *NumPy library*, *Pandas library*, *torch library*, *tqdm library*, *nltk.tokenize library*, *indonlu library* dan *os library* yang dapat terlihat pada Gambar 3.24 berikut ini.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

```
[ ] import os, sys
import random
import numpy as np
import pandas as pd
import torch
from torch import optim
import torch.nn.functional as F
from tqdm import tqdm

from transformers import BertForSequenceClassification, BertConfig, BertTokenizer
from nltk.tokenize import TweetTokenizer

from indonlu.utils.forward_fn import forward_sequence_classification
from indonlu.utils.metrics import document_sentiment_metrics_fn
from indonlu.utils.data_utils import DocumentSentimentDataset, DocumentSentimentDataLoader
```

Gambar 3.24 *Import Library* yang Dibutuhkan

Setelah melakukan impor *library* yang diperlukan, maka dilanjutkan dengan membuat fungsi dengan *code* yang dapat terlihat pada Gambar 2.25 berikut ini.

```
def set_seed(seed):
    random.seed(seed)
    np.random.seed(seed)
    torch.manual_seed(seed)
    torch.cuda.manual_seed(seed)

def count_param(module, trainable=False):
    if trainable:
        return sum(p.numel() for p in module.parameters() if p.requires_grad)
    else:
        return sum(p.numel() for p in module.parameters())

def get_lr(optimizer):
    for param_group in optimizer.param_groups:
        return param_group['lr']

def metrics_to_string(metric_dict):
    string_list = []
    for key, value in metric_dict.items():
        string_list.append('{}:{}'.format(key, value))
    return ' '.join(string_list)
```

Gambar 3.25 Beberapa Fungsi yang Dibutuhkan dalam Pelatihan Model

Apabila sudah membuat fungsi yang dapat dipanggil sesuai dengan Gambar 3.25, maka dapat dilanjutkan dengan memuat model IndoBERT dari folder hasil *cloning* yang sudah dilakukan sebelumnya. Dalam proses pemuatan dapat dilihat pada Gambar 3.26 berikut ini.

```
[ ] # Set random seed
set_seed(26092020)

[ ] # Load Tokenizer and Config
tokenizer = BertTokenizer.from_pretrained('indobenchmark/indobert-base-p1')
#tokenizer = BertTokenizer.from_pretrained('indobenchmark/indobert-large-p1')
config = BertConfig.from_pretrained('indobenchmark/indobert-base-p1')
#config = BertConfig.from_pretrained('indobenchmark/indobert-large-p1')
config.num_labels = DocumentSentimentDataset.NUM_LABELS

# Instantiate model
model = BertForSequenceClassification.from_pretrained('indobenchmark/indobert-base-p1', config=config)
#model = BertForSequenceClassification.from_pretrained('indobenchmark/indobert-large-p1', config=config)

Downloading: 100% ██████████ 224k/224k [00:00<00:00, 4.98kB/s]
Downloading: 100% ██████████ 112/112 [00:00<00:00, 631B/s]
Downloading: 100% ██████████ 2.00/2.00 [00:00<00:00, 19.8B/s]
Downloading: 100% ██████████ 1.50k/1.50k [00:00<00:00, 6.41kB/s]
Downloading: 100% ██████████ 475M/475M [00:14<00:00, 37.5MB/s]
Some weights of BertForSequenceClassification were not initialized from the model checkpoint at indobenchmark/indobert-base-p1. You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
```

Gambar 3.26 Load Model pada Folder Clone

Setelah dilakukan pemuatan Model yang akan digunakan, maka dapat dilanjutkan dengan memuat beberapa *dataset* yang akan diperlukan nantinya pada saat pelatihan model. Proses pemuatan *dataset* yang diperlukan dapat dilihat pada Gambar 3.27 berikut ini.

```
[ ] train_dataset_path = 'indonlu/dataset/smsa_doc-sentiment-prosa/train_preprocess.tsv'
valid_dataset_path = 'indonlu/dataset/smsa_doc-sentiment-prosa/valid_preprocess.tsv'
test_dataset_path = 'indonlu/dataset/smsa_doc-sentiment-prosa/test_preprocess_masked_label.tsv'

[ ] train_dataset = DocumentSentimentDataset(train_dataset_path, tokenizer, lowercase=True)
valid_dataset = DocumentSentimentDataset(valid_dataset_path, tokenizer, lowercase=True)
test_dataset = DocumentSentimentDataset(test_dataset_path, tokenizer, lowercase=True)

train_loader = DocumentSentimentDataLoader(dataset=train_dataset, max_seq_len=512, batch_size=32, num_workers=16, shuffle=True)
valid_loader = DocumentSentimentDataLoader(dataset=valid_dataset, max_seq_len=512, batch_size=32, num_workers=16, shuffle=True)
test_loader = DocumentSentimentDataLoader(dataset=test_dataset, max_seq_len=512, batch_size=32, num_workers=16, shuffle=True)

/usr/local/lib/python3.7/dist-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader will create
cpuset_checked))
```

Gambar 3.27 Load Dataset yang Diperlukan dalam Pelatihan Model

Pemuatan pada model dan *dataset* sudah selesai, maka selanjutnya dilakukan persiapan dalam pelatihan model yaitu dengan mengoptimalkan model IndoBERT dengan bantuan teknologi kartu grafis *Nvidia* CUDA, penggunaan bantuan CUDA ini digunakan sesuai dengan *source codes* yang sudah tersedia dan ada pada panduan dalam melakukan pelatihan model IndoBERT yang mana proses pelatihan menggunakan kartu grafis, tidak dengan menggunakan prosesor pada komputer atau laptop. Proses pengoptimalan dapat dilihat pada Gambar 3.28 berikut ini.

```
▶ optimizer = optim.Adam(model.parameters(), lr=3e-6)  
model = model.cuda()
```

Gambar 3.28 Mengoptimalkan Model dengan Menggunakan Bantuan CUDA dari *Nvidia*

Setelah persiapan pelatihan model sudah selesai, maka dilakukan proses pelatihan model IndoBERT dengan menggunakan *code* yang sudah tersedia. Proses pelatihan model yang akan dilakukan dapat dilihat pada Gambar 3.29 berikut ini.




```

# Train
n_epochs = 5
for epoch in range(n_epochs):
    model.train()
    torch.set_grad_enabled(True)

    total_train_loss = 0
    list_hyp, list_label = [], []

    train_pbar = tqdm(train_loader, leave=True, total=len(train_loader))
    for i, batch_data in enumerate(train_pbar):
        # Forward model
        loss, batch_hyp, batch_label = forward_sequence_classification(model, batch_data[:-1], i2w=i2w, device='cuda')

        # Update model
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()

        tr_loss = loss.item()
        total_train_loss = total_train_loss + tr_loss

        # Calculate metrics
        list_hyp += batch_hyp
        list_label += batch_label

        train_pbar.set_description("(Epoch {}) TRAIN LOSS: {:.4f} LR: {:.8f}".format((epoch+1),
            total_train_loss/(i+1), get_lr(optimizer)))

    # Calculate train metric
    metrics = document_sentiment_metrics_fn(list_hyp, list_label)
    print("(Epoch {}) TRAIN LOSS: {:.4f} {} LR: {:.8f}".format((epoch+1),
        total_train_loss/(i+1), metrics_to_string(metrics), get_lr(optimizer)))

    # Evaluate on validation
    model.eval()
    torch.set_grad_enabled(False)

    total_loss, total_correct, total_labels = 0, 0, 0
    list_hyp, list_label = [], []

    pbar = tqdm(valid_loader, leave=True, total=len(valid_loader))
    for i, batch_data in enumerate(pbar):
        batch_seq = batch_data[-1]
        loss, batch_hyp, batch_label = forward_sequence_classification(model, batch_data[:-1], i2w=i2w, device='cuda')

        # Calculate total loss
        valid_loss = loss.item()
        total_loss = total_loss + valid_loss

        # Calculate evaluation metrics
        list_hyp += batch_hyp
        list_label += batch_label
        metrics = document_sentiment_metrics_fn(list_hyp, list_label)

        pbar.set_description("VALID LOSS: {:.4f} {}".format(total_loss/(i+1), metrics_to_string(metrics)))

    metrics = document_sentiment_metrics_fn(list_hyp, list_label)
    print("(Epoch {}) VALID LOSS: {:.4f} {}".format((epoch+1),
        total_loss/(i+1), metrics_to_string(metrics)))

0% | 0/344 [00:00<?, ?it/s] /usr/local/lib/python3.7/dist-packages/torch/utils/data/dataloader.py:481: UserWarning:
cpuset_checked))
(Epoch 1) TRAIN LOSS:0.3518 LR:0.00000300: 100%|██████████| 344/344 [05:06<00:00, 1.12it/s]
(Epoch 1) TRAIN LOSS:0.3518 ACC:0.87 F1:0.81 REC:0.78 PRE:0.85 LR:0.00000300
0% | 0/40 [00:00<?, ?it/s] /usr/local/lib/python3.7/dist-packages/torch/utils/data/dataloader.py:481: UserWarning:
cpuset_checked))
VALID LOSS:0.1918 ACC:0.92 F1:0.89 REC:0.89 PRE:0.89: 100%|██████████| 40/40 [00:14<00:00, 2.69it/s]
(Epoch 1) VALID LOSS:0.1918 ACC:0.92 F1:0.89 REC:0.89 PRE:0.89

```

Gambar 3.29 Pelatihan Model IndoBERT

Proses pelatihan model IndoBERT dilakukan dengan melalui 5 *Epoch*, setelah berhasil melakukan pelatihan maka dapat dilanjutkan dengan proses evaluasi model IndoBERT hasil pelatihan yang dapat dilihat pada Gambar 2.30 berikut ini.

```

# Evaluate on test
model.eval()
torch.set_grad_enabled(False)

total_loss, total_correct, total_labels = 0, 0, 0
list_hyp, list_label = [], []

pbar = tqdm(test_loader, leave=True, total=len(test_loader))
for i, batch_data in enumerate(pbar):
    _, batch_hyp, _ = forward_sequence_classification(model, batch_data[:-1], i2w=i2w, device='cuda')
    list_hyp += batch_hyp

# Save prediction
df = pd.DataFrame({'label':list_hyp}).reset_index()
df.to_csv('pred.txt', index=False)

print(df)

```

```

0%|          | 0/16 [00:00<?, ?it/s] /usr/local/lib/python3.7/dist-packages/torch/utils/data/dataloader
cpuset_checked))
100%|██████████| 16/16 [00:04<00:00, 3.27it/s]   index   label
0         0   negative
1         1   negative
2         2   negative
3         3   negative
4         4   negative
..      ...   ...
495     495   neutral
496     496   neutral
497     497   neutral
498     498   positive
499     499   positive

```

Gambar 3.30 Evaluasi Hasil Pelatihan Model IndoBERT

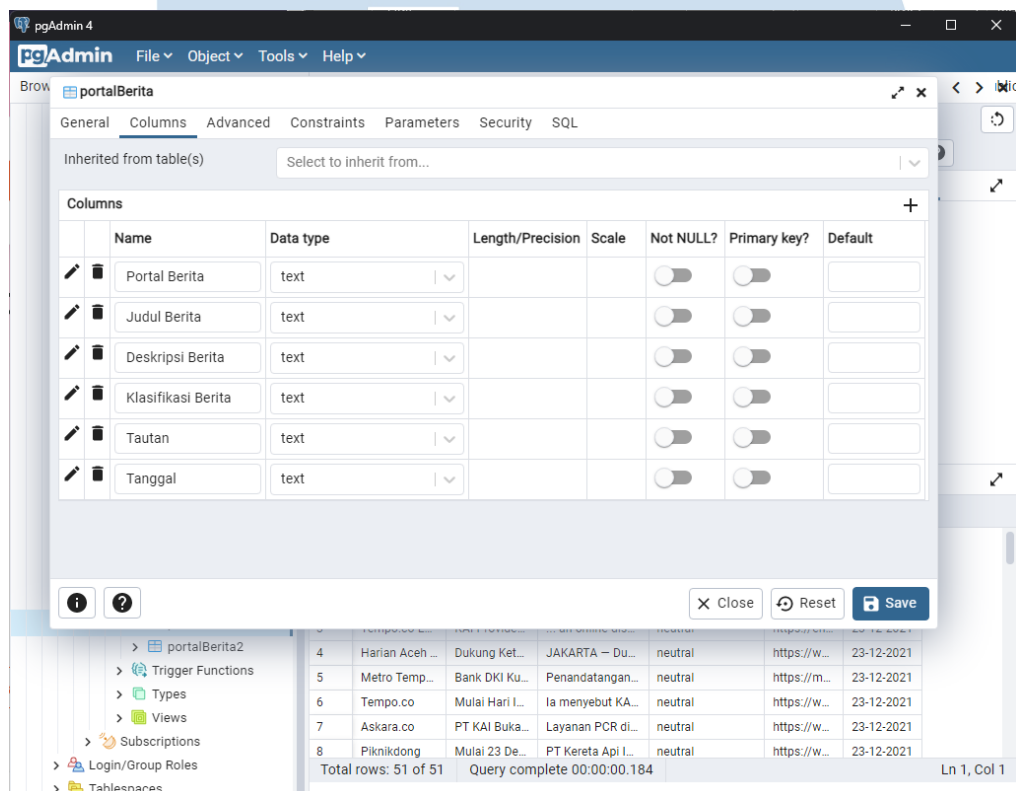
Pada Gambar 3.30 mengenai evaluasi hasil pelatihan model IndoBERT, menunjukkan bahwa model dapat mengklasifikasi menjadi 3 kategori yaitu positif, negatif, atau netral, sehingga klasifikasi dengan model IndoBERT tersebut sesuai dengan target dalam pengimplementasian proses *web crawling* pada *Google News* dengan terdapat analisa berupa klasifikasi *headlines* dari tiap berita yang terambil.

3.3.5 Penyimpanan pada Basis Data PostgreSQL

Hasil dari implementasi proses *web crawling* pada *Google News* sudah didapatkan dan model yang akan digunakan dalam menganalisis berita juga sudah dipersiapkan, namun hasil tersebut diharuskan untuk dapat disimpan sementara sehingga tidak hilang, berdasarkan tugas dari pembimbing lapangan kerja magang, mahasiswa ditugaskan untuk menyimpan hasil keluaran dari *web crawling* pada basis data *PostgreSQL*. Penggunaan basis data *PostgreSQL* dilakukan karena menurut pihak divisi TI lebih baik performanya dan dapat dipelihara dengan mudah. Kemampuan manajemen pada relasional basis data yang dimiliki oleh *PostgreSQL* lebih

baik dari nilai efisiensinya dibandingkan dengan kemampuan sistem manajemen relasional basis data lainnya, seperti *MySQL* [10].

Penyusunan kerangka basis data dilakukan untuk menyesuaikan data yang akan disimpan memiliki struktur yang sesuai dengan struktur tabel penyimpanan data pada basis data. Berikut ini adalah struktur basis data yang sudah dilakukan penyusunan.



Gambar 3.31 Struktur Basis Data PostgreSQL

Berdasarkan Gambar 3.31 mengenai basis data yang akan digunakan dalam penyimpanan hasil *crawling* dan analisa berita, maka dapat disimpulkan bahwa hasil *crawling* dapat disimpan di dalam basis data *PostgreSQL*, namun dilakukan dengan cara manual dan harus *offline*, yaitu tidak dapat dilakukan secara langsung melalui *Google Colab*. Maka dari itu, hasil *crawling* dan analisa berita disimpan ke dalam bentuk *.CSV* dan dilakukan proses penyimpanan menggunakan bantuan perangkat lunak pengganti yaitu *PyCharm*. Berikut ini adalah proses penyimpanan

menggunakan *PyCharm* ke dalam basis data *PostgreSQL* yang ada pada Gambar 3.32.

```
# Function untuk menyimpan hasil crawling ke dalam database PostgreSQL berupa table
def pg_simpan_localhost(df):
    engine = create_engine("postgresql://postgres:password@localhost:5432/webcrawl")

    db = scoped_session(sessionmaker(bind=engine))

    #db.execute("CREATE TABLE IF NOT EXISTS portalBerita();")

    df.to_sql('portalBerita2', engine, if_exists='append', index=False)

    db.commit()

    db.close()
```

Gambar 3.32 Code dalam Proses Penyimpanan pada *PostgreSQL*

Hasil dari proses *web crawling* dan analisa berita pada web *Google News* disimpan dengan menggunakan metode *offline* pada basis data *PostgreSQL* dengan bantuan *PyCharm*. Penyimpanan dilakukan untuk menyimpan hasil sementara sebagai bukti bahwa telah berhasil melakukan proses *web crawling* pada *Google News*.

3.3.6 Finalisasi Program Google News Crawler dan Demo Program

Keberhasilan proses *web crawling* pada portal media daring dan *Google News*, serta berhasilnya analisa berita, maka dapat dilakukan proses evaluasi dan *finishing*. Proses evaluasi dilakukan dengan berdiskusi bersama pembimbing lapangan kerja magang untuk membahas tugas yang sudah mendapatkan hasil dan berhasil. Pada hari akhir yaitu tanggal 31 Desember akan dilakukan presentasi dan demo terhadap proses *web crawling* dan hasil analisa berita yaitu klasifikasi *headlines*. Maka dari itu dari seluruh percobaan yang sudah dilakukan sebelumnya, dapat disimpulkan *code* akhir yang akan digunakan pada saat presentasi dan demo. *Code* tersebut dapat dilihat pada Gambar 3.33 berikut ini.

```

from urllib.request import Request, urlopen
from bs4 import BeautifulSoup
import requests
import pandas as pd
import datetime

def jelajah(link):
    upperframe = []
    # Mengambil tanggal hari ini
    tanggal = datetime.datetime.now()
    root = 'https://www.google.com/'
    # Meminta HTML
    req = Request(link, headers={'User-Agent': 'Edge/95.0.1020.40'})
    webpage = urlopen(req).read()
    with requests.Session() as c:
        # Baca HTML5 menggunakan Library BeautifulSoup
        soup = BeautifulSoup(webpage, 'html5lib')
        frame = []
        # Mencari tiap Classes yang mengandung judul dan headline berita, serta link portal beritanya
        for item in soup.find_all('div', attrs={'class': 'ZINbbc xpd O9g5cc uPjGI'}): #html secara keseluruhan dari inspect google news
            # Variabel link portal berita
            raw_links = (item.find('a', href=True)['href'])
            links = (raw_links.split('/url?q=')[1]).split('&a=U&')[0]
            # Variabel judul berita
            titles = (item.find('div', attrs={'class': 'BNeawe vjwzb AP7Wnd'}).get_text())
            # Variabel nama website dari portal berita
            portal = (item.find('div', attrs={'class': 'BNeawe UPmit AP7Wnd'}).get_text())
            # Variabel headline berita
            desc = (item.find('div', attrs={'class': 'BNeawe s3v9rd AP7Wnd'}).get_text())
            desc = (desc.split('◆')[1])
            # Klasifikasi menggunakan model indobERT
            # Variabel judul berita diklasifikasi
            subwords = tokenizer.encode(titles) #IndoBERT, INI SENTIMENT YANG DINILAI!!!
            subwords = torch.LongTensor(subwords).view(1, -1).to(model.device) #IndoBERT
            logits = model(subwords)[0] #IndoBERT
            label = torch.topk(logits, k=1, dim=-1)[1].squeeze().item() #IndoBERT
            # Variabel hasil klasifikasi per satu judul berita
            klasifikasi = i2w[label] #IndoBERT
            # Menggabungkan seluruh hasil variabel ke dalam bentuk Pandas dataframe
            frame.append((portal, titles, desc, klasifikasi, links, tanggal.strftime('%d-%m-%Y')))
    upperframe.extend(frame)
df_klasifikasi = pd.DataFrame(upperframe, columns=['Portal Berita', 'Judul Berita', 'Deskripsi Berita', 'Klasifikasi Berita', 'Tautan', 'Tanggal'])
# Print dataframe
print(df_klasifikasi)
# Print dataframe to CSV
df_klasifikasi.to_csv('h' + str(tanggal.strftime('%d-%m-%Y')) + '.csv', index=False)

```

Gambar 3.33 Code Proses Web Crawling dan Analisa Berita

Berdasarkan Gambar 3.33 mengenai proses *web crawling* dan analisa berita yang sudah digabung, proses *data cleansing* dilakukan dengan menghilangkan tanda yang ada pada hasil *crawling*, proses penghilangan tersebut dapat terlihat pada *coding* “descs = (descs.split('◆')[1]”. Dalam melakukan klasifikasi, dilakukan dengan memproses variabel ‘titles’ kedalam proses *tokenizer.encode*. Hasil akhir akan berupa *file .CSV* yang nantinya akan dilakukan proses penyimpanan menggunakan *PyCharm* ke dalam basis data *PostgreSQL*. Tugas akhir dari pembimbing lapangan kerja magang yaitu melakukan *crawling* dari tanggal 22 Desember 2021 hingga 31 Desember 2021, menggunakan kata kunci ‘rajawali nusantara Indonesia’ pada web *Google News* dengan kondisi berita yang tersebar yaitu *24-hours ago*.

	Portal Berita text	Judul Berita text	Deskripsi Berita text	Klasifikasi Berita text	Tautan text	Tanggal text
1	Tribun maka...	Lowongan K...	Ada posisi yang...	neutral	https://m...	22-12-2021
2	Pintar Saham	Skydrugz Ra...	Akuisisi Oleh K...	neutral	https://pi...	22-12-2021
3	Tempo.co E...	KAI Provide...	... an online dis...	neutral	https://en...	23-12-2021
4	Harian Aceh ...	Dukung Ket...	JAKARTA – Du...	neutral	https://w...	23-12-2021
5	Metro Temp...	Bank DKI Ku...	Penandatanganan...	neutral	https://m...	23-12-2021
6	Tempo.co	Mulai Hari I...	Ia menyebut KA...	neutral	https://w...	23-12-2021
7	Askara.co	PT KAI Buka...	Layanan PCR di...	neutral	https://w...	23-12-2021
8	Piknikdong	Mulai 23 De...	PT Kereta Api I...	neutral	https://w...	23-12-2021
9	Merdeka.com	Penumpang...	Merdeka.com - ...	neutral	https://w...	23-12-2021
10	Tribun Ponti...	PROFIL BIO...	ID - Satu diantar...	neutral	https://po...	23-12-2021
11	Harian Ekon...	Bank DKI Ku...	NERACA. Jakar...	neutral	https://w...	24-12-2021
12	Majalah Saw...	Menuju Hol...	... PT Rajawali ...	neutral	https://sa...	24-12-2021
13	Warta Kota	Bank DKI Sa...	"Penyaluran kre...	neutral	https://w...	24-12-2021
14	Kominfo Jati...	KAI Siapkan...	Jatim Newsroo...	neutral	http://ko...	24-12-2021
15	BeritaSatu.c...	Dukung Ket...	Berfoto bersam...	neutral	https://w...	24-12-2021
16	NewsSetup	KAI Hadirka...	PT Kereta Api I...	neutral	https://ne...	24-12-2021
17	Medcom.id	Bank DKI Kr...	... ketahanan pa...	neutral	https://w...	24-12-2021
18	Yahoo Berita	Saksikan Ep...	Bank DKI meny...	neutral	https://id...	24-12-2021
19	VOI.id	BUMD Bank...	BUMD Bank DKI...	neutral	https://vo...	25-12-2021
20	VOI.id	Penuhi Keb...	PT Rajawali Nu...	neutral	https://vo...	25-12-2021
21	Pikiran Rakyat	PT RNI Siap ...	"Penyaluran kre...	neutral	https://ja...	25-12-2021
22	Halo Semar...	KAI Berikan ...	HALO SEMARA...	neutral	https://ha...	25-12-2021
23	Finansial Bis...	Bank DKI Sa...	"Penyaluran kre...	neutral	https://fin...	26-12-2021
24	Republika	Anak Usaha...	Anak Perusaha...	neutral	https://re...	26-12-2021
25	Republika	RNI Dorong dengan Pem...	neutral	https://w...	26-12-2021
26	Radarbangsa	KAI Sediaka...	Layanan tes PC...	neutral	https://w...	26-12-2021
27	Majalah Saw...	Potensi Pe...	... Daerah Jawa ...	neutral	https://sa...	27-12-2021
28	Tribun Jateng	Masyarakat ...	Ia menyebutkan...	neutral	https://jat...	27-12-2021
29	Tempo.co	BNI Xpora D...	Rajawali Penta ...	neutral	https://w...	27-12-2021
30	Bisnis Temp...	BNI Xpora D...	Rajawali Penta ...	neutral	https://bi...	27-12-2021
Total rows: 51 of 51		Query complete 00:00:00.184				

Gambar 3.34 Hasil Akhir Web Crawling dan Analisa pada Google News

Hasil akhir harus disimpan pada basis data *PostgreSQL* yang terdapat tanggal pengambilannya. Proses *web crawling* dari tugas akhir tersebut dilakukan per hari. Hasil akhir dari pengerjaan implementasi *web crawling* dan analisa berita pada web *Google News* yang tersimpan pada basis data

PostgreSQL yang dapat terlihat pada Gambar 3.34. Berdasarkan Gambar 3.34, dapat disimpulkan bahwa hasil akhir *web crawling* sudah sesuai target, namun dari hasil analisa berita *headlines* belum sesuai, hal tersebut membuktikan bahwa model IndoBERT yang digunakan masih kurang mendukung untuk mendapatkan hasil yang diinginkan pada proyek kerja magang ini.

3.4 Kendala yang Ditemukan

Adapun kendala yang dihadapi oleh mahasiswa pada saat melaksanakan kerja magang pada divisi TI yang ada pada PT Rajawali Nusantara Indonesia yaitu:

- 1) Tugas yang diberikan oleh pembimbing lapangan kerja magang sudah sesuai dengan profesi dari posisi magang yaitu sebagai *Data Engineer*, namun terdapat beberapa kali perubahan yang berupa *requirements* dalam tugas yang sudah dikasih sebelumnya, *requirements* ini berupa permintaan terhadap variabel-variabel tambahan seperti tanggal pengambilan, nama portal media daring, dan tautan dari berita tersebut. Hal tersebut menimbulkan proses pengerjaan yang kompleks dan memakan waktu yang banyak untuk menyelesaikannya yang mana mahasiswa harus melakukan pengujian dan implementasi lagi untuk mendapatkan variabel yang baru diminta tersebut.
- 2) Proses pemasangan *library* CUDA pada perangkat lunak *Pycharm* yang masih belum berhasil, mahasiswa sudah melakukan percobaan seperti memasang *CUDA Deep Neural Network (CUDNN)*, namun masih belum berhasil, sehingga tidak dapat melakukan pelatihan model IndoBERT pada *Pycharm*.

3.5 Solusi atas Kendala yang Ditemukan

Setelah terdapat kendala-kendala yang sudah tersampaikan, maka adapun solusi atas kendala-kendala yang sudah tertera tersebut, yaitu:

- 1) Solusi terhadap kebutuhan atau *requirements* yang berubah pada tugas yang sudah diterima adalah dengan melakukan diskusi dan perjanjian

bahwa kebutuhan yang ada akan tetap seperti itu hingga hasil akhir sudah didapatkan, namun apabila terdapat kendala dan permasalahan yang kuat dan logis, perubahan dapat terjadi. Perjanjian tersebut harus dapat mengikat pihak mahasiswa dan pihak pembimbing lapangan kerja magang.

- 2) Melakukan riset lebih jauh pada jurnal, web, dan video dan mencoba pemasangan bantuan *tools* lainnya yang terkait dengan penggunaan CUDA pada proses pelatihan model IndoBERT.

