

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Identifikasi

4.1.1 Identifikasi *N-Gram* dan *NP Chunks*

Kumpulan *N-Gram* dan *NP- Chunks* ini didapat dari *dataset* Hulth beserta nilai *co-occurrence*-nya. Berikut ini pada table 4.1 adalah gambaran dari *dataset* dan nilainya.

Tabel 4.1 *dataset* untuk *N-Gram* dan *NP Chunks*.

olympus	0.809	0.501	1.459	-0.262	-0.509	-0.249	0.146
system	0.338	-0.237	0.771	0.208	-0.961	0.217	0.069
information	0.635	0.281	1.103	0.906	0.584	-0.666	-0.588

4.1.2 Identifikasi Data Latih

Data latih terdiri dari beberapa file yaitu :

4.1.2.1 Identifikasi Data Latih Judul dan Isi

Identifikasi data latih terdiri dari judul dan isi yang terdiri dari beberapa kalimat. Kalimat didapatkan dari beberapa *website* yang telah dilakukan *filtering* dan *preprocessing*. Akan tetapi hanya diambil pada *page* utama. Berikut ini pada tabel 4.2 bentuk data latih yang digunakan adalah sebagai berikut :

Tabel 4.2 Data latih

Title	Waiting for the wave to crest [wavelength services]
content	Wavelength services have been hyped ad nauseam for years. But despite their quick turn-up time and impressive margins, such services have yet to live up to the industry's expectations. The reasons for this lukewarm reception are many, not the least of
	which is the confusion that still surrounds the technology, but most industry observers are still convinced that wavelength services with ultimately flourish ...

Pada tabel 4.2 diatas pada kolom pertama yaitu *title*, *title* didapatkan dari *filtering* dan *preprocessing meta title*, sedangkan *content* didapatkan pada dalam `<body> .. </body>`.

4.1.2.2 *Keywords* Terstruktur

Keywords terstruktur adalah *keywords* yang berisi beberapa frase dan beberapa kata dari *keywords* yang telah dimasukan secara terstruktur dan dipisahkan oleh *semicolon* untuk masing - masing *keywords*.

Keywords : optical fibre networks; telecommunication

4.1.2.3 *Keywords* Tidak Terstruktur

Keywords tidak terstruktur adalah *keywords* yang berisi beberapa frase dan beberapa kata dari *keywords* yang telah dimasukan secara tidak terstruktur dan dipisahkan oleh *semicolon* untuk masing - masing *keywords*.

Keywords : wavelength services; fiber optic networks; looking glass networks; point east research

4.2. Hasil Dari Metode

4.2.1 Teknik *Grabbing* Text

Contoh *website* yang akan dicara kata kuncinya yaitu <https://www.guru99.com/>.

4.2.1.1 *Grabbing* link

Pertama mengambil semua link yang terkait dengan *website* tersebut. Adapun beberapa hasil dari *grabbing link* dapat dilihat pada tabel 4.3 yaitu:

Tabel 4.3 Daftar *Link*

Home	https://www.guru99.com/
Blog	https://www.guru99.com/blog/
Agile Testing	https://www.guru99.com/agile-testing-course.html

Tabel 4.3 Daftar *Link* (Lanjutan)

Database Testing	https://www.guru99.com/data-testing.html
Link Facebook	https://www.facebook.com/guru99com/
Link Twitter	https://twitter.com/guru99com

Jika dilihat pada tabel 4.3 yaitu *daftar link* , dari hasil yang didapatkan, maka ada beberapa link yang mengarah ke *social media, email dan link* yang di luar *website* yang ditargetkan. Sehingga perlu adanya pemfilteran yang dikhususkan hanya untuk link yang berada di *website* tersebut.

4.2.1.2 Hasil *grabbing* teks

Adapun salah satu hasil dari *grabbing* teks salah satu *link* diatas yaitu <https://www.guru99.com/>.

```
<html dir="ltr" lang="en-GB">
<head>
<meta content="width=device-width, initial-scale=1.0"
name="viewport"/>
<meta content="IE=edge" http-equiv="X-UA-Compatible">
<link href="/images/logo/favicon1.ico" rel="icon"
type="image/x-icon"/>
<meta charset="utf-8"/>
<base href="https://www.guru99.com/" />
<meta content="home, tutorials, and, for, to" name="keywords"/>
<meta content="width=device-width, initial-scale=1.0"
name="viewport"/>
<meta content=" All Courses include Learn courses from a pro.
Our experts are passionate teachers who share their sound
knowledge and rich experience with learners Variety of
tutorials and Quiz Interactive tutorials" name="description"/>
<meta content="Joomla! - Open Source Content Management">
```

```

name="generator"/>

<title>Meet Guru99 - Free Training Tutorials & Video for IT
Courses</title>

<link          href="/Customization/imagepopup/colorbox.min.css"
rel="stylesheet"/>

<link
href="/plugins/system/maximenuckmobile/themes/custom/maximenuck
mobile.css" rel="stylesheet"/>

<link
href="/modules/mod_maximenuck/themes/custom/css/maximenuck_maxi
menuck243.css" rel="stylesheet"/>

<link          href="/modules/mod_maximenuck/templatelayers/gantry-
navigation.css"          rel="stylesheet"/><link
href="/modules/mod_maximenuck/assets/maximenuckresponsiveck.css"
rel="stylesheet"/>

<style>
...
<h2 style="color:#173a41;font-size:31px;line-height:34px;">
  Guru99 is totally new kind of learning experience.</h2>

<br/><p style="color:#343434;text-align:justify;">Here you
learn by practice. We make tons of efforts to take boredom out
of learning and make education a fun experience.<br/><br/>

Inside, you will find tons of video tutorials</p>

<hr style="background-color:#343434; height:4px;"><h3
style="color:#343434;">All provided FREE!!!</h3>

</div>

...</body></html>

```

4.2.1.3 Filter dengan kalimat dan format utf-8

Setelah mendapatkan semua teks maka selanjutnya adalah memilih *meta tag* yang memiliki teks *utf-8* dan yang berisi sebuah kalimat. Maka jika diambil contoh dari atas akan terpilih kalimat seperti ini :

```

<h2>Guru99 is totally new kind of learning experience.</h2>
<p style="color:#343434;text-align:justify;">Here you learn by

```

```
practice. We make tons of efforts to take boredom out of
learning and make education a fun experience.<br/><br/>Inside,
you will find tons of video tutorials</p>
```

4.2.1.4 Menghilangkan *Tag*

Setelah melakukan pemilihan terhadap *tag* yang memiliki sebuah kalimat, maka selanjutnya adalah menghilangkan *tag* tersebut, sehingga akan menjadi kalimat seperti berikut :

```
Guru99 is totally new kind of learning experience.
Here you learn by practice. We make tons of efforts to take boredom
out of learning and make education a fun experience.Inside, you
will find tons of video tutorials.
```

4.2.1.5 Hasil *Meta Content Keywords* Pada *Website*

Setelah menghilangkan *tag* dan mendapatkan kumpulan dari kalimat – kalimat ataupun frase, maka jika dipisahkan antar *tag body* dan *meta content keywords*, akan terlihat pada kalimat atau frase untuk *meta content keywords* seperti dibawah ini :

```
Courses include Learn courses from a pro. Our experts are
passionate teachers who share their sound knowledge and rich
experience with learners Variety of tutorials and Quiz
Interactive tutorials
```

4.2.1.6 Tahap Terakhir Pada *Preprocessing*

Untuk tahap *preprocessing* pada sebuah *website* bisa dikatakan selesai, karena kalimat telah mudah untuk diidentifikasi oleh sistem.

4.2.2 Hasil Dari Metode *Semantic Relatedness*

Hasil yang didapatkan dari metode *semantic relatedness* adalah dapat mengetahui hubungn antar kata atau frasa beserta nilai hubungan atau keterkaitan

dengan kata atau frasa lainnya. Berikut adalah salah satu dataset yang digunakan dan telah berbentuk *array* dari sebuah kalimat :

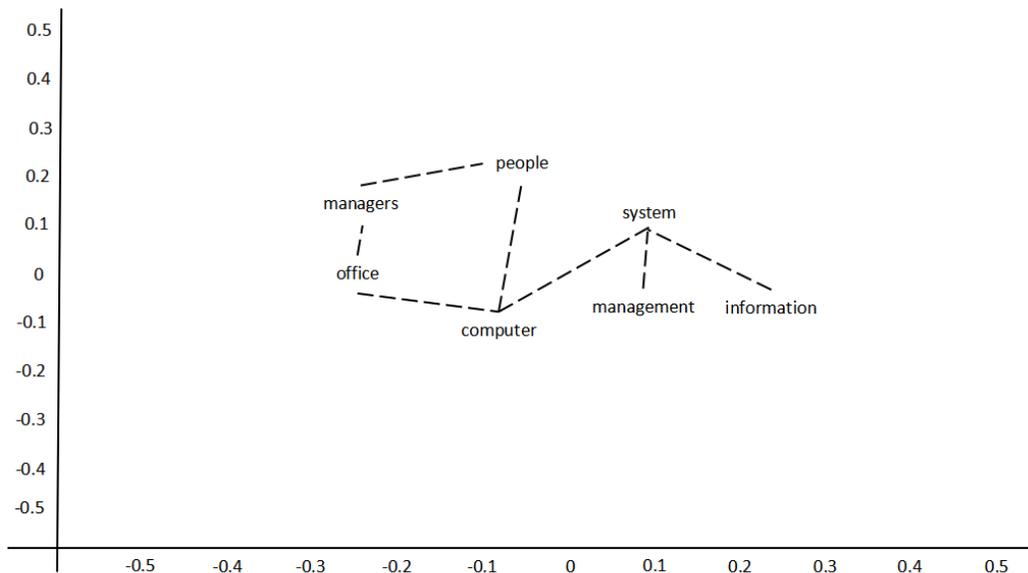
```
[[ 'management', 'information', 'system', 'is', 'a', 'computer', 'based', 'system', 'that', 'collectively', 'provides', 'managers', 'with', 'the', 'information', 'required', 'to', 'organize'], [ 'manage', 'and', 'analysis', 'the', 'ongoing', 'project', 'operation'], [ 'in', 'this', 'course', 'we', 'will', 'cover', 'important', 'management', 'information', 'system']]
```

Dari kalimat diatas yang telah dibagi - bagi berdasarkan kata dan frasanya, maka bisa didapatkan sebuah tabel hubungan semantik seperti pada tabel 4.4 dibawah ini :

Tabel 4.4 *Semantic Relations*

Words	Relations
<i>System</i>	<i>Computer</i>
<i>System</i>	<i>Management</i>
<i>System</i>	<i>Information</i>
<i>Computer</i>	<i>Office</i>
<i>Computer</i>	<i>People</i>

Pada tabel 4.4 diatas memperlihatkan bagaimana hubungan atau relasi semantic dari sebuah kata. Adapun jika digambarkan bagaimana hubungan kata tersebut, maka akan terlihat seperti pada gambar 4.1 yaitu sebagai berikut :



Gambar 4.1 Hubungan semantik

Pada gambar 4.1 digambarkan bagaimana hubungan semantik antar tiap kata terjadi. Dengan menggunakan *word to vector* dengan hubungan semantiknya dan menggunakan data dari *glove* maka ketika memasukan sebuah kata system akan menghasilkan nilai sebagai berikut :

[-0.09447385 0.0290945 0.04032194 -0.04793497 -0.04845755]

4.2.3 Hasil Dari Metode ANN

Berikut adalah hasil yang didapatkan untuk metode ANN tanpa gabungan dari metode *semantic relatedness*.

4.2.3.1 Load Dataset Untuk Metode ANN

Training dan *testing* dilakukan menggunakan data dari Hultth yang berisi data NP dan *Tag POS*. Dengan menggunakan data *training* sekitar 1000 *files* dan data *testing* 500 *files*.

4.2.3.2 Nilai *Precision*, *Recall*, *F1-Score*

Tabel 4.5 Hasil Akurasi dari metode ANN Pada Setiap *Epoch*

Epoch	Precision	Recall	F1	Acc	Loss	Execution Time (s)
1/28	0.0	0.0	0.0	0.9613	2.2581	328
2/28	0.0	0.0	0.0	0.7505	1.3412	402
3/28	0.0	0.0	0.0	0.1483	1.1206	418
4/28	0.0	0.0	0.0	0.9613	1.0989	415
5/28	0.0	0.0	0.0	0.7200	1.0988	415
6/28	0.0	0.0	0.0	0.7413	1.0989	417
7/28	0.0	0.0	0.0	0.7498	1.0989	420
8/28	0.0	0.0	0.0	0.8704	1.0989	420
9/28	0.0	0.0	0.0	0.9310	1.0989	418
10/28	0.0033	0.1985	0.0064	0.2910	1.0989	417
11/28	0.0033	0.1985	0.0064	0.0479	1.0989	424
12/28	0.0	0.0	0.0	0.0185	1.0989	424
13/28	0.0	0.0	0.0	0.6827	1.0989	425
14/28	0.0	0.0	0.0	0.5698	1.0989	423
15/28	0.0	0.0	0.0	0.2370	1.0989	427
16/28	0.0	0.0	0.0	0.7491	1.0989	427
17/28	0.0033	0.1985	0.0064	0.2381	1.0989	428
18/28	0.0033	0.1985	0.0064	0.3207	1.0989	429
19/28	0.0033	0.1985	0.0064	0.1689	1.0988	429

Tabel 4.5 Hasil Akurasi dari metode ANN Pada Setiap *Epoch* (Lanjutan)

20/28	0.0	0.0	0.0	0.0855	1.0989	434
21/28	0.0	0.0	0.0	0.9613	1.0989	435
22/28	0.0	0.0	0.0	0.9613	1.0989	434
23/28	0.0	0.0	0.0	0.2964	1.0989	439
24/28	0.0	0.0	0.0	0.8931	1.0989	435
25/28	0.0	0.0	0.0	0.4789	1.0989	439
26/28	0.0	0.0	0.0	0.3283	1.0989	440
27/28	0.0	0.0	0.0	0.4185	1.0989	439
28/28	0.0033	0.1985	0.0064	0.8707	1.0989	443

Dari tabel 4.5 hasil yang diperoleh setelah melakukan pelatihan dan dilakukan testing serta validasi dengan jumlah *epoch* sampai dengan 28 maka banyak terjadi nilai yang kosong pada beberapa *precision*, *recall* dan *F1-Score*, kemudian nilai akurasi pun semakin lama semakin kecil.

4.2.3.3 Hasil Ekstraksi Kata Kunci Dari Metode ANN 10 Teratas

Adapun hasil dari rekomendasi ekstraksi kata kunci yang diambil 10 besar teratas dengan hasil sebagai berikut ini :

['twenty'], ['years'], ['of'], ['the'], ['literature'], ['on'], ['acquiring'], ['out-of-print'], ['materials'], ['this']

Adapun bentuk kata yang dihasilkan diatas masih dalam bentuk *array*, jika kata kunci dituliskan dalam teks biasa akan menjadi sebagai berikut : *twenty, years, of, the, literature, on, acquiring, out of print, materials, this.*

4.2.3.4 Hasil Akhir Dan Total Keseluruhan Dari Metode ANN

Setelah selesai pelatihan, testing dan validasi maka yang terakhir adalah mengkalkulasikan hasil akurasi dari keseluruhan *epoch*. Berikut ini adalah hasil dari akurasi keseluruhan.

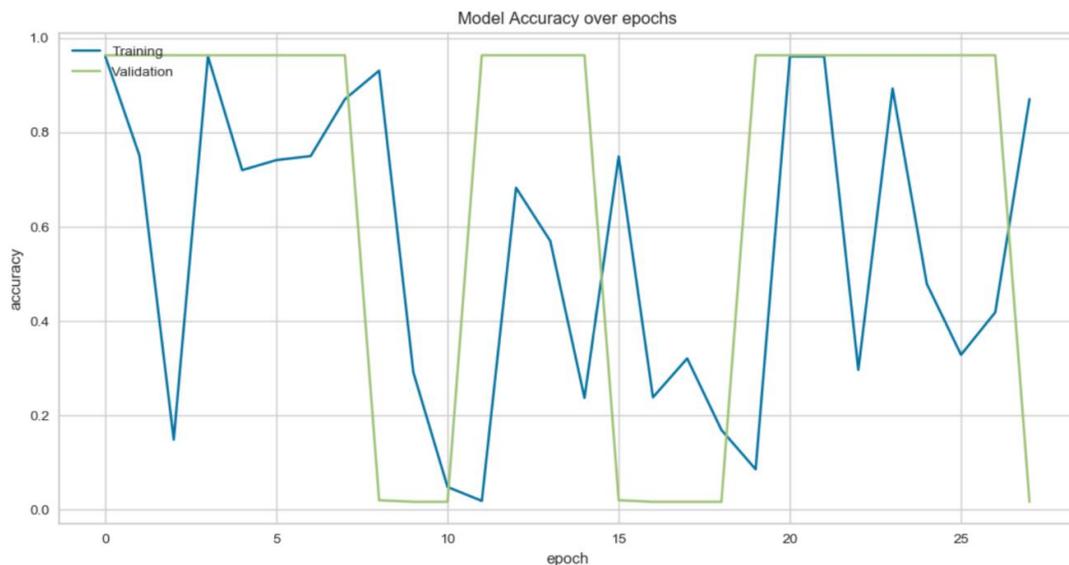
Tabel 4.6 Hasil Akurasi Keseluruhan

<i>Precision</i>	0.0138
<i>Recall</i>	0.1105
<i>F1-Score</i>	0.0246

Pada tabel 4.6 hasil akhir setelah melakukan *training*, maka menghasilkan nilai *precision* sebesar 0.0138, dengan *recall* 0.1105 dan *F1-Score* yaitu 0.0246.

4.2.3.5 Model Accuracy Over Epochs Dari Metode ANN

Adapun jika digambarkan dalam grafik, untuk hasil dari *model accuracy* dengan jumlah *epoch* sampai dengan 28 adalah seperti pada gambar 4.2.

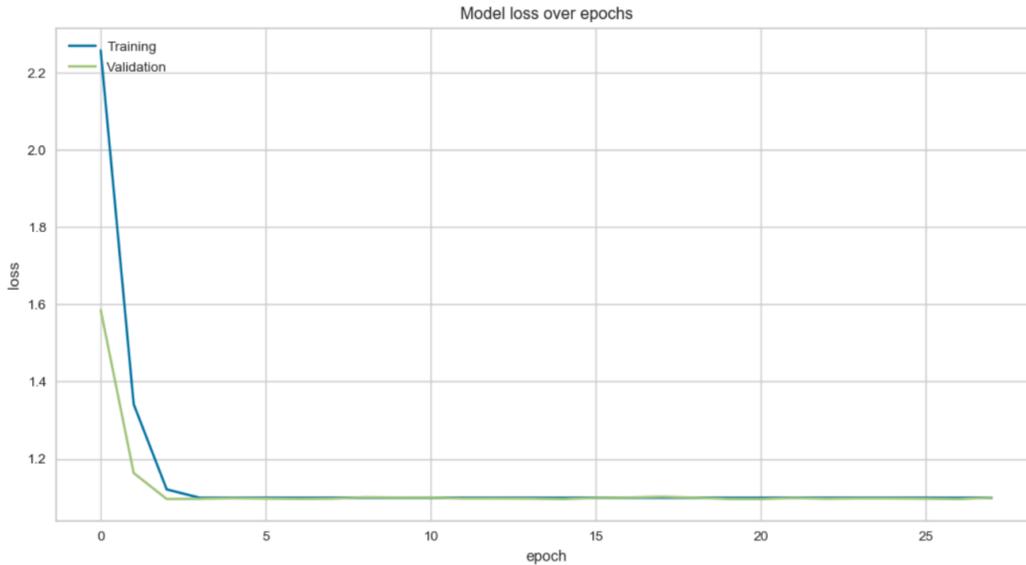


Gambar 4.2 Model Accuracy Over Epoch

Pada gambar 4.2 yaitu model *accuracy over epoch* memperlihatkan bahwa antara nilai *training* dan *validation* sangat tidak beraturan.

4.2.3.6 Model Loss Over Epochs Dari Metode ANN

Adapun jika digambarkan dalam grafik, untuk hasil dari *model loss* dengan jumlah *epoch* sampai dengan 28 adalah seperti pada gambar 4.3.

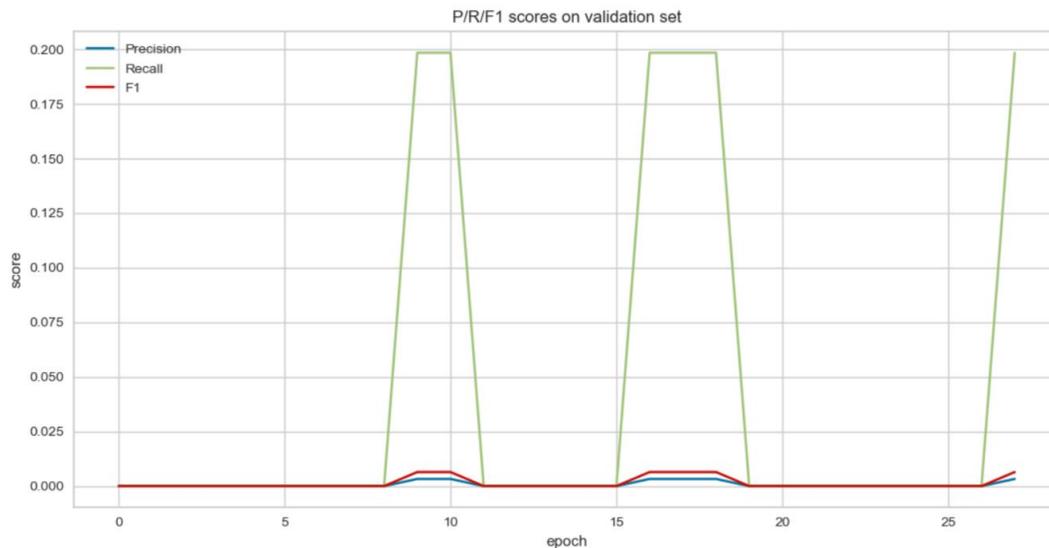


Gambar 4.3 Model Loss Over Epoch

Dari gambar 4.3 Model Loss Over Epoch memperlihatkan *loss* yang terjadi dibawah 1.2.

4.2.3.7 Nilai Precision, Recall dan F1-Score Dari Metode ANN

Setelah memperlihatkan hasil dari *accuracy* dan *loss* maka selanjutnya adalah memperlihatkan bagaimana hasil dari nilai *precision*, *recall* dan *F1-Score* jika digambarkan dalam bentuk grafik. Berikut adalah gambar 4.4 yaitu hasil nilai *precision*, *recall* dan *F1-Score*.



Gambar 4.4 Precision Recall F1 Scores On Validation Set

Dari gambar 4.4 Precision, recall, f1-scores memperlihatkan bahwa nilai recall beberapa kali tinggi, akan tetapi untuk nilai precision dan F1-Score tetap selalu rendah.

4.2.4 Hasil Hybrid Metode Semantic Relatedness dan Metode ANN

Berikut ini adalah hasil yang gabungan dari metode *semantic relatedness* yang menggunakan *glove* sebagai *corpus*-nya dan digabungkan dengan metode ANN. *Training* dan *testing* dilakukan masih menggunakan data dari Hulth yang berisi data NP dan *Tag POS*. *NP Chunks* memberikan presisi yang lebih baik dari pada *n-gram*.

Dengan menggunakan 554 *tokens* atau *dictionary* sebagai *noun phrase* yang diambil dari *glove* dan data *training* sekitar 1000 *files* serta data *testing* 500 *files*.

4.2.4.1 Load Dataset Untuk Hybrid Semantic Relatedness dan ANN

Adapun *capture* pada saat *load dataset* pertama pada gambar 4.5 yaitu sebagai berikut :

```

C:\Python36\Scripts\deepkeyphraseextraction-master>c:\Python36\python.exe SimpleNN.py
Using TensorFlow backend.
2019-07-08 17:45:32,462 INFO      Keras version 2.0.5
2019-07-08 17:45:32,462 INFO      Numpy version 1.16.0
2019-07-08 17:45:32,463 INFO      Tensorflow version 1.2.1
2019-07-08 17:45:32,463 INFO      Loading dataset...
2019-07-08 17:45:32,463 DEBUG     Initialized dataset Hulth, 2003 from folder data/Hulth2003
2019-07-08 17:45:37,351 DEBUG     Loaded training set for dataset Hulth, 2003
2019-07-08 17:45:40,492 DEBUG     Loaded test set for dataset Hulth, 2003
2019-07-08 17:45:41,737 DEBUG     Loaded validation set for dataset Hulth, 2003
2019-07-08 17:45:44,719 INFO      Dataset loaded. Preprocessing data...
2019-07-08 17:45:44,901 DEBUG     Fitting dictionary on 2000 documents...
2019-07-08 17:45:45,028 DEBUG     Dictionary fitting completed. Found 18328 unique tokens
2019-07-08 17:45:45,102 DEBUG     Longest training document : 554 tokens
2019-07-08 17:45:45,103 DEBUG     Longest test document : 380 tokens
2019-07-08 17:45:45,103 DEBUG     Longest validation document : 314 tokens
2019-07-08 17:45:45,301 DEBUG     Training set samples size : (1000, 550)
2019-07-08 17:45:45,301 DEBUG     Training set answers size : (1000, 550, 3)
2019-07-08 17:45:45,302 DEBUG     Test set samples size : (500, 550)
2019-07-08 17:45:45,302 DEBUG     Test set answers size : (500, 550, 3)
2019-07-08 17:45:45,302 DEBUG     Validation set samples size : (500, 550)
2019-07-08 17:45:45,302 DEBUG     Validation set answers size : (500, 550, 3)
path_glove:C:\Python36\Scripts\deepkeyphraseextraction-master\utils
2019-07-08 17:45:45,302 DEBUG     path_glove : C:\Python36\Scripts\deepkeyphraseextraction-master\utils
2019-07-08 17:45:45,303 DEBUG     Loading glove pre-trained embeddings from glove.6B.300d.txt
2019-07-08 17:46:32,718 DEBUG     Total embeddings found: 400000.
2019-07-08 17:46:32,718 DEBUG     Building embedding matrix of size [18329,300]...
2019-07-08 17:46:33,061 INFO      Data preprocessing complete.
2019-07-08 17:46:33,171 INFO      Maximum possible recall: 0.7691837980867088
2019-07-08 17:46:33,172 DEBUG     Building the network...

```

Gambar 4.5 Load Dataset

Pada gambar 4.5 memperlihatkan saat pertama kali *load dataset glove*, *data training*, *data testing* dan *data validation*.

4.2.4.2 Kalkulasi Untuk Total Frase Yang Dimasukan Untuk *Hybrid Semantic Relatedness* dan ANN

Kemudian setelah *load dataset* maka ditotalkan dan mendapatkan jumlah kata - kata serta frasenya yaitu seperti pada gambar 4.6 dibawah ini :

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 550, 300)	5498700
bidirectional_1 (Bidirection	(None, 550, 600)	1442400
dropout_1 (Dropout)	(None, 550, 600)	0
time_distributed_1 (TimeDist	(None, 550, 150)	90150
dropout_2 (Dropout)	(None, 550, 150)	0
time_distributed_2 (TimeDist	(None, 550, 3)	453
=====		
Total params:	7,031,703	
Trainable params:	1,533,003	
Non-trainable params:	5,498,700	

Gambar 4.6 Total Trainable

Gambar 4.6 memperlihatkan kalkulasi total dari jumlah kata atau frase yang dimasukan, dengan total yaitu 1.533.003.

4.2.4.3 Nilai *Precision*, *Recall*, *F1-Score* Untuk *Hybrid Semantic Relatedness* dan ANN

Selanjutnya setelah melakukan kalkulasi awal, maka dilakukan pelatihan mulai dari *epoch* pertama sampai *epoch* terakhir. Tiap masing – masing *epoch* memperlihatkan berapa nilai *precision*, *recall* dan *F1-Score* yang dihasilkan, seperti pada gambar 4.7 dibawah ini :

```

None
2019-07-08 17:46:35,026 INFO Fitting the network...
Train on 1000 samples, validate on 500 samples
Epoch 1/28
992/1000 [=====>.] - ETA: 2s - loss: 2.2416 - acc: 0.7570
### Validation Scores ###
###
### Epoch : 1
### Precision : 0.0390
### Recall : 0.1203
### F1 : 0.0589
###
1000/1000 [=====] - 380s - loss: 2.2352 - acc: 0.7575 - val_loss: 1.2144 - val_acc: 0.8625
Epoch 2/28
992/1000 [=====>.] - ETA: 2s - loss: 1.0076 - acc: 0.8635
### Validation Scores ###
###
### Epoch : 2
### Precision : 0.1598
### Recall : 0.5991
### F1 : 0.2523
###
1000/1000 [=====] - 420s - loss: 1.0059 - acc: 0.8634 - val_loss: 0.5507 - val_acc: 0.9193
Epoch 3/28
992/1000 [=====>.] - ETA: 2s - loss: 0.5289 - acc: 0.8937
### Validation Scores ###
###
### Epoch : 3
### Precision : 0.1689
### Recall : 0.5786
### F1 : 0.2614
###

```

Gambar 4.7 *Validation Score Per-epoch*

Pada gambar 4.7 menggambarkan nilai dari *precision*, *recall*, *F1-Score* dengan menggunakan data *validation*. Gambar 4.7 hanya mengambil *sample* dari *epoch* ke 1 sampai dengan *epoch* ke 3. Adapun untuk lebih detail per masing – masing epoch yaitu dituliskan pada tabel 4.7 sebagai berikut :

Tabel 4.7 Hasil Akurasi Pada Setiap *Epoch*

Epoch	Precision	Recall	F1	Acc	Loss	Execution Time (s)
1/28	0.0390	0.1203	0.0589	0.7575	2.2352	380
2/28	0.1598	0.5991	0.2523	0.8634	1.0059	420
3/28	0.1689	0.5786	0.2614	0.8939	0.5273	424

Tabel 4.7 Hasil akurasi untuk masing - masing *epoch* (Lanjutan)

4/28	0.1786	0.6172	0.2770	0.9010	0.3666	425
5/28	0.1769	0.6190	0.2751	0.9079	0.3268	431
6/28	0.1897	0.6594	0.2946	0.9109	0.2869	437
7/28	0.2132	0.6819	0.3248	0.9146	0.2665	434
8/28	0.2032	0.7046	0.3154	0.9178	0.2504	448
9/28	0.1782	0.6691	0.2815	0.9202	0.2343	435
10/28	0.2172	0.6733	0.3285	0.9232	0.2207	429
11/28	0.2367	0.6550	0.3477	0.9265	0.2046	428
12/28	0.2533	0.6642	0.3668	0.9281	0.1867	431
13/28	0.1938	0.6711	0.3008	0.9302	0.1808	429
14/28	0.2378	0.6775	0.3521	0.9326	0.1658	430
15/28	0.2331	0.7038	0.3502	0.9355	0.1493	440
16/28	0.2259	0.6982	0.3414	0.9381	0.1398	432
17/28	0.2825	0.6587	0.3954	0.9400	0.1272	437
18/28	0.2150	0.6762	0.3262	0.9421	0.1204	443
19/28	0.2722	0.6625	0.3858	0.9434	0.1131	456
20/28	0.1827	0.6254	0.2828	0.9460	0.1050	497
21/28	0.2667	0.6801	0.3831	0.9481	0.1014	442
22/28	0.3102	0.6490	0.4198	0.9489	0.0979	636
23/28	0.2817	0.6649	0.3957	0.9517	0.0920	753
24/28	0.3029	0.6552	0.4143	0.9529	0.0935	784

Tabel 4.7 Hasil akurasi untuk masing - masing *epoch* (Lanjutan)

25/28	0.2524	0.5389	0.3437	0.9556	0.0831	823
26/28	0.2719	0.6552	0.3843	0.9584	0.0824	852
27/28	0.2872	0.6494	0.3983	0.9581	0.0805	814
28/28	0.3628	0.5651	0.4419	0.9617	0.0713	814

Pada tabel 4.7 memperlihatkan nilai *precision*, *recall*, *F1- Score*, *accuracy*, *loss* dan *execution time* pada setiap *epoch*, sampai dengan *epoch* ke 28.

4.2.4.4 Hasil Rekomendasi Ekstraksi Kata Kunci 10 Teratas Untuk *Hybrid Semantic Relatedness* dan ANN

Adapun hasil dari rekomendasi ekstraksi kata kunci yang diambil 10 besar teratas yaitu seperti pada gambar 4.8.

```
'413': [['on-line', 'materials'], ['vertical', 'spacing'], ['web-based', 'instruction'], ['treatment', 'groups'], ['white', 'space'], ['web', 'm
aterials'], ['overall', 'structure'], ['presentation', 'elements'], ['instructors'], ['instructional']], '414': [['electronic', 'telecommunicatio
ns'], ['computer-mediated', 'communications'], ['student', 'attitudes'], ['instructor'], ['interpersonal', 'relationships'], ['fostering', 'inter
personal', 'relationships'], ['students'], ['e-mail'], ['education'], ['fostering']]]
```

Gambar 4.8 Ekstraksi Kata Kunci

Pada gambar 4.8 memperlihatkan hasil rekomendasi ekstraksi kata kunci. Adapun secara detail dapat dilihat pada baris dibawah ini :

```
[[['electronic', 'telecommunications'], ['computer-mediated', 'communications'],
['student', 'attitudes'], ['instructor'], ['interpersonal', 'relationships'], ['fostering',
'interpersonal', 'relationships'], ['students'], ['e-mail'], ['education'], ['fostering']]]
```

Adapun bentuk kata yang dihasilkan diatas masih dalam bentuk *array*, jika kata kunci dituliskan dalam teks biasa akan menjadi sebagai berikut : *electronic telecommunications, computer-metiated communications, student attitudes, instructor, interpersonal relationships, fostering interpersonal relationships, students, e-mail, education, fostering.*

4.2.4.5 Hasil Akhir Dan Total Keseluruhan Untuk *Hybrid Semantic Relatedness* dan ANN

Adapun untuk hasil terakhir atau total dari keseluruhan setelah melakukan *training* sampai dengan *epoch* yang terakhir, dalam hal ini dimasukan epoch sampai dengan 28, maka menghasilkan nilai *precision*, *recall* dan *F1-Score* seperti pada tabel 4.6 yaitu sebagai berikut :

Tabel 4.8 Hasil Akurasi Keseluruhan

<i>Precision</i>	0.3567
<i>Recall</i>	0.4773
<i>F1-Score</i>	0.4083

Pada tabel 4.8 hasil akhir setelah melakukan training, maka menghasilkan nilai *precision* sebesar 0.35, dengan *recall* 0.47 dan *F1-Score* yaitu 0.40. Adapun jika dilakukan capture maka akan terlihat seperti pada gambar 4.6 yaitu sebagai berikut:

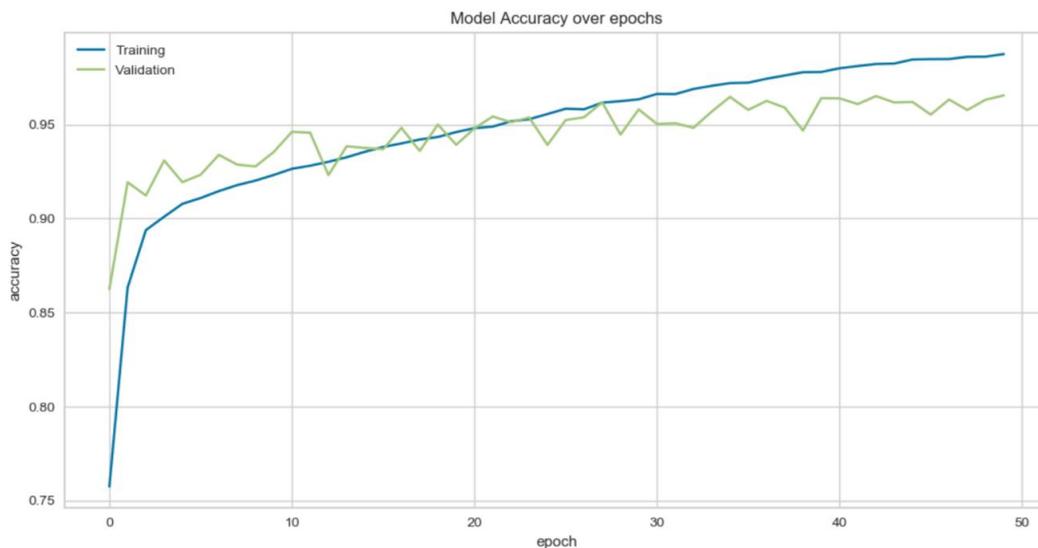
```
2019-07-08 23:34:22,184 DEBUG Shape of output array: (500, 550, 3)
### Obtained Scores ###
### (full dataset) ###
###
### Precision : 0.3567
### Recall : 0.4773
### F1 : 0.4083
###
```

Gambar 4.9 Full Dataset

Pada gambar 4.9 adalah hasil *capture* yang memperlihatkan nilai dari *precision*, *recall* dan *F1-Score*.

4.2.4.6 Model Accuracy Over Epochs Untuk *Hybrid Semantic Relatedness* dan ANN

Adapun jika digambarkan dalam grafik, untuk hasil dari *model accuracy* dengan jumlah *epoch* sampai dengan 50 adalah seperti pada gambar 4.9.

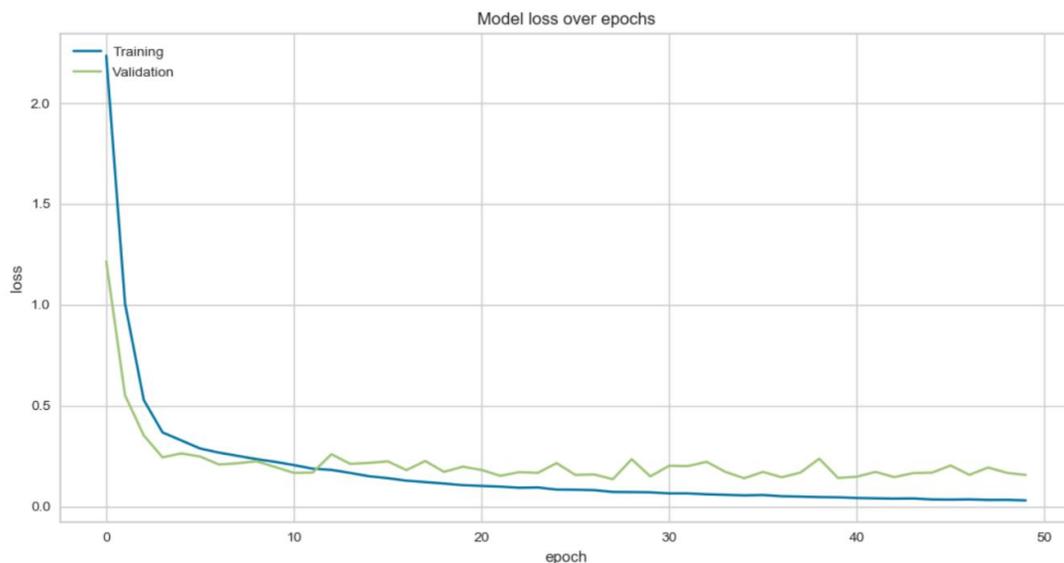


Gambar 4.10 Model Accuracy Over Epochs

Pada gambar 4.10 yaitu model *accuracy over epochs* menyebutkan bahwa semakin tinggi nilai *epoch* yang dimasukkan maka, semakin tinggi pula akurasi karena datanya yang disimpan di memori semakin besar, akan tetapi antara nilai validasi dan *training* yang baik adalah pada *epoch* ke 28 yaitu dengan nilai 0.93.

4.2.4.7 Model Loss Over Epochs Untuk Hybrid Semantic Relatedness dan ANN

Adapun jika digambarkan dalam grafik, hasil dari *Model Loss* dengan jumlah *epoch* 50 adalah seperti pada gambar 4.11.

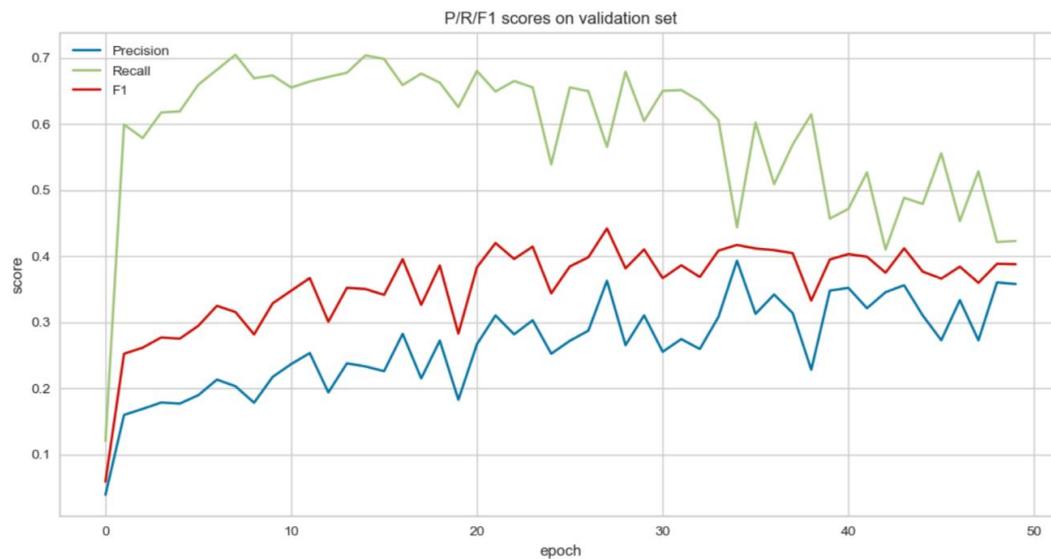


Gambar 4.11 *Model Loss Over Epochs*

Pada gambar 4.11 dapat disimpulkan bahwa semakin tinggi nilai *epoch* yang dimasukan maka akan semakin rendah pula *loss* yang terjadi, akan tetapi nilai *epoch* yang terbaik pada *epoch* ke 28, dikarenakan pada *epoch* ke 28 antara data validasi dan data *training* hampir sama atau berdekatan.

4.2.4.8 Nilai *Precision*, *Recall* dan *F1-Score* Untuk *Hybrid Semantic Relatedness* dan ANN

Setelah memperlihatkan hasil dari *accuracy* dan *loss* maka selanjutnya adalah memperlihatkan bagaimana hasil dari nilai *precision*, *recall* dan *F1-Score* jika digambarkan dalam bentuk grafik. Berikut adalah gambar 4.12 yaitu hasil nilai *precision*, *recall* dan *F1-Score*.



Gambar 4.12 *Precision Recall F1 Scores On Validation Set*

Dari gambar 4.12 yaitu *precision, recall dan F1 score* pada *validation set* memperlihatkan bahwa pada *epoch* ke 28 dan 34 nilai dari *precision, recall* dan *f1 score* saling mendekat akan tetapi nilainya dari data validasi masih dibawah 0.5 persen.