

BAB III

PELAKSANAAN KERJA MAGANG

3.1 Kedudukan dan Koordinasi

Pada hari pertama kegiatan kerja magang di Universitas Multimedia Nusantara di Lab Big Data, bagian *Text Data Mining* merupakan posisi yang diberikan oleh Bapak Yustinus Eko selaku sekretaris prodi system informasi. Selama praktek kerja magang berlangsung, dilakukan berbagai macam rapat yang membahas mengenai koordinasi langsung yang akan dilakukan oleh tim prodi sistem informasi dengan tim prodi ilmu komunikasi mengenai cara berkomunikasi, proses simulasi, *fact checking*, serta penentuan *keyword* yang ada pada debat capres 2019 disaat siaran *live* berlangsung.

3.2 Tugas yang dilakukan

Ada beberapa tugas yang diberikan selama melakukan kegiatan kerja magang di Universitas Multimedia Nusantara, tugas-tugas tersebut adalah:

1. Berkenalan dengan *team fact checking* dari prodi ilmu komunikasi sebagai partner selama proses kegiatan kerja magang berlangsung dan pengenalan bahasa pemograman yang akan digunakan seperti R.
2. Melakukan *convert* pada artikel kompas yang diberikan dalam bentuk *rss xml file* menjadi *csv* menggunakan bahasa Python.

3. Membuat analisis Twitter berdasarkan penggunaan *hashtag*, nama *account* user yang sering melakukan *tweet*, serta analisis kata terbanyak yang sering digunakan oleh *user* di twitter mengenai debat capres yang akan berlangsung.
4. Mencoba melakukan *word count* pada artikel Kompas berdasarkan *keyword* yang ditentukan. Hasilnya akan memunculkan *link* artikel Kompas yang relevan dengan *keyword* beserta tanggal dari artikel tersebut.
5. Melakukan *fact checking* dengan tim ilkom saat debat capres berlangsung secara *live*.

3.3 Uraian Pelaksanaan Kerja Magang

Tabel 1.1 Tabel Uraian Pelaksanaan Kerja Magang

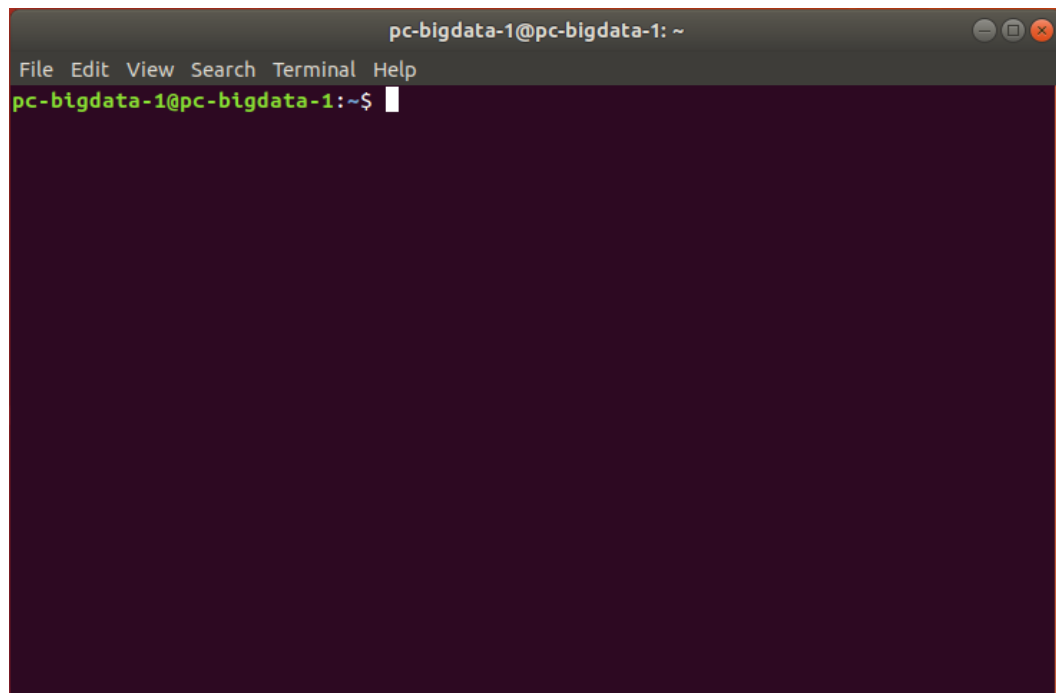
Minggu Ke-	Uraian
1	<ol style="list-style-type: none"> 1. Penkenalan dengan tim Ilkom. 2. Briefing magang. 3. Pengenalan dan beradaptasi dengan bahasa R dan python serta latihan dengan memakai data sampel yang diberikan.
2	<ol style="list-style-type: none"> 1. Membuat <i>script</i> untuk mendownload data <i>rss xml</i> kompas. 2. Membuat <i>script</i> untuk menconvert data <i>rss xml</i> kompas yang sudah di download menjadi data csv.
3	<ol style="list-style-type: none"> 1. Mempelajari <i>script</i> R untuk menganalisis data twitter mengenai debat capres yang diberikan oleh Pak Yustinus Eko. 2. Mejalankan <i>script</i> R yang diberikan dan melakukan analisis terhadap data twitter mengenai debat capres berdasarkan penggunaan <i>hashtag</i>, nama <i>account</i> twitter yang sering melakukan <i>tweet</i> dan kata yang sering

	<p>digunakan oleh user mengenai debat capres tiap harinya.</p> <ol style="list-style-type: none"> 3. Menampilkan <i>wordcloud</i> yang berupa gambar dimana <i>wordcloud</i> tersebut berisikan kata-kata seputar debat capres yang sering digunakan oleh user twitter. 4. Membuat sebuah laporan terhadap analisis yang telah dilakukan.
4	<ol style="list-style-type: none"> 1. Membuat <i>script</i> yang akan digunakan saat debat capres berlangsung. 2. Melakukan rapat dengan tim ilmu komunikasi mengenai debat capres kedua yang akan berlangsung. 3. Melakukan simulasi sebelum debat capres akan berlangsung. 4. Melakukan <i>fact checking</i> bersama tim Ilkom saat debat capres kedua berlangsung secara <i>live</i>.
5	<ol style="list-style-type: none"> 1. Melakukan Evaluasi mengenai kinerja saat debat capres kedua berlangsung bersama tim Ilkom. 2. Menambahkan data agar hasil pencarian link berita untuk debat capres selanjutnya lebih relevan.
6	<ol style="list-style-type: none"> 1. Melakukan simulasi dan mempersiapkan diri bersama tim ilmu komunikasi menjelang debat capres ketigayang akan berlangsung. 2. Melakukan rapat dengan tim ilmu komunikasi mengenai debat capres ketiga yang akan berlangsung. 3. Melakukan <i>fact checking</i> Bersama tim Ilkom saat debat capres ketiga berlangsung secara <i>live</i>.
7	<ol style="list-style-type: none"> 1. Melakukan Evaluasi mengenai kinerja saat debat capres ketiga berlangsung Bersama tim Ilkom. 2. Melakukan rapat dengan tim ilmu komunikasi mengenai debat capres keempat yang akan berlangsung. 3. Melakukan traning kepada anggota magang yang baru untuk melakukan debat cpares keempat yang akan berlangsung.
8	<ol style="list-style-type: none"> 1. Pembuatan laporan dokumentasi

3.3.1 Briefing magang, beradaptasi, pengenalan serta mempelajari data sampel yang diberikan sebelum diberikan data asli yang akan digunakan.

Week1

Briefing magang, beradaptasi, pengenalan serta mempelajari data sampel yang diberikan sebelum diberikan data asli yang akan digunakan. Pada minggu pertama kerja magang, terdapat perkenalan diri di ruang rapat dengan rekan-rekan dari tim Ilkom yang nantinya akan menjadi rekan kerja dan saling berkerjasama selama debat capres berlangsung. Kemudian setelah perkenalan, dilakukan *briefing* magang mengenai tempat dilakukannya kegiatan magang, jam kerja kegiatan magang, jam istirahat, jadwal-jadwal debat capres akan berlangsung, dan pengenalan serta mempelajari bahasa python dan R dengan menggunakan data sampel yang diberikan oleh Pak Yustinus Eko. Selama minggu pertama hanya dilakukan percobaan-percobaan cara untuk menjalankan *script* pada komputer Lab Big Data yang mempunyai system operasi Linux menggunakan Bahasa Python dan R agar lebih terbiasa. Berikut adalah gambar *terminal* pada Linux yang akan digunakan untuk menjalankan *script*.



Gambar 3.1 *Terminal Linux*

Pertama-tama Pak Yustinus Eko memberikan *username* dan *password* untuk mengakses file di server SI. Kemudian Pak Yustinus Eko membimbing untuk memberitahu bagaimana cara mengambil data sampel yang telah dimasukkan Pak Yustinus Eko pada server SI. Lalu menggunakan Linux dapat dilakukan proses download terhadap satu folder yang dipilih dengan perintah “scp-r”.

Pada terminal *Linux* akan dilakukan eksekusi *script* dengan cara mencari letak *script* berada menggunakan perintah change directory yang akan mengarahkan *directory* contohnya seperti “cd Documents”, “cd Downloads”, “cd File”, dan sebagainya. Lalu *script* akan dieksekusi menggunakan python ataupun R dengan memakai perintah “python namascript.py” untuk python dan “Rscript namascript.r”

untuk R. Contohnya seperti “python kompas.py”. Kemudian untuk keluar dapat menggunakan perintah “quit()”.

3.3.2 Mendownload dan menconvert artikel Kompas dari *rss xml* file menjadi csv week 2

Pada minggu kedua, diberikan tugas untuk mendownload rss file yang masih berupa link seperti

[http://rss.kompas.com/umn/\(0sampai667\)?apikey=3f2f8f5be5d853b7f1c776d4134be02b3cbf2df0](http://rss.kompas.com/umn/(0sampai667)?apikey=3f2f8f5be5d853b7f1c776d4134be02b3cbf2df0). dari Kompas yang nantinya akan digunakan sebagai file pendukung untuk melakukan fact checking pada debat capes yang aan berlangsung. Proses download dibuat secara otomatis menggunakan *script* yang bernama “websucker.py”



```
websucker.py
~/Desktop/dataArtikel

import os

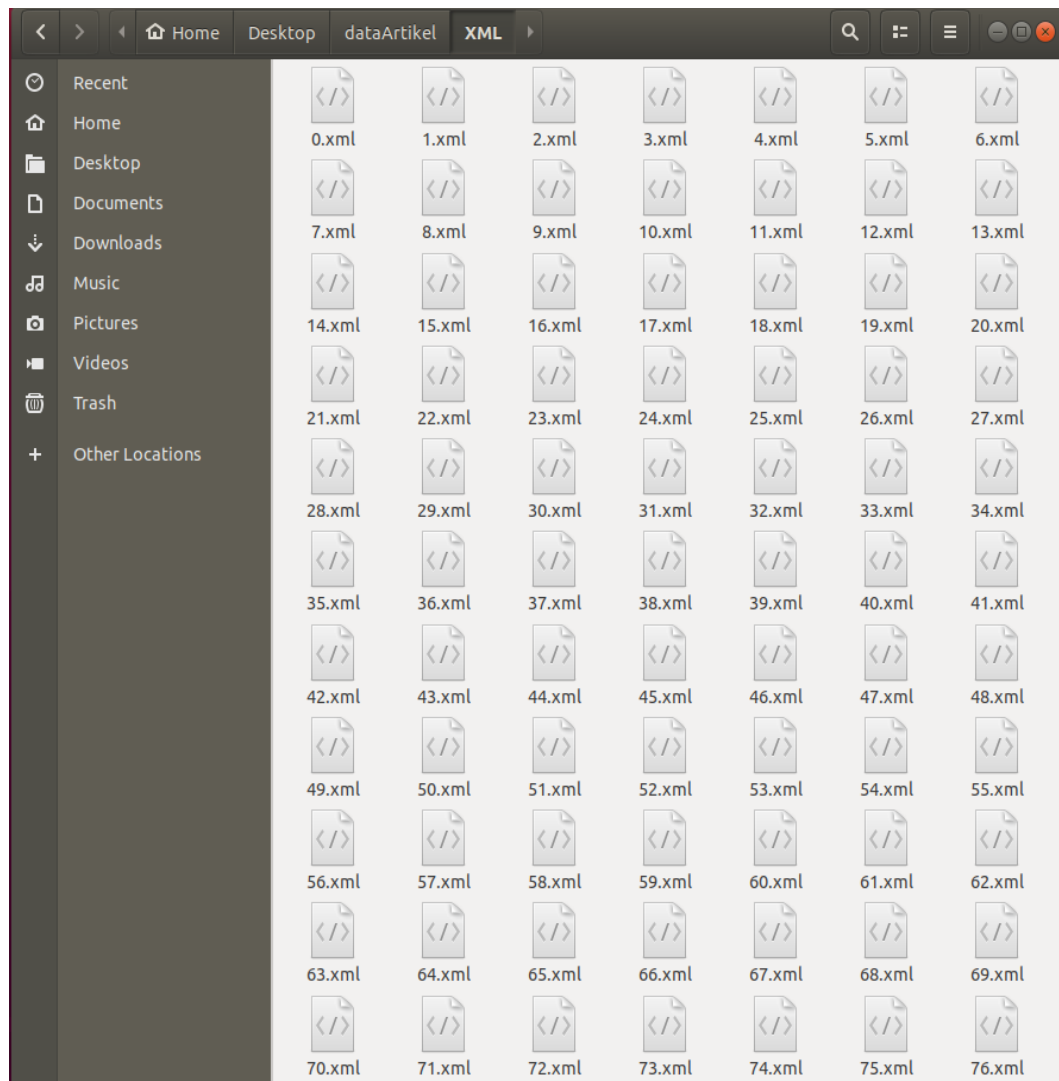
x = 0

for x in range(0, 666):
    link = 'http://rss.kompas.com/umn/' + str(x) + '?
apikey=3f2f8f5be5d853b7f1c776d4134be02b3cbf2df0'
    os.system('wget --output-document=' + str(x) + '.xml ' + link)
    x += 1

Python Tab Width: 8 Ln 3, Col 6 INS
```

Gambar 3.2 websucker.py

Gambar 2.2 merupakan isi dari *script* websucker.py, dimana perintah wget akan melakukan proses *download* link rss Kompas dan menyimpan filenya dalam bentuk *xml* secara otomatis dan melakukannya secara *looping* dari 0 sampai 666.



Gambar 3.3 Hasil dari websucker.py

Gambar 3.3 merupakan hasil dari websucker.py dimana *link rss* kompas telah didownload dalam bentuk file xml



```
<?xml version="1.0" encoding="UTF-8" ?>
<rss xmlns:content="http://purl.org/rss/1.0/modules/content/" version="2.0">
  <channel>
    <title>
      <![CDATA[ Kompas.com - Jernih Melihat Dunia ]]>
    </title>
    <link>
      <![CDATA[ http://www.kompas.com ]]>
    </link>
    <description>
      <![CDATA[
        Kompas.com is Indonesia leading & most credible news and multimedia portal. Previously known as Kompas Cyber Media or Kompas Online. Established in 1995, Kompas.com was reborn in May 2008 with improved contents and architecture. Kompas.com is an affiliate of Kompas Gramedia Digital Group
      ]]>
    </description>
    <pubDate>
      <![CDATA[ Selasa, 29 Januari 2019 ]]>
    </pubDate>
    <totalPage>
      <![CDATA[ 18445 ]]>
    </totalPage>
    <totalPage>
      <![CDATA[ 18445 ]]>
    </totalPage>
    <item>
      <guid isPermaLink="true">
        <![CDATA[ 649c5ab981f3835686308918094ee9ce ]]>
      </guid>
      <title>
        <![CDATA[ Mari, Mengurangi Depresi dengan Olahraga ]]>
      </title>
      <link>
        <![CDATA[
          https://lifestyle.kompas.com/read/2019/01/29/092228129/mari-mengurangi-depresi-dengan-olahraga
        ]]>
      </link>
      <photo>
        <![CDATA[
          https://asset.kompas.com/crop/0x8:1800x667/756x508/data/photo/2018/02/06/638399152.jpg
        ]]>
      </photo>
      <thumb>
        <![CDATA[
          https://asset.kompas.com/crop/0x8:1800x667/195x98/data/photo/2018/02/06/638399152.jpg
        ]]>
      </thumb>
      <description>
        <![CDATA[
          Apakah olahraga memang betul-betul bisa mencegah depresi? Sebuah studi yang dipublikasikan di JAMA Psychiatry menjelaskan tentang hal ini.
        ]]>
      </description>
      <content>
        <![CDATA[
          <strong>Cobalah tanya kepada orang yang meyakini perasaannya lebih baik setelah berolahraga.</strong> Mereka tentu akan mengatakan, olahraga dan kesehatan mental memiliki keterkaitan erat.</p>
          <p>Kesimpulan ini pun rupanya bisa dijelaskan secara ilmiah.</p>
          <p>Banyak penelitian menemukan, aktivitas fisik berkorelasi terhadap penurunan risiko berkembangnya depresi dan memberikan efek positif kepada pelakunya.</p>
          <p>Namun, apakah olahraga memang betul-betul bisa mencegah depresi? Sebuah studi yang dipublikasikan di JAMA Psychiatry menjelaskan tentang hal ini.</p>
          <p>Penelitian tersebut menggunakan data genetik dari 688.888 lebih orang dewasa yang terdaftar dari beberapa genom terkait.</p>
          <p>Para peneliti pun menemukan lebih banyak bukti tentang keterkaitan aktivitas fisik dan penurunan risiko depresi.</p>
          <strong>Baca juga:</strong>
          <a href="http://lifestyle.kompas.com/read/2019/12/17/074549228/nonton-pertunjukan-budaya-ternyata-ampuh-cegah-depresi" target="self">Nonton Pertunjukan Budaya Ternyata Ampuh Cegah Depresi</a></strong>
          <p>Hal itu diungkapkan oleh Karmel Choi, penulis studi sekaligus mahasiswa doktoral di bidang <em>psychiatric and neurodevelopmental genetics</em> di Massachusetts General Hospital.</p>
          <p>Penelitian tersebut meneliti beberapa tolok ukur lain, yaitu gen manusia, sejarah depresi seseorang, dan gejala-gejala depresi yang dialami, serta jumlah aktivitas fisik yang dilakukan.</p>
          <p>Dengan membandingkan data-data tersebut, para peneliti mengidentifikasi adanya beberapa varian gen yang terkait dengan kecenderungan seseorang untuk berolahraga, dan lainnya terkait dengan kecenderungan perkembangan depresi.</p>
          <p>Partisipan yang memiliki kecenderungan gen dengan olahraga cenderung tidak mudah membangun depresi.</p>
          <p>Sebaliknya, orang-orang yang memiliki kecenderungan gen dengan depresi cenderung jarang berolahraga.</p>
          <p>Menurut mereka, penemuan ini menemukan olahraga bisa mencegah kondisi itu. Namun, depresi tidak memiliki keterkaitan atau membuat seseorang menjadi lebih sedikit berolahraga.</p>
          <p>Aktivitas fisik mempunyai banyak manfaat. Tidak hanya untuk aspek kesehatan tapi juga menurunkan risiko perkembangan depresi.</p>
          <strong>Baca juga:</strong>
          <a href="http://lifestyle.kompas.com/read/2019/01/09/094958828/memurut-riset-orang-baik-mudah-depresi" target="self">Menurut Riset, Orang Baik Mudah Depresi</a>
        ]]>
      </content>
    </item>
  </channel>
</rss>
```

Gambar 3.4 Isi dalam 0.xml

Gambar 3.4 merupakan isi dari file 0.xml yang sudah didownload menggunakan websucker.py. Setelah file didownload dalam bentuk xml, file ini akan diconvert dalam bentuk csv, sehingga lebih mudah untuk dibaca. Proses convert file xml ke dalam bentuk csv dilakukan menggunakan *script* python yang bernama “convert.py”



```
Open [icon] convert.py ~/Desktop/dataArtikel Save [icon] [icon] [icon]
import xml.etree.ElementTree as ET
import csv
from xml2csv import xml2csv as x

e = 0

for e in range(0, 666):
    input_path= str(e) + '.xml'
    output_path='Artikel' + str(e) + '.csv'
    data=x(input_path,output_path,encoding='utf-8')
    data.convert(tag="item")
    data.output.close()
    e += 1

Python Tab Width: 8 Ln 9, Col 25 INS
```

Gambar 3.5 convert.py

Gambar 3.5 merupakan isi dari convert.py, dimana isi tag item pada data xml akan diambil dan dikeluarkan dalam bentuk csv dengan nama “Artikel(0-666).csv”. Tag item pada data xml yang diambil terdiri dari *id*, *photo/gambar*, *date/tanggal*, *title/judul*, *description*, *site_name*, *author*, *link*, *editor*, *content*(isiberita), *thumb*, dan *section_name*.

	A	B	C	D	E	F	
1	guid	title	link	photo	thumb	description	content
2	649c5ab9	Mari Mengurangi	https://lifestyle	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x0:1000	Apakah olahraga memang betul-betul bisa	<p><!--img 1-->KOMPAS.com <!--strong> Cobalah tanya
3	697cb82b	Mitsubishi Mulai	https://otomotif	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x0:1000	Ini dilakukan setelah waktu pemenuhan kor	<p class="p1"><!--img 1-->
4	1015c4e4	5 Fakta Bencana	https://regional	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x156:1000	Jumlah korban meninggal akibat bencana	<p><!--img 1--><!--img 1-->KOMPAS.co
5	0e589fce	Peluncuran Indi	https://nasional	https://asset.kompas.com/crop/7	https://asset.kompas.com/crop/78x0:924	Berbeda dari tahun-tahun sebelumnya, peli	<p><!--img 1-->JAKARTA, KOMPAS.com - Tra
6	4ec84723	Samsung Galaxy	https://teknokom	https://asset.kompas.com/crop/1	https://asset.kompas.com/crop/17x0:719	Linji smartphone teranyar Samsung Galaxy	<p><!--img 1-->KOMPAS.com - Lini baru Sam
7	084f6a78	Atletico Klub Imp	https://bola.komp	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x34:988	Saat masih kecil, Morata sempat menjadi	<p><!--img 1-->KOMPAS.com - Pemain baru A
8	5219f97c	KPK Optimis	https://nasional	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x0:997x0	Kita harus optimis, enggak boleh skepti	<p><!--img 1-->JAKARTA, KOMPAS.com - Ko
9	6e84aa7	DKI Data dan Iku	https://megapol	https://asset.kompas.com/crop/3	https://asset.kompas.com/crop/31x40:972	Saat ini pendafar yang berminat mengikut	<p><!--img 1-->JAKARTA, KOMPAS.com <!--strong> Pe
10	096ace4c	Catatan Program	https://ekonomi	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x0:1000	Sebanyak 89 persen e-warong yang merup	<p><!--img 1-->JAKARTA, KOMPAS.com - Mi
11	6fad8c7	Aturan Baru, Pel	https://otomotif	https://asset.kompas.com/crop/3	https://asset.kompas.com/crop/30x0:705	Peraturan baru itu sedang disiapkan untu	<p><!--img 1--><!--img 1-->JAKARTA, K
12	686eeca	Resmi ke Atletico	https://bola.komp	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x34:988	Atletico memang akan menamu Madrid di	<p><!--img 1--><!--img 1-->KOMPAS.com - Alvar
13	2ab05d5b	Komentar Gerin	https://nasional	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x0:1000	Ahmad Dhani divonis 1,5 tahun dan langsu	<p><!--img 1-->JAKARTA, KOMPAS.com - Ke
14	bb021685	Ahmad Dhani Dir	https://entertain	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x46:1000	Ahmad Dhani mendapat dukungan dan doa	<p><!--img 1-->JAKARTA, KOMPAS.com - Ah
15	1bbce0e1	Konstruksi Dike	https://properti	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x40:1000	Sejauh ini, ada enam ruas tol yang digarap	<p><!--img 1-->JAKARTA, KOMPAS.com <!--strong> Pe
16	659349bc	Fakta Barvak Pe	https://otomotif	https://asset.kompas.com/crop/0	https://asset.kompas.com/crop/0x0:900	Mau terlihat keren, tapi mengalalkan sepa	<p><!--img 1--><!--img 1-->JAKARTA, K
17							

Gambar 3.6 Isi dalam Artikel0.csv

Gambar 3.6 merupakan isi dari file csv yang sudah di convert sbelumnya dan isi dari file ini nantinya akan digunakan untuk debat capres yang akan segera berlangsung.

3.3.3 Membuat analisis terhadap data twitter berdasarkan worldcloud. Week 3

Pada minggu ketiga, bertugas untuk menganalisis dan membuat laporan terhadap data twitter kompas yang memiliki topik tetang debat capres. Pak Yustinus Eko memberikan tugas untuk memantau dan membuat laporan analisa pada data Twitter yang membahas tentang debat capres. Pak Yustinus Eko memberikan

Rscript buatannya untuk dapat melakukan analisis penggunaan *hashtag*, *user* yang sering melakukan *tweet*, dan analisis kata terbanyak yang sering diperbicarakan di Twitter.



```
library(Rcurl)
library(plyr)
library(tm)
library(rjson)
library("SnowballC")
library("wordcloud")
library("RColorBrewer")
library(stringr)

raw_data <- ('kompas_all_170119_4.json')
data <- fromJSON(sprintf("[%s]", paste(readLines(raw_data),collapse=",")))
data <- apply(data, "I", "text")
data <- data.frame(data)

numberOfTweet <- nrow(data)
print('Jumlah tweet: ')
print(numberOfTweet)

docs <- Corpus(VectorSource(data$data))

toSpace <- content_transformer(function(x, pattern) gsub(pattern, " ", x))
docs <- tm_map(docs, content_transformer(tolower))
docs <- tm_map(docs, removeWords, c('ini', 'yang', 'ada', 'dan', 'saya', 'nau', 'itu', 'aja', 'sama', 'i', 'aku', 'juga', 'yuk', 'nih', 'dari', 'untuk', 'sih', 'lagi', 'nya', 'jadi', 'baru', 'saat',
'gue', 'apa', 'kan', 'deh', 'gua', 'cont', 'pak', 'tco', 'akan', 'dengan', 'kawi', 'kita'))
# Remove punctuations
docs <- tm_map(docs, removePunctuation)
# Eliminate extra white spaces
docs <- tm_map(docs, stripWhitespace)
# Text stemming
#docs <- tm_map(docs, stemDocument)

dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m), decreasing=TRUE)
d <- data.frame(word = names(v), freq=v)
head(d, 20)
print('Jumlah kata unik: ')
print(nrow(d))
print('Jumlah tweet: ')
print(numberOfTweet)

set.seed(1234)
png('kompas_all_170119_4_00000-90000.png')
wordcloud(words = d$word, freq = d$freq, min.freq = 1,
max.words=200, random.order=FALSE, rot.per=0.35,
colors=brewer.pal(8, "Dark2"))
dev.off()
```

Gambar 3.7 freqCount_4.r

Gambar 3.7 merupakan Rscript yang diberikan untuk menganalisis data twitter berdasarkan *hashtag* yang digunakan, kata yang paling sering diperbincangkan, dan nama user yang sering melakukan tweet. Pada gambar 3.7, diambil data-data twitter dalam format json yang akan diproses dan mengeluarkan output file png. File dalam format png tersebut berisikan gambar *wordcloud* yang terdiri dari kata-kata yang sering diperbincangkan di twitter. Dalam isi freqCount_4.r terdapat bagian `removeWords` yang berguna untuk menyingkirkan kata-kata yang sering digunakan namun tidak berguna. Salah satu contohnya seperti kata *sambung dan*, *atau*, *dari*, *ini*, *dll*.

```

word freq
jokowi                jokowi 8284
01indonesiamaju      01indonesiamaju 4354
hukum                 hukum 4335
prabowo               prabowo 4262
dulu                  dulu 4108
korupsi               korupsi 3822
negara                negara 3476
harus                 harus 3460
kenapa                kenapa 3433
anda                  anda 3386
seperti               seperti 3329
indonesia             indonesia 3282
debatpilpres2019     debatpilpres2019 3270
jurkam                jurkam 3242
debat                 debat 3200
kayak                 kayak 3172
menuduhnuduh         menuduhnuduh 3169
babak                 babak 3147
belur                 belur 3145
😁                    😁 3143
nuduhnuduh           nuduhnuduh 3142
handokotjung         handokotjung 3138
digebukin...         digebukin... 3136
andihiyat            andihiyat 2524
jokowiaminmenangdebat jokowiaminmenangdebat 2514
...                   ... 2500
banyak                banyak 2454
caleg                 caleg 2323
selalu                selalu 2151
bapak                 bapak 2079
kalo                  kalo 2029
menurut               menurut 1978
partai                partai 1958
pimpin               pimpin 1859
icw                   icw 1856
masalah               masalah 1781
pernah                pernah 1777
mardanielisera       mardanielisera 1738
2019gantipresiden   2019gantipresiden 1630
laporkan              laporkan 1402
[1] "Jumlah kata unik: "
[1] 18767
[1] "Jumlah tweet: "
[1] 29429

```

Gambar 3.8 Hasil *output* dari *kompas_all_170119_1.json*

Gambar 3.8 merupakan potongan hasil *output* dari *kompas_all_170119* yang menampilkan jumlah tweet, jumlah kata, serta kata apa saja yang sering digunakan oleh user dalam melakukan tweet seputar debat capres.

	A	B	C	D	E	F	G	H	I	
1	Analisis	kompas_all_1	s/d	kompas_all_5						
2										
3	17 Januari	Jumlah tweet :	49498							
4	Frekuensi									
5	jokowi	prabowo	korupsi	debat	debatpilpres2019	caleg	01indonesiamaju	hukum	bapak	
6	13434	7366	7298	6795	6146	4661	4431	4364	3897	
7	Pembahasan terhadap jokowi lebih banyak dibahas dibanding prabowo dan topik yang dibahas terkait koruptor, debat, debatpilpres2019,caleg, 01indonesiamaju									
8	Terdapat beberapa user id yang sering melakukan tweet yaitu:					andhidayat	erickthohir	ferdinanhean	ustadtengkuzul	
9	Terdapat beberapa hashtag yang sering digunakan yaitu:					prabowosandibawasolusi	werbaper	hoaxbpn	debatpilpres2019	
10										
11	18 Januari	Jumlah tweet :	60994							
12	Frekuensi									
13	jokowi	debat	01indonesiamaju	indonesia	prabowo	presiden	jokowilagi	2019pilihjokowi	2019tetapjokowi	
14	16146	7748	7317	6205	5778	4068	3425	3011	2411	
15	Pembahasan terhadap jokowi lebih banyak dibahas dibanding prabowo dan topik yang dibahas terkait debat, indonesia, 01indonesiamaju, indonesia									
16	Terdapat beberapa user id yang sering melakukan tweet yaitu:					mandanialisera	jajangridwan19	mohmahfudmd	yusufdumdum	
17	Terdapat beberapa hashtag yang sering digunakan yaitu:					01indonesiamaju	2019tetapjokowi	2019pilihjokowi	prabowoindonesiamaju	
18										
19	19 Januari	Jumlah tweet :	60821							
20	Frekuensi									
21	jokowi	prabowo	01indonesiamaju	debat	presiden	indonesia	2019pilihjokowi	sandiuno	eropa	
22	16647	9757	7335	7111	5712	5359	5315	4083	3833	

Gambar 3.11 Potongan laporan data *tweet* menggunakan data asli

27	20 Januari	Jumlah tweet :	66087							
28	Frekuensi									
29	jokowi	jokowiorangbaik	01indonesiamaju	2019tetapjokowi	prabowo	sultengjokowimenanglagi	2019jokowilagi	korupsinyatakseberapa	presiden	
30	16943	12742	8764	8698	6968	6685	6562	6438	5491	
31	Pembahasan terhadap jokowi lebih banyak dibahas dibanding prabowo dan topik yang dibahas terkait debat, indonesia, 01indonesiamaju, indonesia									
32	Terdapat beberapa user id yang sering melakukan tweet yaitu:					nabila10011	taniaiaaa10	sabila0102	jajangridwan19	
33	Terdapat beberapa hashtag yang sering digunakan yaitu:					01indonesiamaju	2019gantipresiden	01jokowilagi	2019pilihjokowi	
34										
35	21 Januari	Jumlah tweet :	7246							
36	Frekuensi									
37	jokowi	01indonesiamaju	prabowo	2019pilihjokowi	01jokowilagi	rakyat	2019gantipresiden	jokowiorangbaik	jokowilagi	
38	2107	908	775	687	561	480	460	453	452	
39	Pembahasan terhadap jokowi lebih banyak dibahas dibanding prabowo dan topik yang dibahas terkait 01indonesiamaju, 2019pilihjokowi, 01jokowilagi, rakyat, 2019gantipresiden									
40	Terdapat beberapa user id yang sering melakukan tweet yaitu:					jajangridwan19	p3nj3l4j4h	heniunique	fadlizon	
41	Terdapat beberapa hashtag yang sering digunakan yaitu:					jokowisatukanindonesia	01indonesiamaju	jokowiorangbaik	jokowudahkasihb	
42										
43	25 Januari	Jumlah tweet :	63590							
44	Frekuensi									
45	01jokowilagi	jokowi	presiden	01indonesiamaju	pekerja	lokal	rakyat	prabowo	2019pilihjokowi	
46	25510	10078	3711	3185	3188	3121	2781	2517	21	

Gambar 3.12 Potongan laporan data *tweet* menggunakan data asli

Gambar 3.11 dan Gambar 3.12 merupakan potongan laporan data *tweet* yang telah dianalisis berdasarkan tanggal, jumlah kata yang sering digunakan, nama

user yang sering melakukan tweet, serta hashtag yang sering digunakan menggunakan data asli.

3.3.4 Melakukan rapat mengenai pembagian tugas terhadap debat capres ke-2 yang akan berlangsung sampai dengan debat capres ke-2 berlangsung. Week 4

Pada minggu keempat penulis bertugas untuk membuat *script* Python bernama “linkkeyword.py” yang terdapat pada gambar 3.13.



```
linkkeyword.py
~/Desktop/dataArtikel/wordcounter

import csv
data = []
a = 1

with open('/home/pc-bigdata-1/Desktop/dataArtikel/wordcounter/Artikel_all.csv', newline='') as csvfile:
    spamreader = csv.reader(csvfile, delimiter=',')
    for row in spamreader:
        data.append(row)

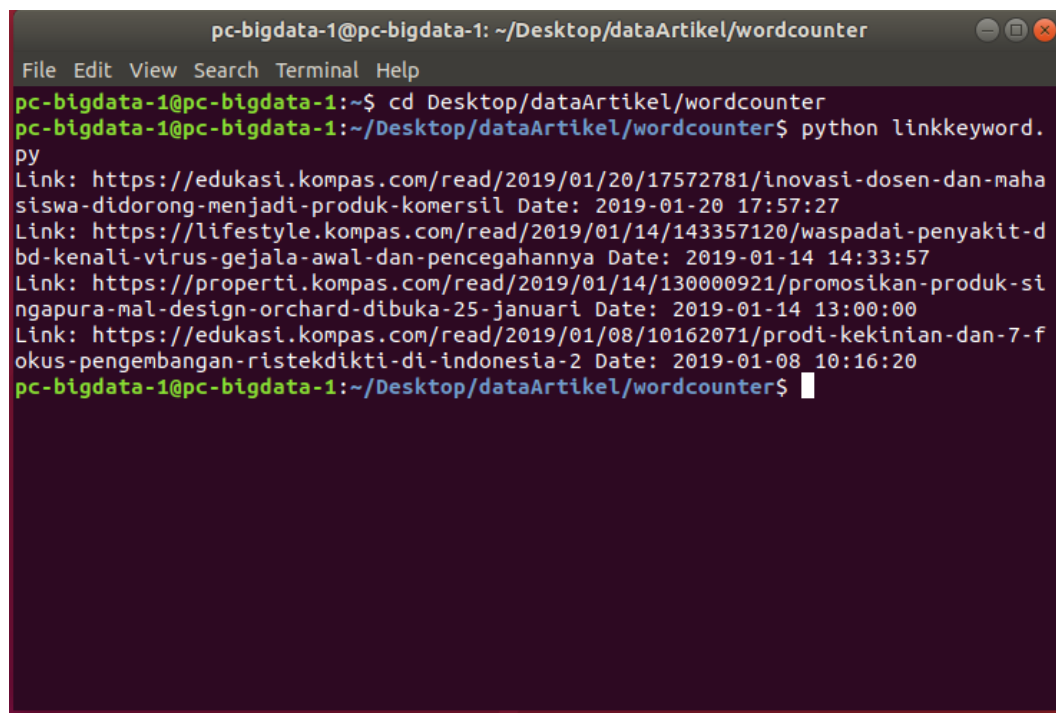
keyword1 = 'inkubasi'
#keyword2 = 'indonesia'
#keyword3 = 'tidak perlu'
#keyword4 = '4 tahun'
#keyword5 = 'negara'
#keyword6 = ''
#keyword7 = ''
namaoutputfile = 'hasil.csv'
datawrite = []

with open(namaoutputfile, 'w', newline='') as csvfile:
    writer = csv.writer(csvfile, delimiter=',')
    for a in range(1,10005):
        if keyword1 in data[a][6]:
            #if keyword2 or keyword3 in data[a][6]:
            #if keyword2 in data[a][6]:
            #if keyword3 in data[a][6]:
            #if keyword4 in data[a][6]:
            #if keyword5 in data[a][6]:
            #if keyword7 in data[a][6]:
            print ('Link: ' + data[a][2] + ' Date: ' + data[a][7])
            datawrite = data[a][2],data[a][7]
            writer.writerow(datawrite)
            a += 1
```

Gambar 3.13 linkkeyword.py

Gambar3.13 merupakan *script* yang akan digunakan saat debat capres akan berlangsung. *Script* ini berfungsi untuk menghasilkan sebuah tanggal dan *link* berdasarkan *keyword* yang terdapat pada file csv yang telah dibuat sebelumnya.

Proses pencarian dilakukan dengan membaca semua semua data yang ada pada “ArtikelAll.csv”. Pada bagian keyword 1 sampai dengan keyword 7 digunakan untuk mengisi kata kunci yang diberikan oleh tim ilkom pada saat debat capres berlangsung. Setelah kata kunci di masukkan *script* “linkkeyword .py” akan menghasilkan output sebagai berikut.



```
pc-bigdata-1@pc-bigdata-1: ~/Desktop/dataArtikel/wordcounter
File Edit View Search Terminal Help
pc-bigdata-1@pc-bigdata-1:~$ cd Desktop/dataArtikel/wordcounter
pc-bigdata-1@pc-bigdata-1:~/Desktop/dataArtikel/wordcounter$ python linkkeyword.py
Link: https://edukasi.kompas.com/read/2019/01/20/17572781/inovasi-dosen-dan-maha
siswa-didorong-menjadi-produk-komersil Date: 2019-01-20 17:57:27
Link: https://lifestyle.kompas.com/read/2019/01/14/143357120/waspada-penyakit-d
bd-kenali-virus-gejala-awal-dan-pencegahannya Date: 2019-01-14 14:33:57
Link: https://properti.kompas.com/read/2019/01/14/130000921/promosikan-produk-si
ngapura-mal-design-orchard-dibuka-25-januari Date: 2019-01-14 13:00:00
Link: https://edukasi.kompas.com/read/2019/01/08/10162071/prodi-kekinian-dan-7-f
okus-pengembangan-ristekdikti-di-indonesia-2 Date: 2019-01-08 10:16:20
pc-bigdata-1@pc-bigdata-1:~/Desktop/dataArtikel/wordcounter$
```

Gambar 3.14 Hasil *script* linkkeyword.py pada terminal

Gambar 3.14 merupakan hasil dari linkkeyword.py pada terminal yang berupa link dan tanggal berita berdasarkan keyword yang diberikan oleh tim ilkom.

	A	B
1	https://edukasi.kompas.com/read/2019/01/20/17572781/inovasi-dosen-dan-mahasiswa-didorong-menjadi-produk-komersil	2019-01-20 17:57:27
2	https://lifestyle.kompas.com/read/2019/01/14/143357120/waspada-penyakit-dbd-kenali-virus-gejala-awal-dan-pencegaha	2019-01-14 14:33:57
3	https://properti.kompas.com/read/2019/01/14/13000921/promosikan-produk-singapura-mal-design-orchard-dibuka-25-jan	2019-01-14 13:00:00
4	https://edukasi.kompas.com/read/2019/01/08/10162071/prodi-kekinian-dan-7-fokus-pengembangan-ristekdikti-di-indonesi	2019-01-08 10:16:20

Gambar 3.15 Hasil *script* linkkeyword.py dalam bentuk csv

Gambar 3.15 merupakan hasil linkeyword.py berisikan link dan tanggal dalam bentuk csv yang nantinya akan diberikan kepada tim ilkom pada saat debat capres berlangsung.

Pada minggu ke-4 juga kan dimulai debat capes kedua, sehingga kami dari tim SI dan tim Ilkom melakukan simulasi live dan rapat megenai ruangan yang akan digunakan nantinya, alat komunikasi yang akan diggunakan pada saat debat capres berlangsung, dan rundown yang akan dilakukan. Berikut merupakan proses yang dilakukan ketika debat capres berlagsung.

	A	B	C	D	E
1	SEGMENT 1	SEGMENT 2	SEGMENT 3	SEGMENT 4	SEGMENT 5
2	PBB tiga kriteria keberhasilan negara pangan, energi, air tanpa impor	index per km infrastruktur jalan tol LRT vietnam murah		kolusi pemerintah swasta	jokowi janji tidak impor pangan
3		15000 pekerja pabrik mobil Jerman	pelanggaran pajak limbah puluhan tahun		impor gula jutaan ton
4		brazil b90	KLH awasi departemen kehutanan	hukum internasional kurang	
5		80% plasma	pencemaran lingkungan perusahaan besar	pelaku eksploitasi kabur	impor 1 bulan sebelum panen
6		jalan tol LRT maroko murah			pajak untuk e-commerce
7			tanah tidak bertambah	nelayan minim akses teknologi	11.400 triliun rupiah di luar negeri
8		jalan tol LRT thailand murah		aturan menghambat nelayan	
9		proyek infrastruktur tidak efisien		nelayan kurang prasarana	impor 2 bulan sesudah panen
10		proyek infrastruktur susah dibayar		nelayan minim akses modal	negara boleh impor saat panen
11				operasional pelabuhan oleh asing	
12		biaya pembangunan mahal			
13					
14		utang pengaruhi efisiensi infrastruktur			
15					
16					

Gambar 3.16 keyword tiap segmen debat capres

Pada debat capres berlangsung kami menggunakan whatsapps sebagai media komunikasi tim Ilkom dan tim SI. Tim SI berada di Lab Big Data yaitu C503 sedangkan tim ilkom berada di News room yaitu B6. Debat Capres akan berlangsung selama 5 segmen, pada gambar 3.16 kami menggunakan link <https://docs.google.com/spreadsheets/d/1tqURIBS238cP8ycHucKfniC8lwvMlucKFg57mTYqxNY/edit#gid=1100111308>. Link tersebut merupakan link google drive yang digunakan sebagai media dimana tim ilkom akan menginput keyword-keyword yang akan diproses. Rundown yang akan dilakukan adalah tim ilkom akan memasukan keyword selama segmen 1 berlangsung, ketika segmen 1 selesai iklan akan berlangsung dan tim SI melakukan proses pencarian link dan tanggal berdasarkan keyword yang telah diberikan, begitupula selanjutnya ketika segmen 2 dimulai.

	A	B	C
8	jalan tol LRT maroko murah	https://properti.kompas.com/read/2017/01/28/131308521/kereta.cepat.pertama.afrika.segera.meluncur.di.maroko	1/24/2019 19:30
9	jalan tol LRT maroko murah	https://properti.kompas.com/read/2019/02/17/212418121/prabowo-anggap-infrastruktur-jokowi-lebih-mahal-dari-luar-negeri	1/23/2019 17:36
10	jalan tol LRT maroko murah	https://ekonomi.kompas.com/read/2016/01/29/211400726/Kereta.Cepat.di.Indonesia.Lebih.Mahal.dari.Iran.Ini.Penjelasan.M	1/22/2019 16:50
11	jalan tol LRT thailand murah		
12	proyek infrastruktur tidak efisien	https://nasional.kompas.com/read/2019/01/25/20432401/fahri-hamzah-sepakat-dengan-kritik-wapres-kalla-ter	1/25/2019 20:43
13		https://nasional.kompas.com/read/2019/01/25/08555611/wapres-iktak-elok-kritik-pembangunan-lrt	1/25/2019 8:55
14		https://properti.kompas.com/read/2019/01/15/150000221/jalan-perbatasan-wajah-indonesia-yang-dulu-dilupa	1/15/2019 15:00
15	proyek infrastruktur susah dibayar	https://kilas.kommentarian.kompas.com/ditjen-penyediaan-perumahan/read/2018/03/14/210000821/kontraktor-swasta-kompla	1/22/2019 14:53
16		https://intern...ibiayai-china	
17	biaya pembangunan mahal	https://internasional.kompas.com/read/2019/01/26/12561521/malaysia-batalkan-proyek-bernilai-rp-281-triliun	1/26/2019 12:56
18		https://nasional.kompas.com/read/2019/01/25/20432401/fahri-hamzah-sepakat-dengan-kritik-wapres-kalla-ter	1/25/2019 20:43
19		https://nasional.kompas.com/read/2019/01/25/08555611/wapres-iktak-elok-kritik-pembangunan-lrt	1/25/2019 8:55
20		https://ekonomi.kompas.com/read/2019/01/23/082237726/mengapa-harga-tiket-pesawat-mahal	1/23/2019 8:22
21		https://regional.kompas.com/read/2019/01/20/13031111/penjelasan-bni-soal-acara-dana-desa-yang-dihadiri-jo	1/20/2019 13:03
22		https://ekonomi.kompas.com/read/2019/01/18/160000426/strategi-prabowo-sandiaga-naikkan-tax-ratio-untuk-	1/18/2019 16:00
23		https://megapolitan.kompas.com/read/2019/01/16/18414291/potret-permukiman-kumuh-warga-petamburan-	1/16/2019 18:41
24		https://nasional.kompas.com/read/2019/01/15/13441781/10-janji-politik-capres-cawapres-prabowo-sandi-pada-	1/15/2019 13:44
25		https://ekonomi.kompas.com/read/2019/01/15/073100126/saat-pembangunan-lrt-jabodebek-dikritik-wapres-ki	1/15/2019 7:31

Gambar 3.17 Hasil link dan tanggal

Pada gambar 3.17 penulis menggunakan link

<https://docs.google.com/spreadsheets/d/1qhZh4vQpBqoH4ZHao-W3oMlthXl4gurndIXtl0xjpPI/edit#gid=1071856252>.

Link tersebut merupakan link google drive dimana tim SI akan memasukkan link dan tanggal berdasarkan keyword-keyword yang telah diberikan oleh tim Ilkom. Setelah link berikan maka tim Ilkom akan melakukan fact checking untuk menentukan apakah pernyataan yang diberikan oleh calon presiden sesuai fakta atau tidak.



Gambar 3.18 Hasil fact checking pada kompas.com

Gambar 3.18 merupakan hasil fact checking yang dimasukkan kedalam artikel kompas.

3.3.6 Melakukan rapat mengenai pembagian tugas terhadap debat capres ke-3 yang akan berlangsung sampai dengan debat capres ke-3 berlangsung. Week 6

Pada minggu keenam magang, debat capes ketiga akan berlangsung. Debat capres ketiga kali ini akan dilakukan oleh bapak ma'aruf amin dan bapak sandiaguno.

	A	B	C	D	E
1	SEGMENT 1	SEGMENT 2	SEGMENT 3	SEGMENT 4	SEGMENT 5
2	1500 kunjungan dalam 7 bulan		Rumah siap kerja link and match	saat ini 7 juta pengangguran	satu pertiga anak indonesia kekurangan
3	BPJS tidak menanggung per 1% dan 2% alokasi anggaran		pengangguran didominasi siswa smk	kualitas lapangan kerja belum optimal	lulusan SMK sulit mendapatkan lapangan
4		riset universitas tidak digunakan	61 persen pengangguran angkatan muda	kewajiban tenaga kerja asing bahasa	61% pengangguran di usia muda 1
5				Pengemudi online belum mendapatkan	pemerintah mengimpor tenaga kerja
6		pelayanan kesehatan berbelit		99,9% unit usaha adalah umkm	pemerintah mengekspor tenaga kerja
7			Prediksi ekonomi Indonesia nomor 7	5,5 juta UMKM belum dimudahkan ke	90.000 pegiat OK OCE Jakarta
8			sandi pernah jadi pengangguran	97% lapangan kerja dari sektor UMKM	90 ribu pegiat OK OCE Jakarta
9			Oke Oce menurunkan pengangguran	serikat pekerja mengeluhkan sistem	69% milenial ingin sistem fleksibel
10				60% dari 400 triliun anggaran pendidikan	riset buktikan milenial ingin jadi entrepreneur
11			1961 Bung Karno bertemu Menteri pendidikan	di daerah gagal	riset buktikan milenial ingin jadi entrepreneur
12				banyak guru belum mendapatkan kes	rumah siap kerja berikan banyak pekerjaan
13			UNESCO Indonesia negara adidaya ke	siswa mengeluhkan UN tidak adil	

Gambar 3.20 keyword tiap segmen debat capres ketiga

Debat capres akan berlangsung berlangsung selama 5 segmen, Tim Si diberikan 2 buah link oleh tim Ilkom. Pada gambar 3.20 kami menggunakan link https://docs.google.com/spreadsheets/d/1C9v_wonsfKr6aspBdEcjI34chBmNiDL5Qxt8amvC4qA/edit#gid=1942389541. Link tersebut merupakan link google drive yang digunakan sebagai media dimana tim ilkom akan menginput keyword-keyword yang akan diproses. Rundown yang akan dilakukan sama seperti sebelumnya yaitu tim ilkom akan memasukan keyword selama segmen 1 berlangsung, ketika segmen 1 selesai iklan akan berlangsung dan tim SI melakukan proses pencarian

link dan tanggal berdasarkan keyword yang telah diberikan, begitupula selanjutnya ketika segmen 2 dimulai.

	A	B
5	1% dan 2% alokasi anggaran riset teknologi	https://megapolitan.kompas.com/read/2019/02/11/19301961/ambil-alih-pengelolaan-ai
6	1% dan 2% alokasi anggaran riset teknologi	https://nasional.kompas.com/read/2019/01/16/06021581/mendagri-1166-lembaga-tela
7	riset universitas tidak digunakan	https://money.kompas.com/read/2019/03/01/093155726/menilik-besarnya-anggaran-ris
8	riset universitas tidak digunakan	https://ekonomi.kompas.com/read/2019/01/18/114501726/dukung-pembangunan-papu
9	riset universitas tidak digunakan	https://ekonomi.kompas.com/read/2019/01/08/122411226/pembangunan-2019-tak-har
10	riset universitas tidak digunakan	https://ekonomi.kompas.com/read/2019/01/07/170000826/terungkap-zodiak-ini-paling
11	riset universitas tidak digunakan	https://ekonomi.kompas.com/read/2018/11/05/093300226/sri-mulyani-human-capital-l
12	pelayanan kesehatan berbelit	tidak ada
13	pengangguran didominasi siswa smk	https://nasional.kompas.com/read/2019/03/06/20492021/jokowi-dilaporkan-ke-bawaslu
14	61 persen pengangguran angkatan muda	https://money.kompas.com/read/2019/03/11/093303326/9-konsep-keuangan-yang-perli
15	62 persen pengangguran angkatan muda	https://nasional.kompas.com/read/2019/03/07/00131491/timses-prabowo-janji-jokowi-l
16	63 persen pengangguran angkatan muda	https://nasional.kompas.com/read/2019/03/04/19495321/bpn-prabowo-nilai-kartu-pra-l
17	sandi pernah jadi pengangguran	https://ekonomi.kompas.com/read/2017/05/23/21564941/anies.status.saya.sekarang.penganggur

Gambar 3.21 Hasil link dan tanggal

Pada gambar 3.21 penulis menggunakan link

<https://docs.google.com/spreadsheets/d/1r3QIkX6TCLR8q0mReAnDNOAffauU0-VV8-Ch6K7tnu0/edit#gid=257447014>

Link tersebut merupakan link google drive dimana tim SI akan memasukkan link dan tanggal berdasarkan keyword-keyword yang telah diberikan oleh tim Ilkom. Setelah link berikan maka tim Ilkom akan melakukan fact checking seperti debat capres sebelumnya untuk menentukan apakah pernyataan yang diberikan oleh calon wakil presiden sesuai fakta atau tidak.

3.3.7 Melakukan evaluasi mengenai debat capres ke-3 yang telah berlangsung, melakukan rapat mengenai pembagian tugas terhadap debat capres ke-4 yang akan berlangsung sampai dengan debat capres ke-4 berlangsung serta melakukan training terhadap mahasiswa calon kerja magang selanjutnya. Week7

Pada minggu ketujuh magang, dilakukan evaluasi dengan tim Ilkom yang membahas mengenai kendala-kendala yang terjadi saat debat capres ketiga berlangsung. Pada rapat tersebut tim Ikom menyampaikan bahwa data hasil *link* yang diberikan jauh lebih banyak yang relevan dan memenuhi kriteria. Dari tim Ilkom juga menyambakan bahwa untuk selanjutnya mereka ingin mengambil beberapa artikel slain dari *website* kompas sehingga hasil pencarian bisa lebih luas.

Pada minggu ketujuh, penulis dikenalkan pada mahasiswa calon magang baru yang akan meneruskan tugas penulis untuk debat capres keempat. Penulis juga diberikan tugas oleh pak eko untuk memberikan *training* kepada mahasiswa calon magang seputar tugas yang akan dilakukan pada debat capres keempat yang akan berlangsung, sehingga nantinya dapat berjalan dengan baik seperti sebelumnya.

3.3.8 Pembuatan laporan dokumentasi. Week 8

Pada minggu kedelapan, dilakukan tugas untuk membuat laporan dokumentasi selama kegiatan kerja magang berlangsung. Laporan tersebut berisikan *script-script* yang telah dibuat, penjelasan mengenai *script* tersebut, serta hasil yang dikeluarkan oleh *script* tersebut.

3.4 Kendala yang dihadapi

Selama melakukan kegiatan kerja magang di Universitas Multimedia Nusantara, terdapat beberapa kendala yang dialami diantaranya:

1. Kurangnya pemahaman terhadap penggunaan system oprasi Linux serta kurangnya pemahaman terhadap bahasa pemograman R dan Python, sehingga masih harus belajar terlebih dahulu.
2. Hasil yang dikeluarkan ada yang tidak relevan dikarenakan *keyword-keyword* yang diberikan oleh tim Ilkom terlalu luas.

3.5 Solusi atas kendala

Solusi atas permasalahan di atas yang dilakukan dalam menyelesaikan tugas magang yaitu:

1. Belajar serta menanyakan hal yang tidak dimengerti kepada Pak Eko selaku pembimbing kerja magang.
2. Melakukan evaluasi dan diskusi kepada tim Ilkom agar mempersempit *keyword-keyword* yang diberikan agar lebih spesifik.