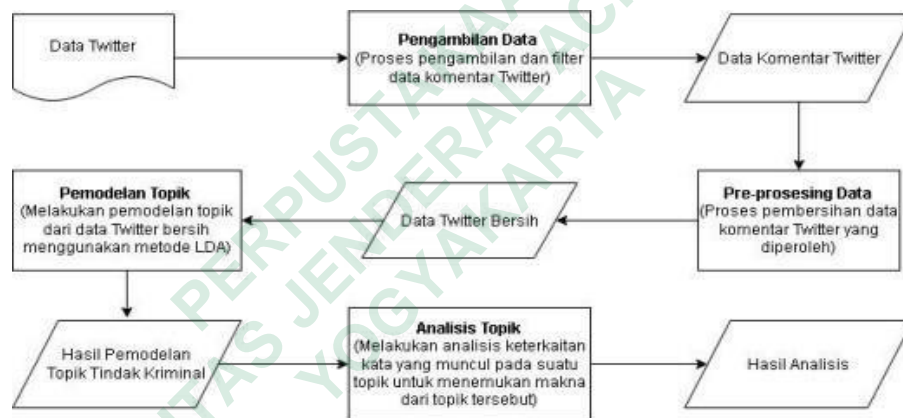


BAB 4

HASIL PENELITIAN

4.1 Ringkasan Hasil Penelitian

Penelitian ini bertujuan untuk membuat pemodelan topik data komentar twitter yang berkaitan dengan *cryptocurrency* untuk dilakukan analisis dari hasil model topik yang diperoleh. Analisis topik tersebut dilakukan untuk mencari keterkaitan kata yang muncul pada suatu topik, sehingga suatu kumpulan kata pada topik tersebut memiliki arti atau makna. Gambar 4.1 menunjukkan proses yang dilakukan untuk melakukan analisis topik data *cryptocurrency* di media sosial twitter.



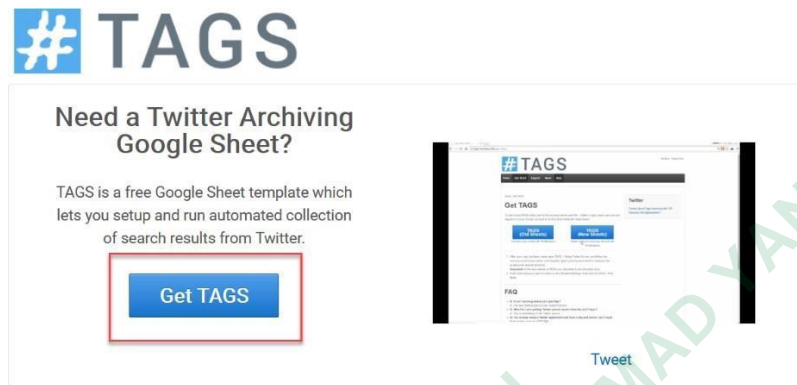
Gambar 4.1 Bagan Analisis Topik *Cryptocurrency*

4.2 Pengambilan Data

Proses pengambilan data tweet di twitter menggunakan TAGS. TAGS adalah singkatan dari *Twitter Archiving Google Sheet*. TAGS adalah alat pencari tweet data Twitter yang memungkinkan pengguna mengumpulkan tweet secara otomatis ke dalam lembar google. TAGS memiliki kemampuan untuk mengumpulkan Tweet dari orang tertentu, tagar (#) tertentu, atau istilah pencarian. Untuk mendapatkan TAGS, pengguna harus memiliki Akun Google dan Akun Twitter. Setelah itu dilanjutkan dengan membuat Google Sheet TAGS dan mengatur akses ke Twitter.

Setelah memiliki akun yang diperlukan, selanjutnya pengguna dapat membuat GoogleSheet TAGS sendiri dengan beberapa tahapan sebagai berikut.

- 1 Arahkan ke situs web TAGS dan tekan tombol Get TAGS



Gambar 4.2 Proses instalasi TAGS

- 2 Pilih versi TAGS 6.1



Gambar 4.3 Pilih versi TAGS

- 3 Pilih Make a copy

Google Sheets

Copy document

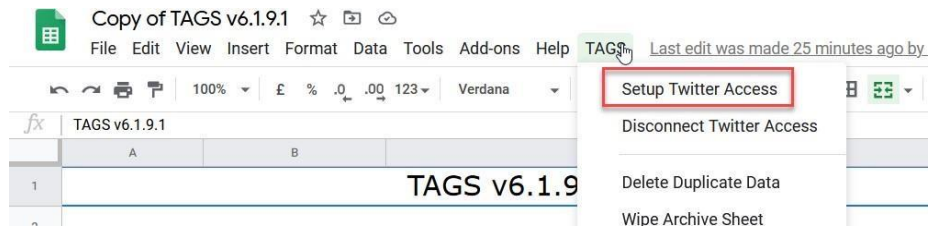
Would you like to make a copy of TAGS v6.1.9.1?

Make a copy



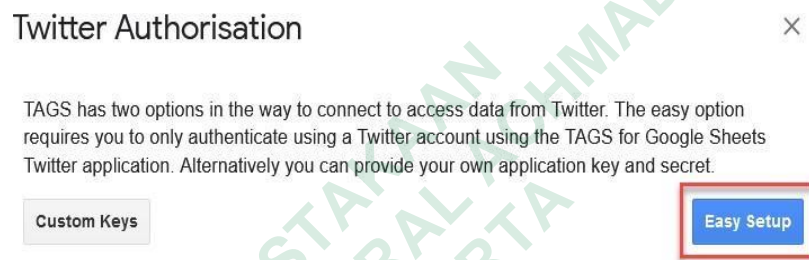
Gambar 4.4 Akhiri proses instalasi TAGS

Proses awal mengumpulkan tweet, pengguna harus memberikan akses *Google Sheet* ke account Twitter dengan tahapan memilih tab TAGS pada menu seperti pada gambar 4.5.



Gambar 4.5 Proses Setup TAGS ada account twitter

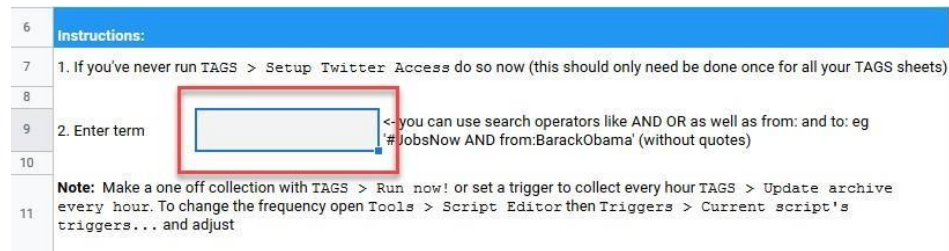
Sebuah pop-up akan muncul menanyakan bagaimana pengguna ingin mengatur otorisasi Twitter. Tekan Easy Setup dan ikuti petunjuknya



Gambar 4.6 Proses Otentifikasi API TAGS ke Twitter

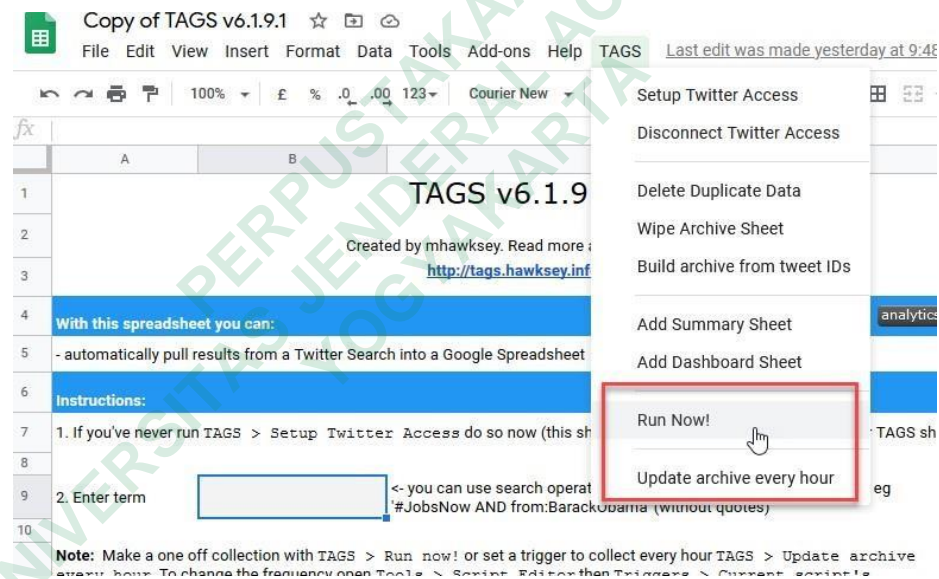
Setelah pengguna memberikan semua izin yang sesuai untuk TAGS, selanjutnya pengguna menyiapkan parameter pencarian dan menjalankan skrip. Kemudian dapat mengunduh data yang dibutuhkan untuk memulai proses analisis. Pengguna dapat memasukkan istilah pencarian apa pun pada dalam kotak pencarian, dan tweet apa pun yang berisi istilah pencarian tersebut akan muncul.

Pengguna dapat melakukan filter terhadap semua tweet untuk tagar tertentu dengan mengetikkan “#crypto”. TAGS akan hanya menampilkan data tweet dari orang tertentu atau tweet yang dikirim ke orang tertentu menggunakan namapengguna mereka dalam format berikut: “from:username”; atau “ke: nama pengguna”.



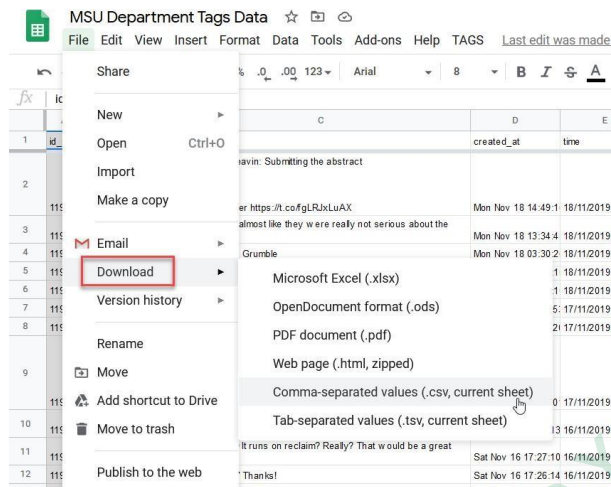
Gambar 4.7 Proses Filter Tweet pada Twitter

Twitter hanya memungkinkan pengguna untuk dapat mengumpulkan data tweet maksimal dari 7 hari terakhir. Jadi jika pengguna ingin menganalisis tweet sebelum dan sesudah peristiwa sesuai dengan topik yang diinginkan, pengguna dapat menyiapkan TAGS untuk menghapus tweet sebelumnya. Tapi, pastikan untuk mengawasi apakah lembar TAGS Anda telah mencapai batas tweet pengguna.



Gambar 4.8 Proses Menjalankan TAGS

Proses selanjutnya adalah unduh data dengan cara memilih format yang pengguna inginkan untuk data yang akan diunduh. Dalam penelitian ini data yang digunakan oleh pengguna menggunakan format CSV. Alasan pemilihan data dengan tipe CSV ini adalah agar data mudah diimport ke dalam aplikasi Python.



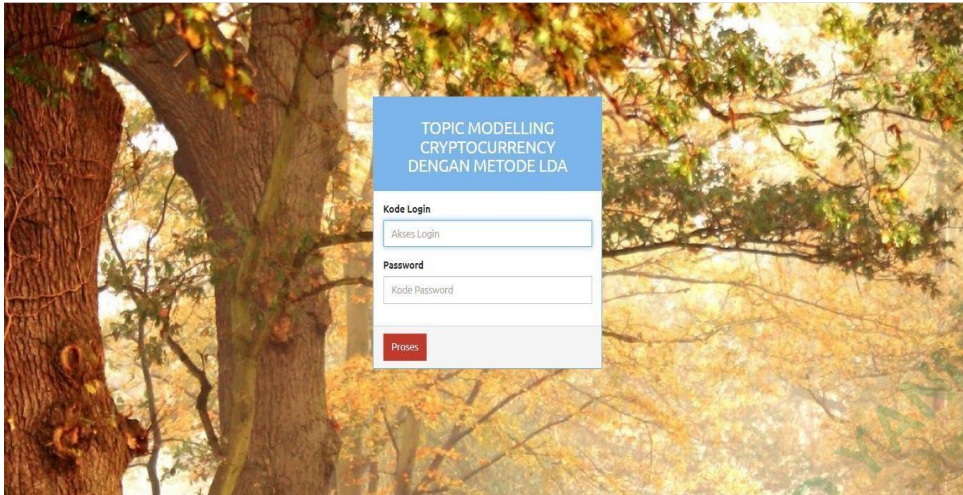
Gambar 4.9 Proses Unduh Data Tweet dalam Bentuk File CSV

TAGS bekerja dengan membuat permintaan untuk tweet (yang memenuhi kriteria pengguna) ke antarmuka program aplikasi (API) Twitter. API pada dasarnya bertindak sebagai administrator yang memutuskan apakah akan memberikan *Tweet* yang pengguna minta atau tidak. Artinya, Twitter mengawasi tweet mana yang dikirimkan kepada pengguna.

4.3 Implementasi Antar Muka

4.3.1 Halaman *Login*

Halaman *login* merupakan halaman awal ketika pengguna akan menggunakan aplikasi. Fungsi halaman *login* adalah sebagai salah satu komponen keamanan sistem agar tidak semua orang dapat mengakses aplikasi. Pada halaman *login* pengguna diminta untuk memasukkan login dan password yang dimiliki agar dapat mengakses halaman utama aplikasi. Tampilan halaman *login* dapat dilihat pada gambar 4.10.



Gambar 4.10 Halaman Login

Inputan data *login* dan *password* yang dimasukan oleh pengguna akan dilakukan verifikasi oleh sistem apakah datatersebut adadalam database atau tidak. Apabila data terdaftar dalam database, sistem secara otomatis akan menampilkan halaman utama dari aplikasi. Apabila data belum diketahui sistem akan menampilkan informasi kesalahan input data. Proses verifikasi *login* dan *password*apalikasi dapat dilihat pada gambar 4.11.

```

if request.method == 'POST':
    form = penggunaForm(request.POST)
    if form.is_valid():
        kd_login = form.cleaned_data.get('kd_login')
        kd_pass = form.cleaned_data.get('kd_pass')
        user = authenticate(kd_login=kd_login, kd_pass=kd_pass)
        if user is not None:
            login(request, user)
            return redirect('/home')

```

Gambar 4.11 Script Verfikasi Data *Password*

4.3.2 Halaman Utama Aplikasi

Halaman utama aplikasi berisikan informasi dan link dengan berbagai halaman lainnya terkait dengan proses *modelling topic Cryptocurrency*. Pada halaman utama aplikasi pengguna dapat mengakses semua halaman baik halaman pengaturan pengguna, halaman data *cryptocurrency* dan halaman proses pengujian LDA. Tampilan halaman utama aplikasi dapat dilihat pada gambar 4.12.



Gambar 4.12 Halaman Utama Aplikasi

Proses untuk menghubungkan halaman utama aplikasi dengan halaman-halaman lainnya menggunakan *script* bahasa pemrograman Django. Proses implementasi *script* pemrograman tersebut dapat dilihat pada gambar 4.13.

```
urlpatterns = [
    path('admin/', admin.site.urls),
    path('', views.index, name="index"),
    path('home/', views.home, name="home"),
    path('login_act/', views.login_act, name="login_act"),
    path('loginaction/', views.loginaction, name="loginaction"),
    path('logout_request/', views.logout_request, name="logout_request"),
    path('user_view/', views.user_view, name="user_view"),
    path('user_delete/<int:id>', views.user_delete, name="user_delete"),
    path('user_add/', views.user_add, name="user_add"),
    path('user_edit/<int:id>', views.user_edit, name="user_edit"),
    path('dtuser_edit/<int:id>', views.dtuser_edit, name="dtuser_edit"),
    path('uji_lda/', views.uji_lda, name="uji_lda"),
    path('proses_ujilda/', views.proses_ujilda, name="proses_ujilda"),
    path('lda_view/', views.lda_view, name="lda_view"),
    path('lda_add/', views.lda_add, name="lda_add"),
    path('lda_delete/<int:id>', views.lda_delete, name="lda_delete"),
    path('lda_edit/<int:id>', views.lda_edit, name="lda_edit"),
    path('dtlda_edit/<int:id>', views.dtlda_edit, name="dtlda_edit"),
]
```

Gambar 4.13 Script Konkoneksi *Link* Halaman Utama

4.3.3 Halaman Daftar *User*

Halaman daftar *user* merupakan menu yang digunakan untuk proses tambah, edit dan hapus data pengguna yang dapat mengakses aplikasi. Pada halaman ini pengguna dapat melakukan proses pembatasan jumlah orang yang

dapat mengakses aplikasi. Tampilan halaman daftar user dapat dilihat pada gambar 4.1.1.

ID	Nama User	Kode Login	Kode Password	Aksi
1	Admin Sistem	12345	54321	Edit Hapus
2	Staf Administrasi	admin	admin	Edit Hapus
3	Sekretaris	333	333	Edit Hapus
4	5 edit	5	5	Edit Hapus

Gambar 4.14 Halaman Daftar User

Proses penambahan pengguna pada menu daftar user dapat dilakukan pengguna dengan menekan tombol *Add New Record*. System secara otomatis akan masuk pada halaman tambah data pengguna seperti pada gambar 4.12.

Gambar 4.15 Tambah Data Pengguna Baru

Proses penambahan data pengguna baru diawali dengan proses *input* data pengguna. Proses selanjutnya adalah proses simpan pengguna baru dengan menekan tombol simpan. *Script* yang digunakan untuk proses penambahan data baru dapat dilihat pada gambar 4.13.


```

def user_add(request):
    if request.method == "POST":
        form = penggunaForm(request.POST)
        if form.is_valid():
            try:
                form.save()
            except:
                pass
            return redirect('/user_view')

```

Gambar 4.16 *Script* penambahan data

Menu lainnya yang pada pada menu daftar *user* adalah proses edit data pengguna. Menu ini berfungsi untuk melakukan proses *editing* pada data pengguna yang dapat mengakases aplikasi. Tampilan halaman edit data pengguna dapat dilihat pada gambar 4.14.



Gambar 4.17 Halaman Edit Data Pengguna

Proses edit data pengguna dapat direalisasikan dengan menggunakan *script* pemrograman Django seperti pada gambar 4.15.

```

def user_edit(request, id):
    dtuser = pengguna.objects.get(id=id)
    form = penggunaForm(request.POST, instance = dtuser)
    if form.is_valid():
        form.save()
        return redirect("/user_view/")
    return render(request, 'user_edit.html', {'dtuser': dtuser})

```

Gambar 4.18 *Script* Edit Data Pengguna

Proses lain yang ada pada halaman daftar pengguna adalah proses hapus data. Apabila pengguna hendak melakukan proses hapus data, sistem akan melakukan proses konfirmasi terlebih dahulu sebelum melakukan proses hapus data. Apabila konfirmasi disetujui oleh pengguna sistem akan melanjutkan proses hapus data. Script yang digunakan untuk proses hapus data dapat dilihat pada gambar 4.16.

```
def user_delete(request, id):
    dtpengguna = pengguna.objects.get(id=id)
    dtpengguna.delete()
    return redirect("/user_view")
```

Gambar 4.19 *Script* Hapus Data Pengguna

Halaman data cryptocurrency adalah halaman yang digunakan untuk memasukan data uang kripto yang *topic modelling* proses LDA pada aplikasi. Pada halaman ini pengguna dapat melakukan proses tambah, edit dan hapus jenis mata uang kripto yang akan di proses dalam perhitungan *topic modelling* dalam perhitungan LDA. Tampilan halaman data *cryptocurrency* dapat dilihat pada gambar 4.17.



ID	Nama Cryptocurrency	Singkatan	Keterangan	Aksi
1	Bitcoin	BTC	-	Edit Hapus
2	Ethereum	ETH	-	Edit Hapus
3	Tether	USDT	-	Edit Hapus
5	BNB	BNB	-	Edit Hapus

Gambar 4.20 Halaman Data Cryptocurrency

Pada halaman data *cryptocurrency* pengguna diberikan fasilitas untuk melakukan proses tambah data *cryptocurrency* ke dalam database. Data jenis mata uang kripto ini akan menjadi *topic modelling* yang akan diproses menggunakan metode LDA untuk mengetahui distribusi topik dan membuat rangking topik mata uang apa yang paling sering diperbincangkan dan muncul dalam tweet

media sosial twitter. Halaman untuk proses menambah data uang krypto dapat dilihat pada gambar 4.18.

Gambar 4.21 Halaman Tambah Data Uang Kripto

```
def lda_add(request):
    if request.method == "POST":
        form = dataaldaForm(request.POST)
        if form.is_valid():
            try:
                form.save()
                return redirect('/lda_view')
            except:
                pass
    else:
        form = dataaldaForm()
    return render(request, 'lda_add.html', {'form':form})
```

Gambar 4.22 Script Tambah Data Uang Kripto

Proses lain yang terdapat pada menu data uang krypto adalah proses edit data uang krypto. Pada halaman ini pengguna diizinkan untuk melakukan proses editing terhadap mata uang krypto apabila dianggap terdapat kesalahan. Tampilan halaman edit data dapat dilihat pada gambar 4.20.

Gambar 4.23 Halaman Edit Data Uang Kripto

Script bahasa pemrograman yang penulis gunakan untuk proses edit data uang kripto dapat dilihat pada gambar 4.21.

```
def dtlda_edit(request, id):
    dtlda = data_lda.objects.get(id=id)
    form = dataaldaForm(request.POST, instance = dtlda)
    if form.is_valid():
        form.save()
        return redirect("/lda_view")
    return render(request, 'lda_edit.html', {'dtlda': dtlda})
```

Gambar 4.24 Script Edit Data Uang Kripto

Proses lainnya yang merupakan fasilitas pada menu mata uang kripto adalah proses hapus data uang kripto dari database aplikasi. Proses yang penulis gunakan untuk menghapus data uang kripto dapat dilihat pada gambar 4.22.

```
def lda_delete(request, id):
    dtlda = data_lda.objects.get(id=id)
    dtlda.delete()
    return redirect("/lda_view")
```

Gambar 4.25 Script Hapus Data Uang Kripto

4.3.4 Halaman Proses Pengujian LDA

Halaman pengujian topik *modeling* dengan menggunakan LDA akan memberikan menu untuk proses untuk file CSV yang berisikan data tweet dari media sosial twitter yang berisikan informasi percakapan menggunakan twitter terkait dengan mata uang krypto. Tampilan halaman proses uji LDA dapat dilihat pada gambar 4.23.



Gambar 4.26 Halaman Proses Perhitungan Metode LDA

Proses perhitungan LDA diawali dengan proses input file CSV. Proses selanjutnya aplikasi akan membaca isi dari file CSV tersebut. Proses pembacaan file CSV menggunakan script pemrograman seperti pada gambar 4.24

```
data = pd.read_csv('data_tweet.csv', error_bad_lines=False);
data_text = data[['headline_text']]
data_text['index'] = data_text.number
documents = data_text
```

Gambar 4.27 Script Baca File CSV

Tahap selanjutnya adalah proses *Pre-Processing* data hasil tweet. Proses awal dari tahap *Pre-Processing* adalah dengan melakukan proses pembersihan data. Pada tahapan ini aplikasi akan menghapus kata yang terdapat *stopword*. *Stopword* merupakan proses penghapusan kata yang sering muncul namun tidak memiliki makna. *Tokenizing* adalah proses memisah atau memecah kalimat menjadi potongan berdasarkan tiap kata yang menyusunnya. Lalu kita menghapus kalimat yang hanya terdiri dari 2 huruf kata dengan fungsi *wordcount*. Hasil dari tahapan *Stopword* dapat dilihat pada tabel 4.1.

Tabel 4.1 Hasil dari tahapan *stopword*

No	Tweet Sebelum <i>Stopword</i>	Tweet Setelah <i>Stopword</i>
1	<p>Introducing Safeguard Program, which provides protection & support for individuals and institutions affected by security incidents. This includes assistance with technical support, evidence retention, asset confirmation and more.</p> <p>More info: https://land.kucoin.com/guardian?lang=en_US...</p>	<p>Introducing Safeguard Program, which provides protection & support individuals and institutions affected security incidents. includes assistance technical support, evidence retention, asset confirmation and more.</p> <p>More info: https://land.kucoin.com/guardian?lang=en_US...</p>
2	<p>#KuCoin CEO Livestream Recap - Latest Updates About Security Incident 0930 Details: https://kucoin.com/news/en-kucoin-ceo-livestream-recap-latest-updates-about-security-incident-0930...</p>	<p>#KuCoin Livestream Recap - Latest Updates About Security Incident 0930 Details: https://kucoin.com/news/en-kucoin-ceo-livestream-recap-latest-updates-about-security-incident-0930...</p>
3	<p>IM TRYING TO SELL ALL OF MY ALTS TO GO ALL IN ON \$RSR AND IFUCKING FORGOT KUCOIN HASNT ENABLED DEPOSITS/WITHDRAWALS YETFUCK MY LIFE</p>	<p>IM TRYING ALTS \$RSR FUCKING FORGOTKUCOIN HASNT ENABLED DEPOSITS/WITHDRAWALS FUCK LIFE</p>

4	<p>Due to the suspension of the deposits and withdrawals of \$NIM on #KuCoin, 'Nimiq (NIM) New Pair Trading Competition, 4,000,000 NIM to be Won!' will be postponed.</p> <p>Please stay tuned for the official announcement from KuCoin.</p> <p>Details: https://kucoin.com/news/en-nim-trading-competition-will-be-postponed...</p>	<p>suspension deposits withdrawals \$NIM #KuCoin, 'Nimiq (NIM) Pair Trading Competition, 4,000,000 NIM Won!' will postponed.</p> <p>Please stay tuned for the official announcement from KuCoin.</p> <p>Details: https://kucoin.com/news/en-nim-trading-competition-will-be-postponed...</p>
5	<p>Tokens recovered from the KuCoin hack have been returned to KuCoin wallets. All OCEAN user funds now reside with KuCoin.</p> <p>#TogetherStronger @lyu_johnny @kucoincom</p>	<p>Tokens recovered from KuCoin hack have been returned KuCoin wallets. OCEAN user funds reside with KuCoin.</p> <p>#TogetherStronger @lyu_johnny @kucoincom</p>

Script bahasa pemrograman yang penulis gunakan untuk mengimplementasikan proses *Stopword* menggunakan script bahasa pemrograman Python seperti pada gambar 4.25.

```
def removeStopword(str):
    factory = StopWordRemoverFactory()
    more_stopword = ['due','to','the','and','now','all','new','off','on',
, 'of','yet','my','i','to be','by','with','for','have','been']
    stopwords = factory.get_stop_words()+more_stopword
    word_tokens = str.split()
    filtered_sentence = [w for w in word_tokens if not w in
stopwords]
    return ' '.join(filtered_sentence)
```

Gambar 4.28 Script proses *Stopword*

Tahapan selanjutnya adalah proses untuk membersihkan data teks seperti remove URLs, *punctuation*, membuat teks setara dengan format *lowercase*. Hasil dari proses *remove URLs, punctuation* dan mengubah semua data menjadi huruf kecil dapat dilihat pada tabel 4.2.

Tabel 4.2 Hasil dari proses *remove URLs*

No	Tweet Sebelum Proses	Tweet Setelah Proses
1	<p>Introducing Safeguard Program, which provides protection & support individuals and institutions affected security incidents. includes assistance technical support, evidence retention, asset confirmation and more.</p> <p>More info: https://land.kucoin.com/guardian?lang=</p>	<p>introducing safeguard program which provides protection support individuals and institutions affected security incidents includes assistance technical support evidence retention asset confirmation info</p>

2	#KuCoin Livestream Recap - Latest Updates About Security Incident 0930 Details: https://kucoin.com/news/en-kucoin-ceo-livestream-recap-latest-updates-about-security-incident-0930...	kucoin livestream recap latest updates about security incident 0930 details
3	IM TRYING ALTS \$RSR FUCKING FORGOT KUCOIN HASNT ENABLED DEPOSITS/WITHDRAWLS FUCK LIFE	trying alts rsr fucking forgot kucoinhasnt enabled deposits withdrawls fuck life
4	suspension deposits withdrawals \$NIM #KuCoin, 'Nimiq (NIM) Pair Trading Competition, 4,000,000 NIM Won!' will postponed. Please stay tuned for the official announcement from KuCoin. Details: https://kucoin.com/news/en-nim-trading-competition-will-be-postponed...	suspension deposits withdrawals nim kucoin, nimiq nim pair trading competition 4000000 nim won postponed please stay tuned official announcement kucoin
5	Tokens recovered from KuCoin hack have been returned KuCoin wallets. OCEAN user funds reside with KuCoin. #TogetherStronger @lyu_johnny @kucoincom	tokens recovered from kucoin hack returned kucoin wallets ocean user funds reside kucoin togetherstronger

Script pemrograman yang penulis gunakan untuk proses implementasi tahapan *casefolding*, *punctuation*, menghapus URL dan email serta menghapus tanda baca yang tidak dikehendaki dapat dilihat pada gambar 4.26.

```

def cleaning(str):
    #hapus huruf non-ascii
    str = unicodedata.normalize('NFKD', str).encode('ascii', 'ignore').decode('utf -8',
'ignore')
    #hapus URLs dan email
    str = re.sub(r'(?:https?://www\d{0,3}[.][a-z0-9.\-]+[.][a-
z]{2,4})|(?:(^\\s(<>)+|(( [^\\s() <> ]+|\\ )+)) *\\ )+(?:\\ ([^\\s() <> ]+| \\ ([^\\s() <> ]+
))) *\\ |([\\s\\!()\\[\\{};:\\'\".,<>?«»“”‘’])', '', str)
    #hapus punctuations
    str = re.sub(r'[^\\w] |_', ' ', str)
    #hapus tanda yang tidak penting pada string
    str = re.sub("\\S*d\\S*", "", str).strip()
    #hapus tanda titik pada numerik
    str = re.sub(r"\\b\\d+\\b", " ", str)
    #ubah semua huruf menjadi lowercase
    str = str.lower()
    #hapus spasi yang tidak penting
    str = re.sub('\\[\\s]+', ' ', str)
    return str

def preprocessing(str):
    str = removeSentence(str)
    str = cleaning(str)
    str = removeStopword(str)
    return str

```

Gambar 4.29 Script proses *Cleanning* Data

Langkah berikutnya adalah melakukan implementasi *preprocessing* pada data teks tweet. Hasil dari data teks yang telah di lakukan *preprocessing* adalah kalimat dalam satu tweet telah dihilangkan tanda baca, disamakan format *lower case*, telah dilakukan *tokenizing* kata per kata dalam kalimat, selain itu kata yang terdiri dari 1 huruf juga sudah dihapus. Hasil dari proses *tokenizing* dapat dilihat pada tabel 4.3.

Tabel 4.3 Hasil *Tokenizing*

No	Tweet Sebelum Tokenizing	Tweet Setelah Tokenizing
1	introducing safeguard program which provides protection support individuals and institutions affected security incidents includes assistance technical support evidence retention asset confirmation info	'introducing', 'safeguard', 'program', 'which', 'provides', 'protection', 'support', 'individuals', 'institutions', 'affected', 'security', 'incidents', 'includes', 'assistance', 'technical', 'support', 'evidence', 'retention', 'asset', 'confirmation', 'info'
2	kucoin livestream recap latest updates about security incident 0930 details	'kucoin', 'livestream', 'recap', 'latest', 'updates', 'about', 'security', 'incident', '0930', 'details'

3	trying alts rsr fucking forgot kucoin hasnt enabled deposits withdrawlsfuck life	'trying', 'alts', 'rsr', 'fucking', 'forgot', 'kucoin', 'hasn't', 'enabled', 'deposits', 'withdrawls', 'fuck', 'life'
4	suspension deposits withdrawals nimkucoin, nimiq nim pair trading competition 4000000 nim won postponed please stay tuned official announcement kucoin	'suspension', 'deposits', 'withdrawals', 'nim', 'kucoin', 'nimiq', 'nim', 'pair', 'trading', 'competition', '4000000', 'nim', 'won', 'postponed', 'please', 'stay', 'tuned', 'official', 'announcement', 'kucoin'
5	tokens recovered from kucoin hack been returned kucoin wallets ocean user funds reside kucoin togetherstronger	'tokens', 'recovered', 'kucoin', 'hack', 'returned', 'kucoin', 'wallets', 'ocean' 'user', 'funds', 'reside', 'kucoin', 'togetherstronger'

Script yang penulis gunakan untuk proses *tokenizing* menggunakan bahasapemrograman phyton seperti pada gambar 4.27.

```

from helper import *
import warnings
warnings.filterwarnings('ignore')

import pandas as pd

data = data['tweet'].apply(lambda x: preprocessing(x))
print(data)

text_list = [i.split() for i in data]
print(len(text_list))
print(text_list)

```

Gambar 4.30 Script Proses Tokenizing

Setelah tahap *text Tokenizing* dilakukan aplikasi akan melakukan proses *bigram* dan *trigram model* yaitu dengan membuat dua kata yang sering muncul bersama dalam tweet atautiga kata yang sering muncul bersamaan menggunakan package gensim seperti pada script pemrograman gambar 4.28

```

import gensim

#Membuat Bigram dan Trigram Models
from gensim.models import Phrases
# membuat data bigram/trigram bila kata diulang min 10 kali .
bigram = Phrases(text_list, min_count=10)
trigram = Phrases(bigram[text_list])
for idx in range(len(text_list)):
    for token in bigram[text_list[idx]]:
        if '_' in token:
            # beri tanda sebagai bigram
            text_list[idx].append(token)
    for token in trigram[text_list[idx]]:
        if '_' in token:
            # beri tanda sebagai trigram.
            text_list[idx].append(token)

```

Gambar 4.31 Script Bigram dan Trigram Model

Hasil dari proses *bigram* dan *trigram* terhadap dokumen dari tweet dapat dilihat pada tabel 4.4. Tabel 4. 4 Hasil Proses *Bigram* dan *Trigram*

Tabel 4.4 Hasil Proses *Bigram* dan *Trigram*

No	Tweet Awal	Tweet Setelah Proses Bigram	Tweet Setelah Proses Trigram
1	'introducing', 'safeguard', 'program', 'which', 'provides', 'protection', 'support', 'individuals', 'institutions', 'affected', 'security', 'incidents', 'includes',	'introducing_ safeguard', 'safeguard _program', 'program_which', 'which_provides', 'provides_protection', 'protection'_support', 'support'_individuals', 'individuals'_institutions' , 'institutions_affected', 'affected'_security', 'security'_incidents', 'incidents	'introducing_safeguard_program', 'safeguard_program which', 'program_which_provides', 'which_provides_ protection', 'provides_protection_support', 'protection_support_individuals', 'support_individuals_institution', 'individuals_institutions_affected', 'institutions_affected_security', 'affected'_security_incidents', 'security'_incidents_includes', 'incidents_includes_assistance', 'includes_assistance_technical',

	<p>‘assistance’, ‘technical’, ‘support’, ‘evidence’, ‘retention’, ‘asset’, ‘confirmation’, ‘info’</p>	<p>_includes’, ‘includes_assistance’, ‘assistance_technical’, ‘technical_support’, ‘support_evidence’, ‘evidence_retention’, ‘retention_asset’, ‘asset_conf irm ation’, ‘confirmation_info’</p>	<p>‘assistance_technical_support’, ‘technical_support_evidence’, ‘support_evidence_retention’, ‘evidence_retention_asset’, ‘retention_asset_confirmation’ , ‘asset_confirmation_info’</p>
2	<p>‘kucoin’, ‘livestream’, ‘recap’, ‘latest’, ‘updates’, ‘about’, ‘security’, ‘incident’, ‘0930’, ‘details’</p>	<p>‘kucoin_livestream’, ‘livestream_recap’, ‘recap_latest’, ‘latest_updates’, ‘updates_about’, ‘about_security’, ‘security_incident’, ‘incident_0930’, ‘0930_ details’</p>	<p>‘kucoin_livestream_recap’, ‘livestream_recap_latest’, ‘recap_latest_updates’, ‘latest_updates_about’, ‘updates_about_security’, ‘about_security_incident’, ‘security_incident_0930’, ‘incident_0930_details’</p>
3	<p>‘trying’, ‘alts’, ‘rsr’, ‘fucking’, ‘forgot’, ‘kucoin’, ‘enabled’, ‘deposits’, ‘withdrawals’, ‘fuck’, ‘life’</p>	<p>‘trying_alts’, ‘alts_rsr’, ‘rsr_fucking’, ‘fucking_forgot’, ‘forgot_kucoin’, ‘kucoin_enabled’, ‘enabled_deposits’, ‘deposits_withdrawals’, ‘withdrawals_fuck’, ‘fuck_life’</p>	<p>‘trying_alts_rsr’, ‘alts_rsr_fucking’, ‘rsr_fucking_forgot’, ‘fucking_forgot_kucoin’, ‘forgot_kucoin_enabled’, ‘kucoin_enabled_deposits’, ‘enabled_deposits_withdrawals’, ‘deposits_withdrawals_fuck’, ‘withdrawals_fuck_life’</p>

4	<p>‘suspension’, ‘deposits’, ‘withdrawals’, nim’, ‘kucoin’, ‘nimiq’, ‘nim’, ‘pair’, ‘trading’, ‘competition’, ‘4000000’, ‘nim’, ‘won’, ‘postponed’, ‘please’, ‘stay’, ‘tuned’, ‘official’, ‘announcement’, ‘kucoin’</p>	<p>‘suspension_deposits’, ‘deposits_withdrawals’, ‘withdrawals_nim’, nim_kucoin’, ‘kucoin_nimiq’, ‘nimiq_nim’, ‘nim_pair’, ‘pair_trading’, ‘trading_competition’, ‘competition_4000000’, ‘4000000_nim’, ‘nim_won’, ‘won_postponed’, ‘postponed_please’, ‘please_stay’, ‘stay_tuned’, ‘tuned_official’, ‘official_announcement’, ‘announcement_kucoin’</p>	<p>‘suspension_deposits_withdrawals’, ‘deposits_withdrawals_nim’, ‘withdrawals_nim_kucoin’, ‘nim_kucoin_nimiq’, ‘kucoin_nimiq_nim’, ‘nimiq_nim_pair’, ‘nim_pair_trading’, ‘pair_trading_competition’, ‘trading_competition_4000000’, ‘competition_4000000_nim’, ‘4000000_nim_won’, ‘nim_won_postponed’, ‘won_postponed_please’, ‘postponed_please_stay’, ‘please_stay_tuned’, ‘stay_tuned_official’, ‘tuned_official_announcement’, ‘official_announcement_kucoin’</p>
5	<p>‘tokens’, ‘recovered’, ‘kucoin’, ‘hack’, ‘returned’, ‘kucoin’, ‘wallets’, ‘ocean’ ‘user’, ‘funds’, ‘reside’, ‘kucoin’, ‘togetherstronger’</p>	<p>‘tokens_recovered’, ‘recovered_kucoin’, ‘kucoin_hack’, ‘hack_returned’, ‘returned_kucoin’, ‘kucoin_wallets’, ‘wallets_ocean’, ‘ocean_user’ ‘user_funds’, ‘funds_reside’, ‘reside_kucoin’, ‘kucoin_’ togetherstronger’</p>	<p>‘tokens_recovered_kucoin’, ‘recovered_kucoin_kucoin’, ‘kucoin_hack_returned’, ‘hack_returned_kucoin’, ‘returned_kucoin_wallets’, ‘kucoin_wallets_ocean’, ‘wallets_ocean_user’, ‘ocean_user_funds’ ‘user_funds_reside’, ‘funds_reside_kucoin’, ‘reside_kucoin_togetherstronger’</p>

Selanjutnya adalah tahap mengkonversi dokumen ke dalam bentuk *dictionary*, tujuan dari tahap ini adalah menyimpan data token ke dalam bentuk *dictionary*, kemudian memberikan indeks yang berfungsi untuk mengidentifikasi kata tersebut.

Script yang penulis gunakan untuk proses konversi dokumen ke dalam *dictionary* dapat dilihat pada gambar 4.29

```
from gensim import corpora, models
# buat kamus.
dictionary = corpora.Dictionary(text_list)
dictionary.filter_extremes(no_below=1, no_above=0.2)
#nilai min utk integer adalah 1 dan desimal adalah 0.2
print(dictionary)
```

Gambar 4.32 Script Proses *Dictionary*

Proses dari LDA adalah tahapan yang dilakukan untuk membentuk jumlah topik. Terdapat dua hal penting yang menjadi perhatian dalam tahap ini, yaitu alur pemodelan topik dengan *Latent Dirichlet Allocation* dan eksperimen pemodelan topik. Script untuk proses pemodelan topik dengan LDA dapat dilihat pada gambar 4.30

```

doc_term_matrix = [dictionary.doc2bow(doc) for doc in text_list]
print(len(doc_term_matrix))
print(doc_term_matrix[1])
tfidf = models.TfidfModel(doc_term_matrix)
#proses pembuatan model TF-IDF
corpus_tfidf = tfidf[doc_term_matrix]
from gensim.models.coherencemodel import CoherenceModel
from gensim.models.ldamodel import LdaModel
from gensim.corpora.dictionary import Dictionary
from numpy import array
#fungsi untuk menghitung nilai coherence
def compute_coherence_values(dictionary, corpus, texts, limit, start,
step):
    coherence_values = []
    model_list = []
    for num_topics in range(start, limit, step):
        model = LdaModel(corpus=corpus, id2word=dictionary,
num_topics=num_topics, iterations=300)
        model_list.append(model)
        coherencemodel = CoherenceModel(model=model, texts=texts,
dictionary=dictionary, coherence='c_v')
        coherence_values.append(coherencemodel.get_coherence())

return model_list, coherence_values

```

Gambar 4.33 Script Permodelan Topik dengan LDA

Dalam menentukan jumlah model topik yang dilakukan dalam analisis *coherence score*. *Coherence score* digunakan untuk evaluasi model topik, semakin tinggi *coherence score* maka model yang dihasilkan semakin tinggi. Secara otomatis aplikasi akan membandingkan data modeling topik yang dihasilkan oleh LDA apa sudah terdaftar dalam database aplikasi. Apabila sudah terdaftar sistem akan melakukan proses penambahan data skor koheren ke dalam database. Proses penambahan data skor koheren hasil perhitungan topik modeling dengan menggunakan metode LDA dapat dilihat pada gambar 4.31

```

m = sql.connect(host="localhost", user="root", passwd="",
database='lda')
cursor = m.cursor()
b = "select * from data_lda where nama_crypto='{}' and
singkatan='{}'".format(nama, sing)
cursor.execute(b)
t = tuple(cursor.fetchall())

        for row in t:
            id = row[0]
            nama = str(row[1])
            singk = str(row[2])
            nilai_koheren= coherence_values

c = "INSERT INTO detail_tes_lda (id_tes, id, nilai_koheren) VALUES
(%s, %s, %s)"
val = (id_tes, id, nilai_koheren)
cursor.execute(c, val)
m.commit()
cursor.close()
m.close()
return redirect('/view_hasil_lda')

```

Gambar 4.34 Script Tambah Data Skor Koheren Topik Modeling

Tampilan hasil dari proses permodelan topik modeling dengan menggunakan metode LDA dapat dilihat pada gambar 4.32. Hasil dari proses perhitungan LDA ditampilkan dalam bentuk tabel agar mempermudah pengguna membaca hasil. Sistem akan secara otomatis akan melakukan proses sorting data nilai skor koheren perhitungan LDA.

Pada aplikasi proses penambahan atau edit data mata uang krypto akan disesuaikan dengan data yang ada dalam database aplikasi. Mata uang krypto yang terdaftar dalam sistem saja yang akan disimpan nilai skor koherennya. Sistem akan melakukan proses pencarian data yang sesuai dari hasil modeling topik LDA dan dibandingkan dengan database pada aplikasi.



ID	Nama Cryptocurrency	Nilai Koheren
1	Bitcoin	0.817
2	Ethereum	0.714
3	Tether	0.710
5	BNB	0.587
6	DOGE COIN	0.512
8	XRP	0.247

Gambar 4.35 Nilai Skor Koheren Perhitungan LDA

4.3.6 Halaman Hasil Proses Data *Stopword*

Proses *stop word* adalah proses filtering dengan memilih kata-kata yang dianggap mewakili dokumen dengan menghilangkan kata-kata yang dianggap tidak bermakna seperti, kata tanya, kata seru dan kata hubung. Tampilan dari hasil pemrosesan data *stop word* dapat dilihat pada gambar 4.33.



No	Isi Chat
1	Introducing Safeguard Program, which provides protection & support individuals and institutions affected security incidents. includes assistance technical support, evidence retention, asset confirmation more. More info: https://lanid.kucoin.com/gua
2	#KuCoin Livestream Recap - Latest Updates About Security Incident 0930 Details: https://kucoin.com/news/en-kucoin-ceo-livestream-recap-latest-updates-about-security-incident-0930...
3	IM TRYING ALTS \$RSR FUCKING FORGOT KUCOIN HASNT ENABLED DEPOSITS/WITHDRAWLS FUCK LIFE
4	suspension deposits withdrawals \$NIM #KuCoin, 'Nimiq (NIM) Pair Trading Competition, 4,000,000 NIM Won!' will postponed. Please stay tuned for the official announcement from KuCoin. Details: https://kucoin.com/news/en-nim-trading-competition-wil
5	Tokens recovered from KuCoin hack have been returned KuCoin wallets. OCEAN user funds reside with KuCoin. #TogetherStronger @lyu_johnny @kucoincom

Gambar 4.36 Hasil Pemrosesan Data *Stop Word*

4.3.7 Halaman Proses Data *Cleaning* dan *Folding*

Proses *cleaning* dan *folding* adalah membersihkan isi chat dari semua emoji, tanda baca, tagar, link html dan mengubah semua isi chat menjadi huruf kecil. Data proses *cleaning* dan *folding* didapatkan setelah sistem melakukan proses data *stop word* terlebih dahulu untuk selanjutnya dilakukan proses *cleaning* dan *folding* pada data isi chat pada tweeter. Tampilan data hasil *cleaning* dan *folding* dapat dilihat pada gambar 4.34.

**TOPIC MODELLING TERKAIT CRYPTOCURRENCY
PADA PLATFORM TWITTER MENGGUNAKAN METODE LDA
(LATENT DIRICHLET ALLOCATION)**
Jully Sulistyowati - 182102019

FAKULTAS TEKNIK & TEKNOLOGI INFORMASI
UNIVERSITAS JENDERAL AHMAD YANI
YOGYAKARTA
TAHUN 2022

Home Daftar User Data Cryptocurrency Proses Pengujian LDA Data Stopword Data Cleaning Data Tokenizing Data Stemming Hasil Modelling Logout

DATA HASIL PROSES CLEANING

No	Isi Chat
1	introducing safeguard program which provides protection support individuals and institutions affected security incidents includes assistance technical support evidence retention asset confirmation info
2	kucoin livestream recap latest updates about security incident 0930 details
3	trying alts rsr fucking forgot kucoin hasnt enabled deposits withdrawls fuck life
4	suspension deposits withdrawls nim kucoin. nimiq nim pair trading competition 4000000 nim won postponed please stay tuned official announcement kucoin
5	tokens recovered from kucoin hack been returned kucoin wallets ocean user funds reside kucoin togetherstronger

Gambar 4.37 Halaman Data Hasil Proses Cleaning dan Folding

4.3.8 Halaman Hasil Proses *Tokenizing*

Proses *tokenizing* merupakan proses memisahkan setiap kata yang menyusun suatu dokumen. Setelah proses *cleaning* dan *folding* selesai, tahap selanjutnya aplikasi akan memisahkan masing kata-kata pembentuk kalimat dalam chat menjadi kata yang berdiri sendiri. Tampilan halaman hasil proses *tokenizing* dapat dilihat pada gambar 4.35.

**TOPIC MODELLING TERKAIT CRYPTOCURRENCY
PADA PLATFORM TWITTER MENGGUNAKAN METODE LDA
(LATENT DIRICHLET ALLOCATION)**
Jully Sulistyowati - 182102019

FAKULTAS TEKNIK & TEKNOLOGI INFORMASI
UNIVERSITAS JENDERAL AHMAD YANI
YOGYAKARTA
TAHUN 2022

Home Daftar User Data Cryptocurrency Proses Pengujian LDA Data Stopword Data Cleaning Data Tokenizing Data Stemming Hasil Modelling Logout

DATA HASIL PROSES TOKENIZING

No	Isi Chat
1	'introducing', 'safeguard', 'program', 'which', 'provides', 'protection', 'support', 'individuals', 'institutions', 'affected', 'security', 'incidents', 'includes', 'assistance', 'technical', 'support', 'evidence', 'retention', 'asset', 'confirmati
2	'kucoin', 'livestream', 'recap', 'latest', 'updates', 'about', 'security', 'incident', '0930', 'details'
3	'trying', 'alts', 'rsr', 'fucking', 'forgot', 'kucoin', 'hasnt', 'enabled', 'deposits', 'withdrawls', 'fuck', 'life'
4	'suspension', 'deposits', 'withdrawals', 'nim', 'kucoin', 'nimiq', 'nim', 'pair', 'trading', 'competition', '4000000', 'nim', 'won', 'postponed', 'please', 'stay', 'tuned', 'official', 'announcement', 'kucoin'
5	'tokens', 'recovered', 'kucoin', 'hack', 'returned', 'kucoin', 'wallets', 'ocean', 'user', 'funds', 'reside', 'kucoin', 'togetherstronger'

Gambar 4.38 Halaman Hasil *Tokenizing*

4.3.9 Halaman Hasil Proses *Stemming*

Proses stemming merupakan proses lanjutan setelah sistem melakukan proses tokenizing. Pada proses *stemming*, aplikasi akan mengubah kata berimbuhan menjadi kata dasar. Hasil proses *stemming* dapat dilihat pada gambar 4.36.



No	Isi Chat
1	'introducing', 'safeguard', 'program', 'which', 'provides', 'protection', 'support', 'individuals', 'institutions', 'affected', 'security', 'incidents', 'includes', 'assistance', 'technical', 'support', 'evidence', 'retention', 'asset', 'confirmati
2	'kucoin', 'livestream', 'recap', 'latest', 'updates', 'about', 'security', 'incident', '0930', 'details'
3	'trying', 'alts', 'rsr', 'fucking', 'forgot', 'kucoin', 'hasn't', 'enabled', 'deposits', 'withdrawals', 'fuck', 'life'
4	'suspension', 'deposits', 'withdrawals', 'nini', 'kucoin', 'nimiq', 'nini', 'pair', 'trading', 'competition', '4000000', 'nini', 'won', 'postponed', 'please', 'stay', 'tuned', 'official', 'announcement', 'kucoin'
5	'tokens', 'recovered', 'kucoin', 'hack', 'returned', 'kucoin', 'wallets', 'ocean', 'user', 'funds', 'reside', 'kucoin', 'togetherstronger'

Gambar 4.39 Hasil Proses *Stemming*

4.4 Pembahasan

Sistem ini dapat mempermudah para penanam saham pemula yang ingin menginvestasikan dananya pada mata uang *crypto* dengan mengikuti perkembangan informasi pertumbuhan mata uang kripto yang menjadi topik pembicaraan pada sosial media Twitter guna mengurangi resiko kerugian finansial.

Sistem yang dibuat untuk mengelola data csv yang dihasilkan dari twitter ini memiliki kelebihan yaitu dapat menampilkan hasil dari data *stopword*, data *cleanning*, data *tokenizing* dan data *stemming* yang berguna untuk memproses persiapan untuk melakukan proses permodelan tweet terkait *cryptocurrenccy* serta menampilkan data dari hasil proses pengujian LDA. Sistem ini juga masih adanya kekurangan yaitu belum terdapat fitur grafik untuk lebih memudahkan dalam memonitoring para investor disetiap hari/bulannya