

FINAL ROJECT

KELOMPOK 6

Big Data kelas A

*[Big Data Analysis : Analisa Perubahan Iklim, dan Hubungannya dengan Penyakit di Indonesia
“Juni – Oktober 2018”]*

Anggota :

Muhammad Musta'in

Muzzammil Fadli

Nashrul Fatah

Novi Dwiasih

Nurul Hestinintyas

A. PENDAHULUAN

Bumi satu-satunya tempat tinggal bagi manusia kini telah banyak berubah. Perubahan itu terjadi dikarenakan banyak hal ada faktor alam dan faktor manusia. Salah satu yang terjadi adalah perubahan iklim. Iklim adalah rata-rata cuaca dimana cuaca merupakan keadaan atmosfer pada suatu saat di waktu tertentu. Iklim juga didefinisikan sebagai ukuran rata-rata dan variabilitas kuantitas yang relevan dari variabel tertentu (seperti temperatur, curah hujan atau angin), pada periode waktu tertentu.

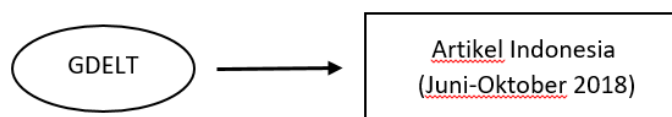
Perubahan iklim merupakan fenomena perubahan jangka panjang dalam distribusi pola cuaca secara statistik sepanjang periode waktu. Perubahan cuaca ini dapat disebabkan baik secara langsung atau tidak langsung oleh aktivitas manusia sehingga mengubah komposisi dari atmosfer global dan variabilitas iklim alami pada periode waktu yang dapat diperbandingkan (Konvensi PBB tentang Kerangka Kerja Perubahan Iklim dalam ditjenppi.menlhk.go.id). Iklim berubah secara terus menerus karena interaksi antara komponen-komponennya dan faktor eksternal seperti erupsi vulkanik, variasi sinar matahari, dan faktor-faktor disebabkan oleh kegiatan manusia seperti misalnya perubahan penggunaan lahan dan penggunaan bahan fosil.

Perubahan iklim ini berpengaruh kepada banyak aspek seperti ketahanan pangan, ekonomi, kesehatan, sosial, dan politik. Dalam ketahanan pangan, perubahan iklim mempengaruhi hasil pertanian dan peternakan. Dalam kesehatan perubahan iklim yang terjadi dari titik ekstrem ke titik ekstrem lainnya berpengaruh pada tubuh manusia. Masih banyak lagi contoh kasus lain yang disebabkan oleh perubahan iklim. Semua yang terjadi tersimpan dalam data sehingga dapat kita analisis bagaimana dampak perubahan iklim terhadap kehidupan manusia

B. METODE PENELITIAN

1. Dataset

Dataset yang didapatkan dari GDELT (Global Database of Events, Language, and Tone) berupa parameter/unsur yang terdapat dalam suatu artikel. Dalam proyek ini hanya difokuskan pada peristiwa yang terjadi di Indonesia, dengan memilih kode Negara (ID) pada periode waktu Juni-Oktober 2018.



Dari kumpulan artikel peristiwa di Indonesia pada periode tersebut dilakukan filter untuk memilih artikel yang membahas mengenai Perubahan Iklim (Climate Change). Beberapa parameter dari GDELT yang akan digunakan untuk analisis perubahan iklim diantaranya :

- | | |
|--------------------|--------------------------|
| a. SQLDATE, | g. ActionGeo_FullName |
| b. Actor1Code, | h. ActionGeo_CountryCode |
| c. Actor1Name | i. ActionGeo_Lat |
| d. Actor1Type1Code | j. ActionGeo_Long |
| e. QuadClass | k. SOURCEURL |
| f. GoldsteinScale | |

DATE	Actor1Code	Actor1Name	Actor1Type1Code	QuadClass	GoldsteinScale	ActionGeo_FullName	ActionGeo_CountryCode	ActionGeo_Lat	ActionGeo_Long	
70901	COP	PRISON	COP	4	-5.0	Jakarta, Jakarta Raya, Indonesia	ID	-6.17444	106.829	http://vn
30901	None	None	None	1	2.8	Tamansari, Jawa Tengah, Indonesia	ID	-6.9556	110.057	http://vn
30901	None	None	None	1	1.0	Jakarta, Jakarta Raya, Indonesia	ID	-6.17444	106.829	http://vn
30901	None	None	None	1	7.0	Jakarta, Jakarta Raya, Indonesia	ID	-6.17444	106.829	http://vn
30901	None	None	None	4	-5.0	Jakarta, Jakarta Raya, Indonesia	ID	-6.17444	106.829	http://vn
30901	None	None	None	1	2.8	Tamansari, Jawa Tengah, Indonesia	ID	-6.9556	110.057	http://vn
30901	COP	POLICE CHIEF	COP	1	1.9	Lokasari, Jawa Barat, Indonesia	ID	-7.0588	108.42	http://vn

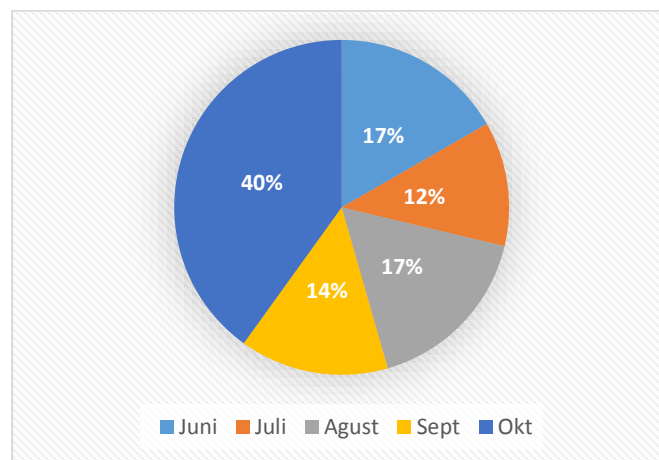
Tabel 1. Contoh Dataset dari GDELT

Iklim/Unsur	Penyebab	Proses	Dampak	Mitigasi
Atmosfer: - Oksigen - Karbon - Karbondioksida - Karbonmonoksida - Ozon - Sinar matahari - Cuaca - Suhu - Celsius - Curah Hujan - Awan - Uap - Global Warming	Bumi : - Motor - Pembakaran - Alam - Hutan - Erupsi - Gunung api - Polusi - Efek rumah kaca - Plastik	- Kondensasi - Peningkatan - Penurunan - Dingin - Panas	- Makhluk hidup - Hewan - Manusia - Tumbuhan - Lingkungan - Laut - Air - Banjir - Kekeringan - Kutub - Tenggelam - Pulau	- Energi alternative - Hemat - Penghijauan - Pencegahan - Forum

Tabel 2. Pengelompokan kata kunci dalam analisis perubahan iklim

2. Eksplorasi Data

Data diambil dari GDELT berupa sekumpulan url yang berisi berita pada periode Juni-Oktober 2018, kemudian diambil berita yang berkaitan dengan tema yang telah ditentukan. Total berita yang didapat yaitu sebanyak 62763 berita dengan rata-rata 91 pemberitaan per harinya. Berikut ditampilkan persentase berita perbulan dibanding total berita yang diambil.



Gambar 1. Pie chart data berita per bulan

Berdasarkan Gambar 1, berita yang berkaitan dengan tema paling banyak ditemukan pada bulan Oktober. Sedangkan untuk bulan lainnya cenderung sama terkait dengan jumlah pemberitaan yang terkait dengan tema yang diteliti. Dalam penelitian ini digunakan lima variabel yaitu :

a. Global Warming

Variabel yang berisi banyaknya berita yang mengandung unsur global warming. *Keyword* yang digunakan untuk menentukan suatu berita masuk variabel global warming adalah *pollution, deforestation, greenhouse, plastic waste, ilegal logging, carbon dioxide, methane, emission, emissions, global warming*.

b. Penyebab Alami

Variabel yang berisi banyaknya berita yang mengandung unsur penyebab alami terjadinya perubahan iklim. *Keyword* yang digunakan untuk menentukan suatu berita masuk variabel penyebab alami adalah *volcanic eruptions, volcanic eruption, volcano eruption, eruption, volcano eruptions*.

c. Perubahan Iklim

Variabel yang berisi banyaknya berita yang mengandung unsur perubahan iklim. *Keyword* yang digunakan untuk menentukan suatu berita masuk variabel perubahan iklim adalah *flood damage, flood damages, extreme drought, climate change, climate changes, desertification, extreme weather*.

d. Mitigasi

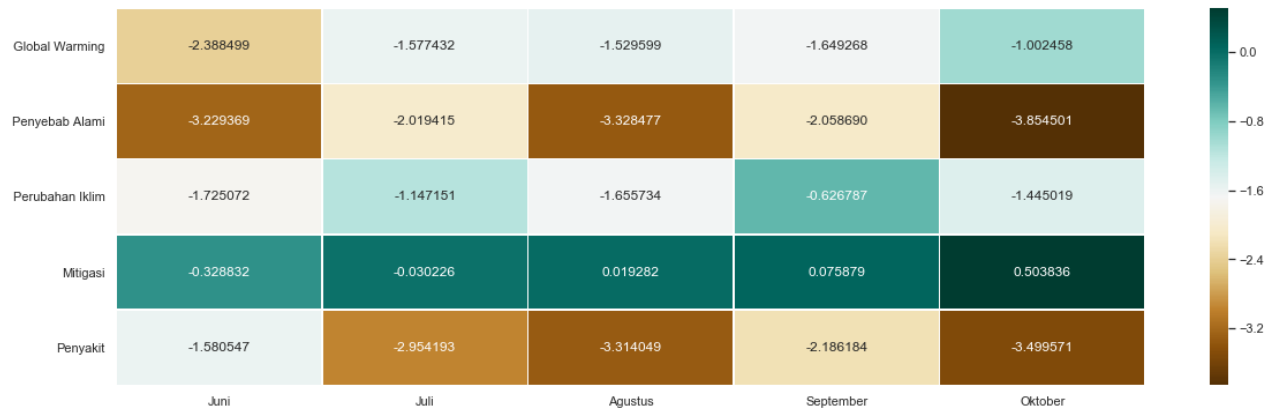
Variabel yang berisi banyaknya berita yang mengandung unsur mitigasi. *Keyword* yang digunakan untuk menentukan suatu berita masuk variabel mitigasi adalah *alternative energy, renewable energy, energy efficiency, greening, reforestation, greenpeace, eco-, waste reduction, sustainability, sustainable development, sustainable developments, renewable resource, renewable resources, kyoto protocol*.

e. Penyakit

Variabel yang berisi banyaknya berita yang mengandung unsur penyakit. *Keyword* yang digunakan untuk menentukan suatu berita masuk variabel penyakit adalah *epidemic, endemic, disease, virus, dengue fever, mumps, zika, mosquito-borne, contaminated water, influenza*.

Berita berdasarkan Tonasi

Berikut diberikan gambar plot berita tonasi rata-rata variabel/label bulan Juni-Oktober 2018



Gambar 2. Headmap Tonasi rata-rata per bulan

Dalam eksplorasi tonasi ini (averageTone) angka-angka diatas menunjukan rata-rata dari semua dokumen yang mengandung satu atau lebih dokumen/artikel yang menyebutkan “topic” tertentu selama pembaruan 15 menit dimana pertama kali dilihat. Rentang nilai pada umumnya antara -10 dan +10, dengan 0 (nol) menunjukan netral.

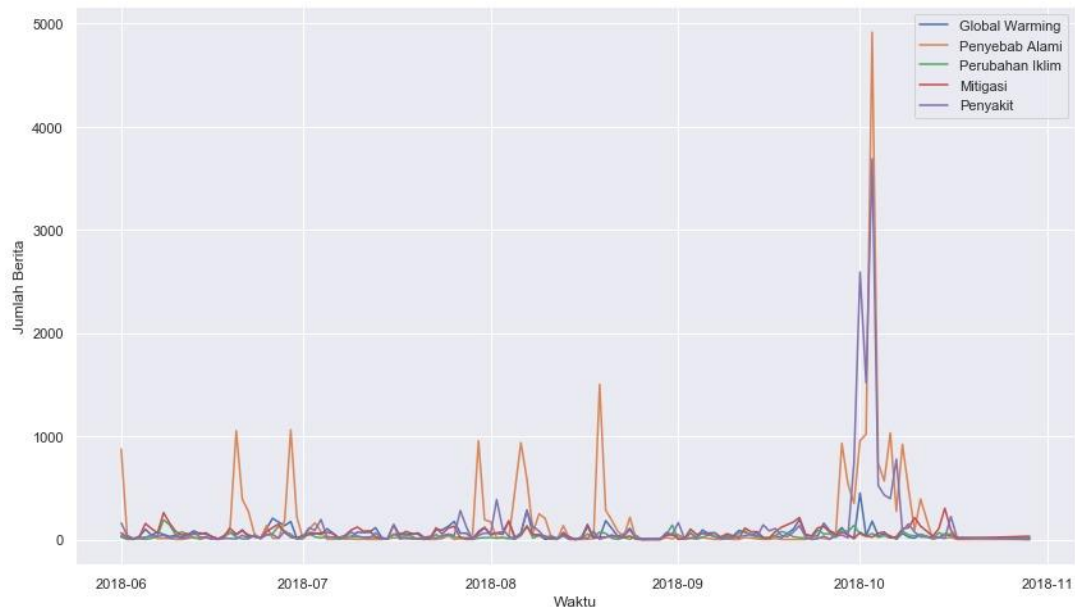
Angka tonasi dapat digunakan sebagai metode filter konteks dari peristiwa dan sebagai wakil untuk dampak dari peristiwa tersebut. Angka dengan tonasi lebih negative (-) menunjukan peristiwa yang lebih serius.

Pada Gambar 2, *keyword* dengan warna yang semakin cokelat memiliki peran lebih penting/serius dalam kejadian perubahan iklim dan paling sering dalam pemberitaan. Dimana perubahan iklim dipengaruhi oleh penyebab alami dengan angka tonasi tinggi (-3,85401) pada bulan Oktober serta memiliki dampak terhadap *keyword* yang berhubungan dengan adanya penyakit pada angka tonasi (-3,499571) juga di bulan Oktober.

Keyword yang berhubungan dengan perubahan iklim yang sering dibahas adalah mengenai penyebab dan dampak (timbulnya penyakit). Untuk *keyword* yang berhubungan dengan mitigasi perubahan iklim mengalami penurunan dalam pemberitaan maupun aksi di lapangan dalam periode 5 bulan (Juni-Oktober 2018).

Berita berdasarkan variabel/label

Berikut diberikan plot banyaknya berita berdasarkan kategori berita dari semua variabel/label yang digolongkan.



Gambar 3. Jumlah pemberitaan pada bulan Juni-Oktober 2018



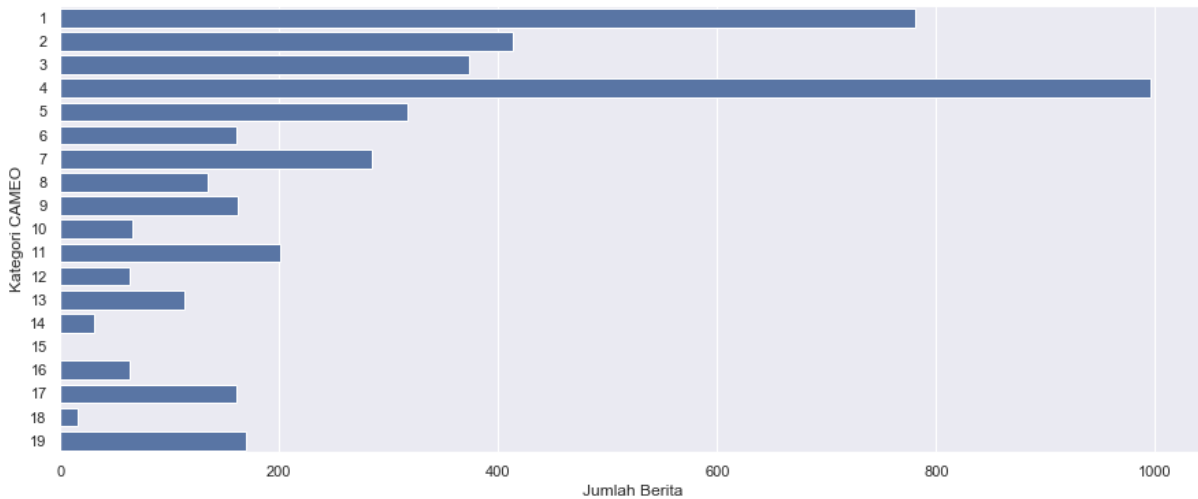
Gambar 4. Jumlah pemberitaan per variabel/label pada bulan Juni-Oktober 2018

Berdasarkan Gambar 3, terdapat pemberitaan pada variabel penyebab alami yang menjadi *trending* karena banyak diberitakan. Salah satu berita yang populer yaitu terjadinya gempa bumi di Palu yang meningkatkan pemberitaan tentang penyakit arena kejadian alam yaitu gempa bumi, tsunami mengakibatkan terjadinya peningkatan penyakit seperti malnutrisi pada anak-anak, diare, demam, dan lain-lain.

Berdasarkan Gambar 4, nampak bahwa kejadian gempa bumi di Palu pada bulan Oktober menaikkan intensitas pemberitaan tentang Indonesia dengan label *global warning*, penyebab alami, perubahan iklim, dan penyakit namun tidak memberikan pengaruh terhadap pemberitaan dengan label mitigasi.

Berita berdasarkan kategori CAMEO

Berikut diberikan plot banyaknya berita berdasarkan kategori berita dari semua label yang digolongkan berdasarkan CAMEO code yang ada pada GDELT

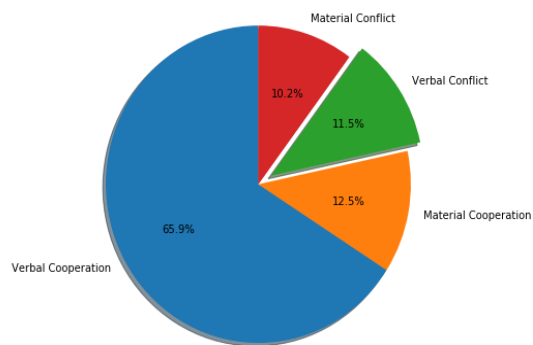


Gambar 5. Plot banyaknya berita per kategori CAMEO code

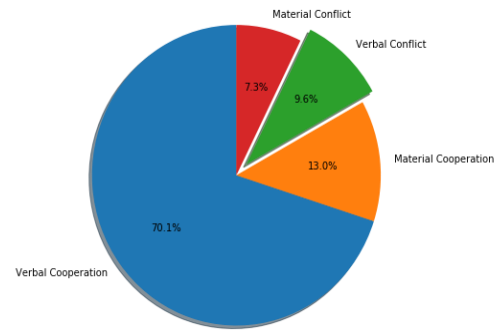
Berdasarkan Gambar 5 nampak bahwa kategori CAMEO 4 adalah yang paling banyak jumlah beritanya yaitu hampir 1000 berita dibandingkan kategori yang lain. Artinya bahwa kategori 4 menjadi kategori yang paling populer dalam rentang bulan Juni-Oktober 2018 jika dibandingkan dengan yang lain.

Berikut diberikan *pie chart* sebaran kategori berita penanganan perubahan iklim per bulan.

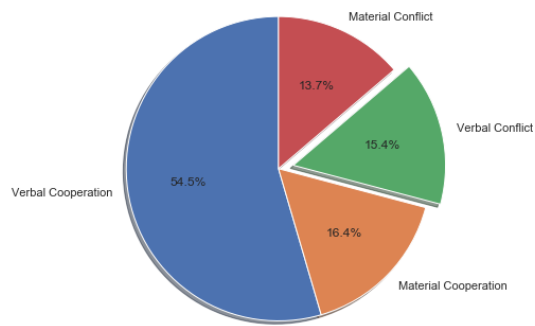
Sebaran Kategori Penanganan Perubahan Iklim bulan Agustus



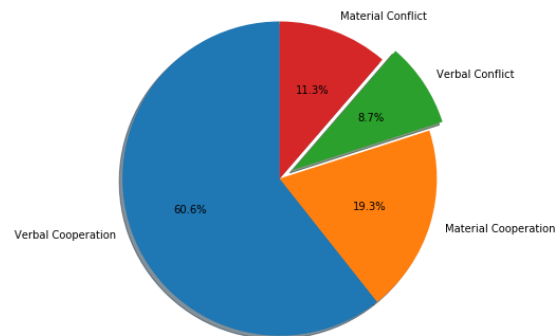
Sebaran Kategori Penanganan Perubahan Iklim bulan September



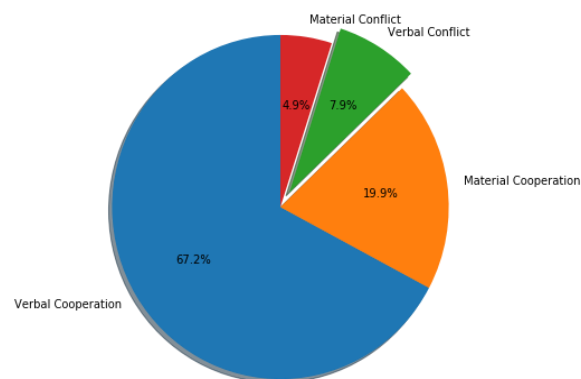
Sebaran Kategori Penanganan Perubahan Iklim bulan Juni



Sebaran Kategori Penanganan Perubahan Iklim bulan Juli



Sebaran Kategori Penanganan Perubahan Iklim bulan oktober



Gambar 6. *Pie chart* sebaran kategori berita penanganan perubahan iklim per bulan.

Berdasarkan Gambar 6 nampak bahwa kerjasama verbal paling banyak digunakan dalam penanganan perubahan iklim pada periode Juni-Oktober.

3. Rumusan Masalah

Research question adalah sebuah pertanyaan yang jelas, terfokus, singkat, kompleks dan diperdebatkan dan menjadi pertanyaan utama dari riset. Dalam penelitian ini *research question* yang digunakan adalah sebagai berikut

- a. Faktor apa yang paling berpengaruh terhadap perubahan iklim di Indonesia?
- b. Apakah perubahan iklim mempengaruhi penyakit?
- c. Apa faktor yang paling berpengaruh dalam perubahan iklim yang dapat mempengaruhi timbulnya penyakit?
- d. Apa jenis penyakit yang sering muncul akibat perubahan iklim?
- e. Bagaimana prediksi pengaruh perubahan iklim terhadap penyakit?
- f. Apa langkah yang sebaiknya dilakukan sebagai tindak lanjut dari hasil prediksi yang telah dilakukan?

C. MODEL PEMBELAJARAN MESIN

Sebagaimana yang telah diketahui, kolom “NumArticles” dalam dataset GDELT memuat jumlah pemberitaan yang dapat dijadikan tolok ukur untuk menentukan apakah suatu kejadian memiliki tingkat signifikansi yang tinggi. Karakteristik seperti yang tercantum dalam Codebook GDELT, semakin banyak pemberitaannya, maka berita tersebut semakin penting. Oleh karena itu, untuk pembelajaran mesin yang akan diterapkan dalam GDELT ini salah satunya memperhatikan jumlah pemberitaan tersebut. Data yang sudah dikelompokkan dalam label tertentu selanjutnya dianalisis. Metode pertama yang dilakukan adalah regresi linear sederhana, dengan memperhatikan jumlah pemberitaan mengenai “Perubahan Iklim” sebagai variabel dependen, dan global warming serta penyebab alami sebagai variabel independen.

1. Analisis Linear Regresi Sederhana

Sebelum dilakukan analisis regresi linear sederhana, dilakukan analisis korelasi untuk variabel independen dan variabel dependen yang akan dimasukkan dalam model regresi. Diperhatikan bahwa jumlah pemberitaan yang terdapat dalam semua kolom (label) tersebut merupakan data dalam runtun waktu. Dengan demikian, untuk mencari korelasinya, yang dibandingkan adalah percentage of change (perubahan level) dari jumlah pemberitaan tersebut.

2. Model Runtun Waktu

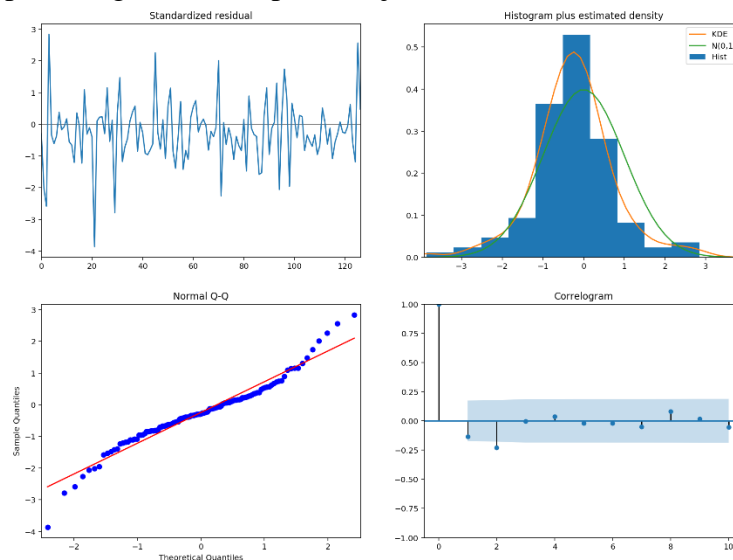
Dalam pemodelan runtun waktu ini, data yang digunakan adalah data *Average Tone* dari pemberitaan Global Warming dari dataset GDELT. Dataset yang digunakan telah diperiksa tidak memiliki data yang hilang. Data yang tidak lengkap dapat menyebabkan akurasi analisis model ini menjadi kurang baik.

Selanjutnya, dalam memilih metode *time series* (runtun waktu) yang tepat adalah dengan mempertimbangkan pola data, sehingga metode yang paling tepat dengan pola tersebut data diuji. Teknik yang bisa digunakan apabila data stasioner : *Naïve*, *Simple averaging*, *Moving average methods*, *Autoregressive integrated moving average (ARIMA)*. Dalam penelitian ini digunakan metode *ARIMA (Auto Regressive Integrated Moving Average)*.

3. Model Diagnostic Checking

Residual dari model dihitung sebagai perbedaan antara nilai aktual dan nilai yang *fitted*: $e_i = y_i - \hat{y}_i$. Setiap residu adalah komponen observasi yang tidak dapat diprediksi. Ketika data merupakan *time series*, harus melihat plot ACF residual. Plot ACF residual akan memperlihatkan jika ada autokorelasi di residual (menunjukkan bahwa ada informasi yang belum diperhitungkan dalam model).

Outliers menunjukkan ada sesuatu yang tidak biasa terjadi. Akan sangat berharga untuk menyelidiki pencilan itu untuk melihat apakah ada keadaan atau kejadian yang tidak biasa. **Histogram** baik untuk memeriksa apakah residu terdistribusi secara normal. Seperti yang dijelaskan sebelumnya, ini tidak penting untuk meramalkan, tetapi itu membuat perhitungan interval prediksi jauh lebih mudah.

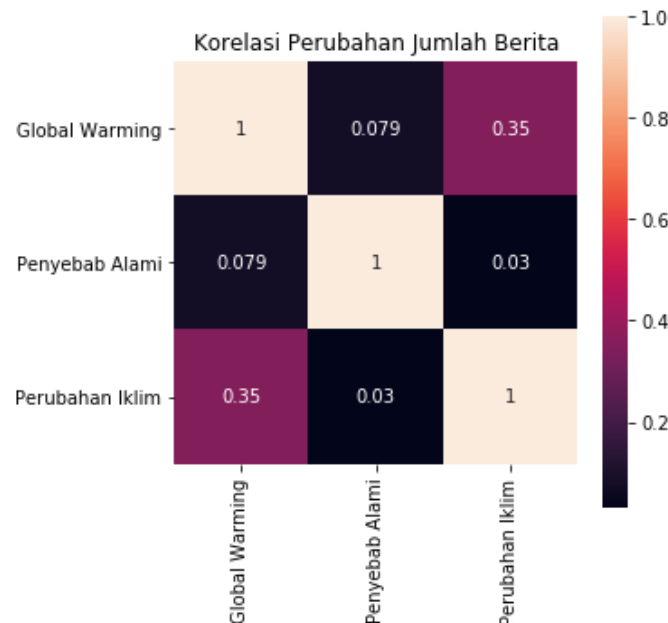


Gambar 7. Analisa Outliers dan Histogram

D. HASIL

1. Analisis Linear Regresi Sederhana

Berdasarkan analisis linear regresi sederhana diperoleh tabel grafik korelasi antar variabelnya adalah sebagai berikut.



Gambar 8. korelasi antar variabel

Dengan memperhatikan korelasi pada Gambar 7 diatas, diperoleh kesimpulan bahwa ternyata antar variabel tidak berpengaruh sangat signifikan terhadap kejadian perubahan iklim. Korelasi paling tinggi adalah dari kejadian pemanasan global yang memberikan nilai Pearson Correlation sebesar 0,35. Dengan demikian, sebetulnya dapat secara tidak langsung disimpulkan bahwa kejadian pemanasan global dan kejadian perubahan iklim berhubungan positif, meskipun tidak terlalu kuat.

Selanjutnya, dilakukan analisis regresi linear sederhana untuk menentukan model regresi serta melihat koefisien yang paling mempengaruhi kejadian perubahan iklim.

Dengan menggunakan library *statmodels* di Python, diperoleh keluaran untuk model regresinya adalah sebagai berikut :

Dep. Variable:	Perubahan Iklim	R-squared:	0.085
Model:	OLS	Adj. R-squared:	0.071
Method:	Least Squares	F-statistic:	6.223
Date:	Tue, 27 Nov 2018	Prob (F-statistic):	0.00260
Time:	13:01:21	Log-Likelihood:	-667.81
No. Observations:	137	AIC:	1342.
Df Residuals:	134	BIC:	1350.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	24.0002	3.827	6.272	0.000	16.431	31.569
Global Warming	0.1421	0.047	3.010	0.003	0.049	0.235
Penyebab Alami	0.0056	0.006	0.983	0.328	-0.006	0.017

Omnibus:	75.772	Durbin-Watson:	1.437
Prob(Omnibus):	0.000	Jarque-Bera (JB):	295.693
Skew:	2.095	Prob(JB):	6.18e-65
Kurtosis:	8.851	Cond. No.	748

Gambar 9. Model Regresi

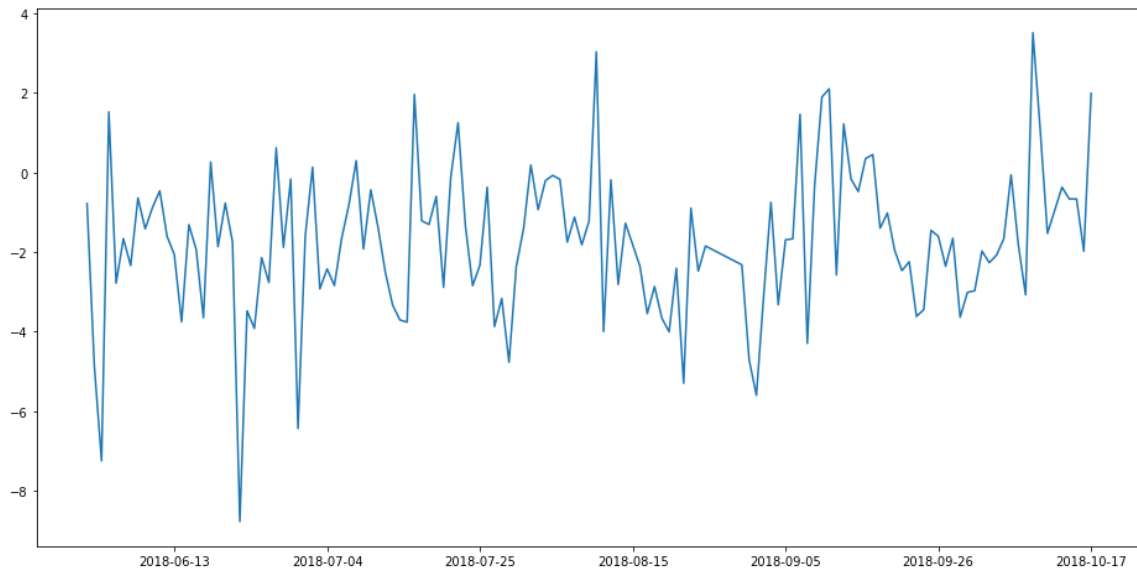
Diperoleh bahwa koefisien dari kolom pemanasan global ternyata hanya berpengaruh sebesar 0,1421 terhadap perubahan iklim. Nilai R-squared dari model regresi ini sendiri ternyata terlampau kecil, yaitu hanya 0,085. Artinya, 8,5% kejadian perubahan iklim dipengaruhi oleh pemanasan global dan penyebab alami (bencana alam), sedangkan sisanya yaitu 91,5% dipengaruhi oleh faktor lain yang belum diperhatikan di dalam model. Adapun model regresi linear sederhana yang diperoleh dari output tersebut adalah

$$PI = 24.0002 + 0.1421 \times GW + 0.0056 \times PA \quad (1)$$

dengan *PI* adalah kejadian perubahan iklim, *GW* adalah global warming, dan *PA* adalah penyebab alami.

2. Model Runtun Waktu

Berikut adalah plot data runtun waktu dari jumlah kejadian Global Warming.



Gambar 10. Grafik data runtun waktu dari jumlah kejadian Global Warming

Dapat diperhatikan bahwa data observasi berubah-ubah sekitar tingkatan atau rata-rata yang konstan disebut pola horizontal. Tipe ini pada data runtun waktu disebut stationer dalam rata-rata. Data stasioner didefinisikan sebagai data yang nilai rata-ratanya tidak berubah dari waktu ke waktu atau dapat dikatakan data bersifat stabil.

Uji data stationer menggunakan *ARIMA (Auto Regressive Integrated Moving Average)* dilakukan menggunakan Phyton dengan syntax :

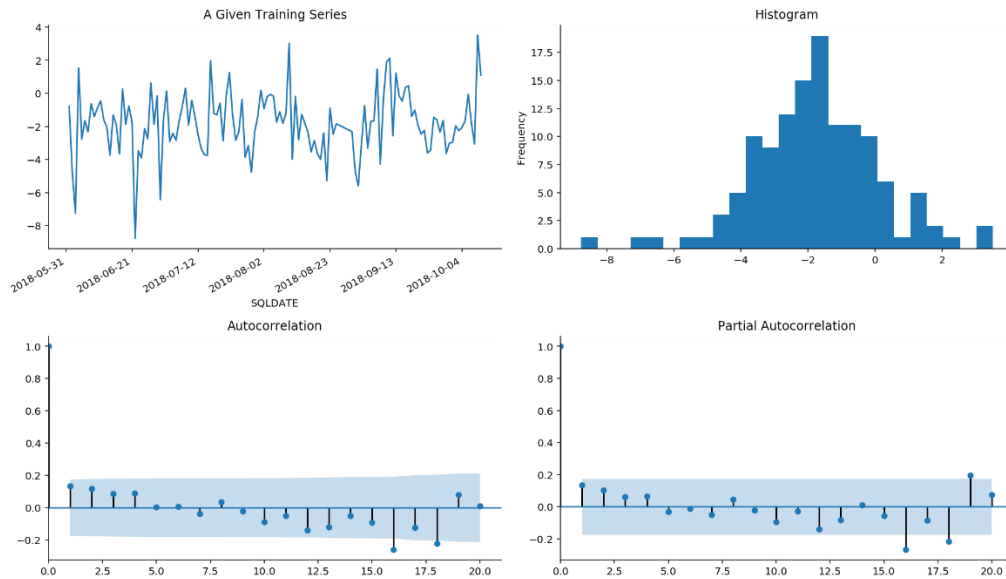
```

1 %load_ext autoreload
2 %autoreload 2
3 %matplotlib inline
4 %config InlineBackend.figure_format='retina'
5 import warnings; warnings.simplefilter('ignore')
6
7 from __future__ import absolute_import, division,
  print_function
8
9 import sys
10 import os
11
12 # TSA from Statsmodels
13 import statsmodels.api as sm
14 import statsmodels.formula.api as smf
15 import statsmodels.tsa.api as smt
16
17 pd.set_option('display.float_format', lambda x: '%.5f' % x) #
  pandas
18 np.set_printoptions(precision=5, suppress=True) # numpy
19
20 pd.set_option('display.max_columns', 100)
21 pd.set_option('display.max_rows', 100)

```

Gambar 11. Syntax ARIMA

Data yang dianalisis selanjutnya di-split menjadi data *training* dan data *test*. Selanjutnya, data *training* diplot runtun waktu, histogram, autokorelasi, dan autokorelasi parsialnya.



Gambar 12. Plot data training

Hasil dari model ARIMA yang dianalisis dari datanya adalah sebagai berikut.

Dep. Variable:	AvgTone	No. Observations:	127			
Model:	SARIMAX(2, 0, 0)	Log Likelihood	-271.190			
Date:	Tue, 27 Nov 2018	AIC	548.379			
Time:	19:02:53	BIC	556.912			
Sample:	0	HQIC	551.846			
- 127						
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.3488	0.083	4.183	0.000	0.185	0.512
ar.L2	0.3615	0.099	3.655	0.000	0.168	0.555
sigma2	4.1697	0.376	11.080	0.000	3.432	4.907
Ljung-Box (Q):	47.98	Jarque-Bera (JB):	24.91			
Prob(Q):	0.18	Prob(JB):	0.00			
Heteroskedasticity (H):	0.64	Skew:	-0.03			
Prob(H) (two-sided):	0.15	Kurtosis:	5.17			

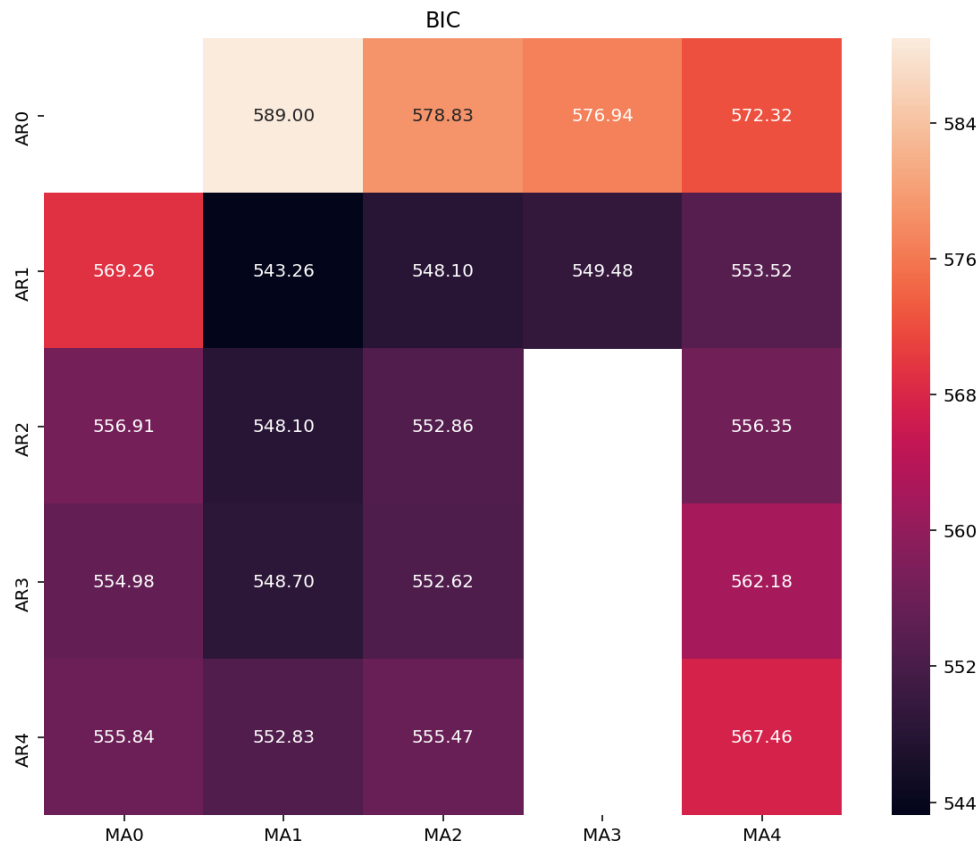
Gambar 13. Hasil Model ARIMA

Selanjutnya, dicari lebih dari beberapa model menggunakan petunjuk visual dari plotting sebelumnya sebagai titik awal.

```
1 import itertools
2
3 p_min = 0
4 d_min = 0
5 q_min = 0
6 p_max = 4
7 d_max = 0
8 q_max = 4
9
10 # Initialize a DataFrame to store the results
11 results_bic = pd.DataFrame(index=['AR{}'.format(i) for i in
12 range(p_min,p_max+1)],
13                             columns=['MA{}'.format(i) for i in
14 range(q_min,q_max+1)])
15
16 for p,d,q in itertools.product(range(p_min,p_max+1),
17                                range(d_min,d_max+1),
18                                range(q_min,q_max+1)):
19     if p==0 and d==0 and q==0:
20         results_bic.loc['AR{}'.format(p), 'MA{}'.format(q)] =
21         np.nan
22         continue
23
24     try:
25         model = sm.tsa.SARIMAX(data_train, order=(p, d, q),
26                                #enforce_stationarity=False,
27                                #enforce_invertibility=False,
28                                )
29         results = model.fit()
30         results_bic.loc['AR{}'.format(p), 'MA{}'.format(q)] =
31         results.bic
32     except:
33         continue
34 results_bic = results_bic[results_bic.columns].astype(float)
```

Gambar 14. Syntax model visualisasi dari plotting data

Berikut merupakan model dari hasil visualisasi Bayesian Information Criterion (BIC) :



Gambar 15. Model dari hasil visualisasi Bayesian Information Criterion

Metode pemilihan model alternatif, terbatas hanya mencari parameter AR dan MA

```
1 train_results = sm.tsa.arma_order_select_ic(data_train,
2 ic=['aic', 'bic'], trend='nc', max_ar=4, max_ma=4)
3 print('AIC', train_results.aic_min_order)
4 print('BIC', train_results.bic_min_order)
```

```
AIC (3, 1)
BIC (1, 1)
```

3. Model Diagnostic Checking

Beberapa tes (*formal testing*) yang dilakukan menghasilkan keluaran sebagai berikut.

Test heteroskedasticity of residuals (breakvar): stat=0.641, p=0.154

Test normality of residuals (jarquebera): stat=24.906, p=0.000

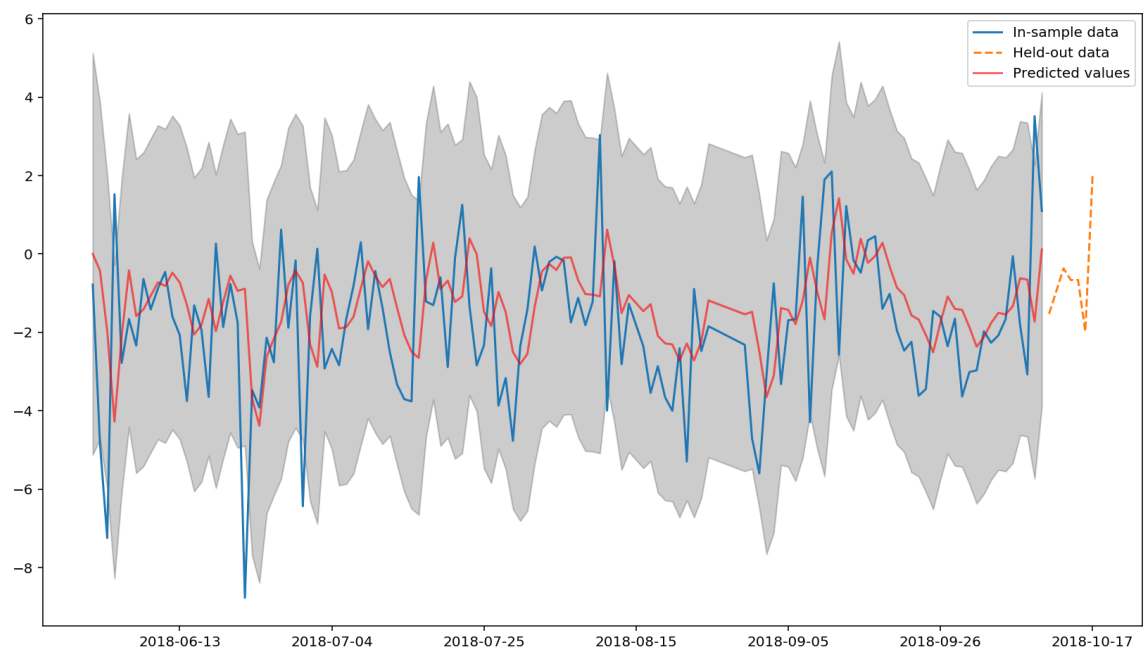
Test serial correlation of residuals (ljungbox): stat=47.984, p=0.181

Durbin-Watson test on residuals: d=2.13

(NB: 2 means no serial correlation, 0=pos, 4=neg)

Test for all AR roots outside unit circle (>1): True

Test for all MA roots outside unit circle (>1): True



Gambar 16. Evaluasi performansi model

Dilakukan prediksi untuk data test yang merupakan data Average Tone dari tanggal 11 Oktober 2018 sampai 17 Oktober 2018, dan data yang belum ada yaitu data pada tanggal 18 Oktober 2018. Diketahui, data test aktualnya dari data yang sudah ada adalah sebagai berikut.

SQLDATE	
2018-10-11	-1.52889
2018-10-13	-0.36819
2018-10-14	-0.66522
2018-10-15	-0.66332
2018-10-16	-1.97722
2018-10-17	1.98598

Name: AvgTone, dtype: float64

Diperoleh, data yang diprediksi adalah sebagai berikut.

128	0.97286
129	0.93672

```

130    0.67844
131    0.57529
132    0.44594
133    0.36353
dtype: float64

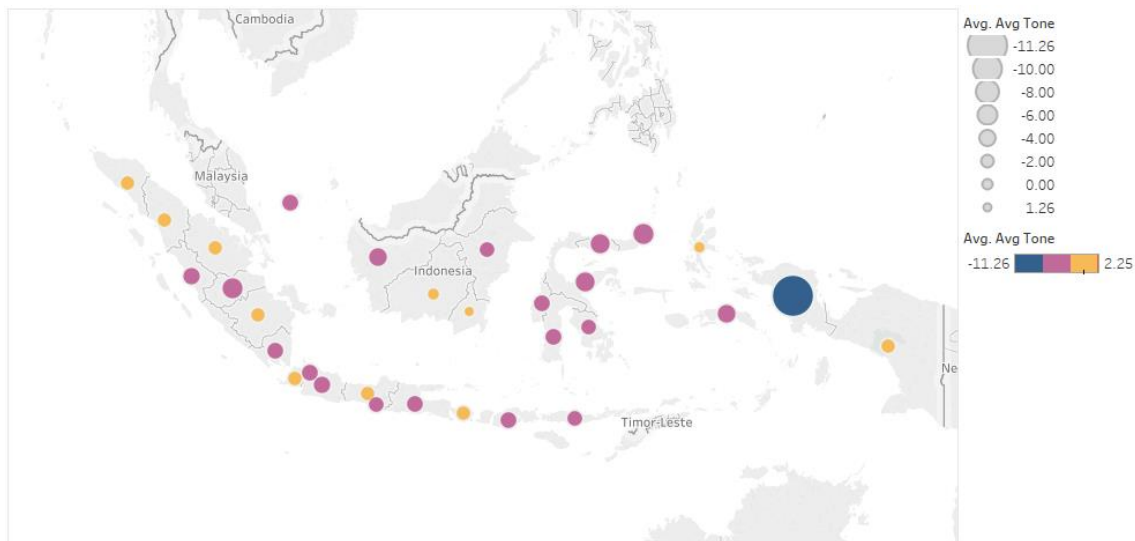
```

Selanjutnya, untuk data average tone yang belum ada pada tanggal 18 Oktober 2018 dihasilkan average tone Global Warmingnya adalah 0,28802, yang berarti ada kejadian positif dari global warming (yang dapat berbentuk penanganan global warming atau mitigasi). Dalam tahapan lebih lanjut, prediksi dengan runtun waktu ini dapat digunakan untuk memperkirakan apakah dalam rentang waktu ke depan pemberitaan mengenai global warming ini berangsur membaik atau memburuk.

4. Model Sebaran dan Visualisasi

Model sebaran yang ditunjukkan dalam subbab berikut memperhatikan tonasi kejadian yang diambil dari angka AVGTone dalam kolom yang termuat dalam dataset GDELT. Tonasi tersebut kemudian dipetakan ke peta Indonesia untuk melihat daerah-daerah terdampak dari label-label kejadian yang telah dikelompokkan. Tonasi dengan angka negatif menunjukkan kejadian negatif, 0 menunjukkan netral, dan positif menunjukkan kejadian positif.

a. Penyakit

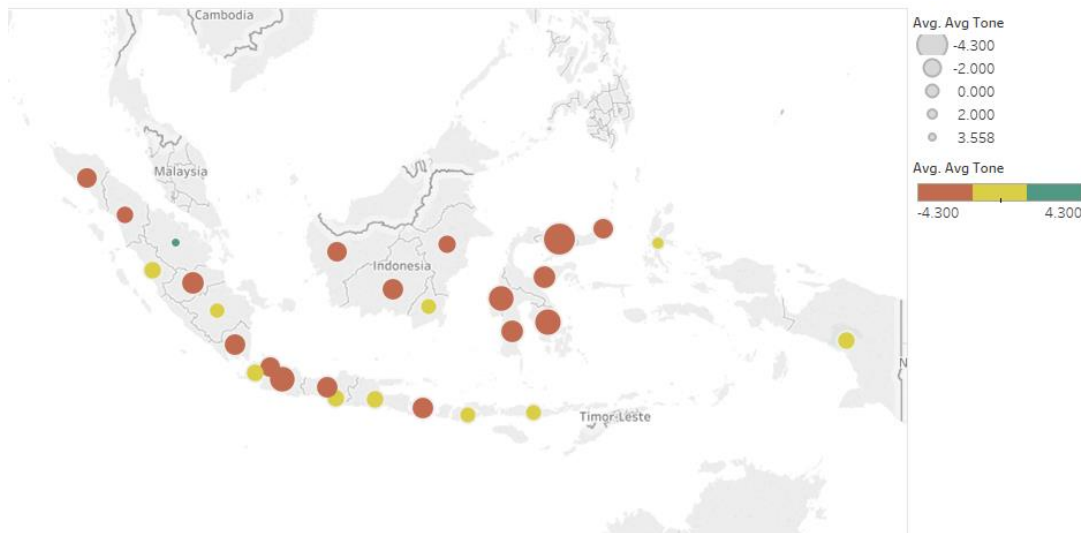


Gambar 17. Persebaran nilai Avg.Tone terhadap Pemberitaan tentang “Penyakit”

Persebaran rata-rata tonasi terhadap berita/kejadian yang berhubungan dengan “Penyakit” di Indonesia memiliki nilai antara -11,26 sampai 2,25. Nilai tonasi paling negative (-) berada di Papua Barat dengan nilai tonasi -11,26.

Daerah di Sorong, Papua Barat memiliki dampak/kejadian yang serius terhadap pemberitaan yang berhubungan dengan “Penyakit”.

b. Global Warming (Pemanasan Global)

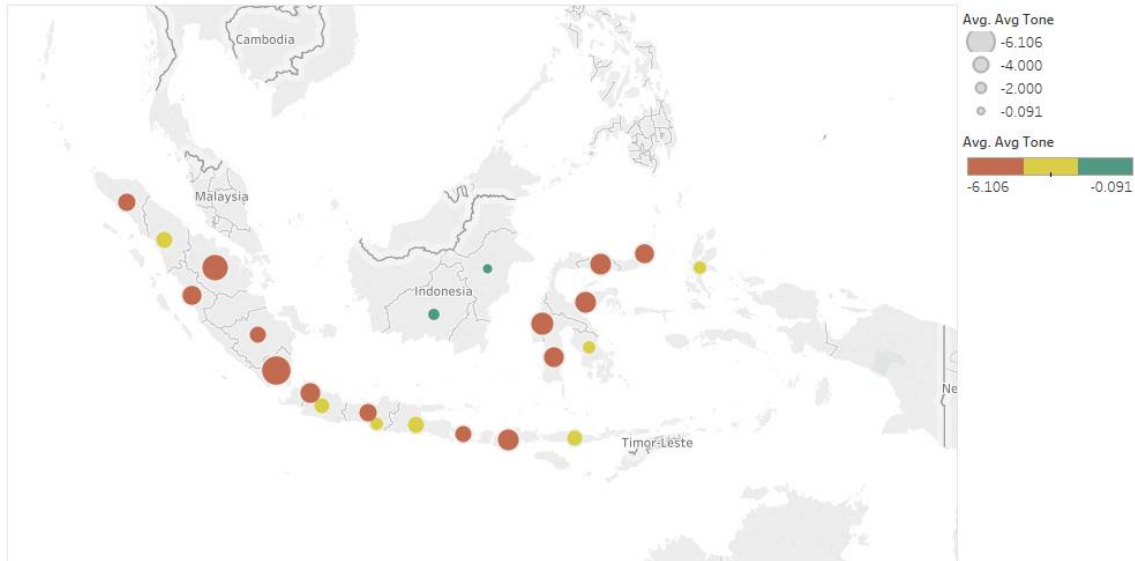


Gambar 18. Persebaran nilai Avg.Tone terhadap Pemberitaan tentang “Global Warming”

Daerah di Indonesia yang memiliki nilai tonasi paling negative (-) hampir berada di pulau-pulau besar di Indonesia, kecuali Pulau Irian. Sedangkan kota dengan tonasi paling negative berada di Kota Gorontalo dengan nilai notasi -4,30.

Sebagian besar kota-kota besar di Indonesia memiliki dampak/event yang berdampak serius terhadap issue “Global Warming”.

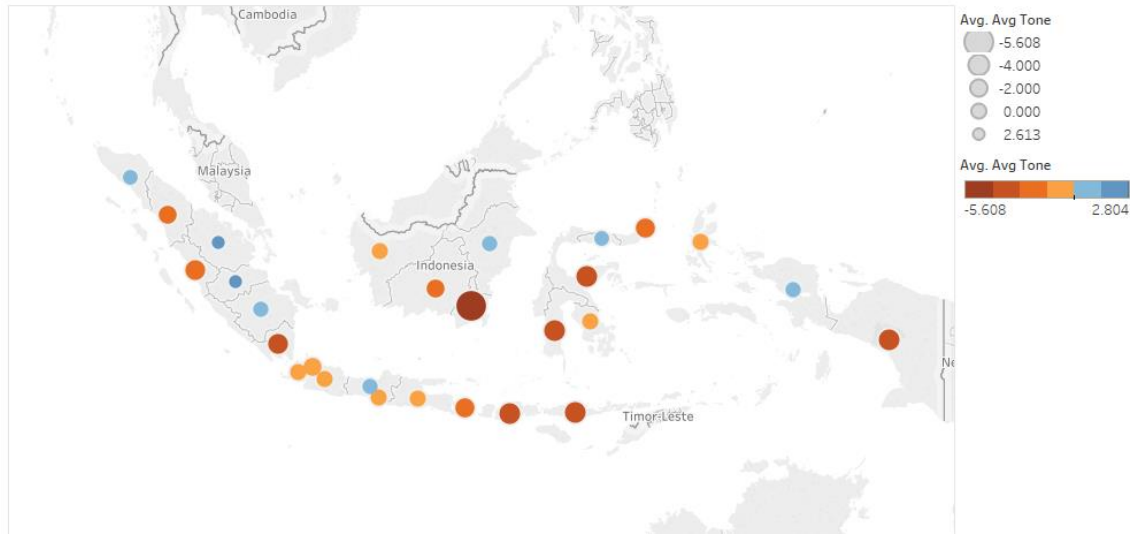
c. Penyebab Alami Perubahan Iklim (Bencana Alam Gunung Meletus)



Gambar 19. Persebaran Nilai Avg.Tone terhadap Pemberitaan tentang Penyebab Perubahan Iklim yang Berasal dari Alam.

Wilayah di Indonesia yang memiliki dampak serius (tonasi negative) terhadap penyebab perubahan iklim karena factor alam adalah sebagian besar wilayah di Sulawesi, Jawa dan Sumatera. Sedangkan daerah dengan tonasi positif berada di wilayah Pulau Kalimantan. Secara logis, hal ini juga dipengaruhi bahwa tidak ada gunung berapi di Pulau Kalimantan. Hal ini juga terjadi di Pulau Papua yang tidak memiliki indikasi terdapat kejadian di sana.

d. Perubahan Iklim



Gambar 20 . Pesebaran Nilai Tonasi terhadap Pemberitaan tentang “Perubahan Iklim”

Daerah dengan tonasi paling negative terhadap “Perubahan Iklim” adalah Kalimantan Selatan. Sedangkan daerah di Pulau Jawa juga memiliki dampak yang cukup serius terhadap “Perubahan Iklim”. Hal ini berbanding terbalik dengan daerah-daerah di Pulau Sumatera yang sebagian besar wilayah memiliki tonasi yang positif.

E. PEMBAHASAN

Dalam penelitian ini didapatkan data pemberitaan dari periode Juni-Oktober 2018 mengenai hal-hal yang sering berhubungan dengan perubahan iklim yaitu variabel dependen yang berupa “Perubahan iklim” dan variabel independen: “Global Warming (penyebab dari manusia)”, dan “Penyebab alami”. Serta hubungannya dengan tingkat “Penyakit” akibat kedua variabel tersebut.

Data pemberitaan yang didapatkan dari GDELT untuk analisa adanya keterkaitan hubungan antara perubahan iklim dengan variabel penelitian dapat dilihat dari diagram tonasi pemberitaan. Berdasarkan diagram tonasi pemberitaan variabel yang hampir berbanding lurus dengan pemberitaan perubahan iklim adalah variabel independen yaitu *global warming*, dimana kata kunci yang berhubungan dengan perubahan iklim dari variabel *global warming* adalah : “*pollution*”, “*deforestation*”, “*greenhouse*”, “*plastic waste*”, “*illegal logging*”, “*carbon dioxide*”, “*methane*”, “*emission*”, “*emissions*”, “*global warming*”. Faktor penyebab perubahan iklim yang terjadi saat ini lebih diakibatkan oleh aktivitas manusia yang menimbulkan “efek rumah kaca” dibandingkan faktor dari alam, seperti letusan gunung api.

Hubungan antara perubahan iklim dengan timbulnya penyakit yang didapatkan dalam penelitian ini menunjukkan terdapat hubungan antara kedua variabel yaitu berbanding lurus (trend positif), walaupun nilainya cukup kecil.

Faktor yang paling berpengaruh dalam perubahan iklim yang memengaruhi timbulnya suatu penyakit adalah *global warming* atau pemanasan global yang disebabkan oleh emisi gas di udara yang membentuk efek rumah kaca sehingga menjadikan suhu di permukaan bumi meningkat. Peningkatan suhu permukaan bumi ini dapat menyebabkan perubahan ekosistem di bumi (makhluk hidup).

Salah satu akibat yang paling signifikan dalam perubahan ekosistem di bumi akibat peningkatan suhu adalah pertumbuhan dan perkembangan nyamuk. Dimana nyamuk akan cepat berkembang di lingkungan yang hangat. Perkembangan dan pertumbuhan nyamuk ini akan mengakibatkan timbulnya penyakit seperti demam berdarah dan malaria.

Berdasarkan analisa data Avg.Tone (tonasi), ada beberapa daerah di Indonesia rawan dengan timbulnya penyakit akibat perubahan iklim adalah Kabupaten Sorong (Irian Jaya Barat) yang memiliki keterkaitan dengan penyakit malaria, dan Minahasa (Sulawesi Utara) dengan adanya penyakit rabies akibat virus dari hewan.

Penanganan yang serius perlu dilakukan di daerah Sorong dan sekitarnya, karena daerah tersebut sudah terdeteksi daerah endemik malaria. Hal yang paling penting dilakukan adanya penyadaran masyarakat tentang pola hidup sehat, baik kebersihan lingkungan maupun fisik, karena faktor penyebab utama malaria adalah nyamuk, maka seminimal mungkin membatasi pertumbuhan dan perkembangan nyamuk pembawa malaria tersebut. Begitu pula untuk kasus penyebaran rabies di wilayah Sulawesi Utara, hal ini juga diperlukan penanganan yang serius, karena penyakit ini disebabkan oleh virus, dan penyebaran penyakit ini tergolong cukup cepat serta akibat dari penyakit ini yang berpengaruh pada saraf manusia.

Daerah – daerah di Indonesia yang rawan akan penyakit akibat perubahan iklim diperlukan sosialisasi menumbuhkan kesadaran akan hidup bersih, menjaga kebersihan lingkungan tempat tinggal, antara lain pengolahan limbah/sampah (organik/plastik).

Berdasarkan model rutun waktu menggunakan *Model Diagnostic Checking* dapat digunakan untuk tahapan lebih lanjut dalam memperkirakan apakah dalam rentang waktu ke depan pemberitaan mengenai global warming ini berangsur membaik atau memburuk. Dari grafik evaluasi performansi model nilai prediksi menunjukkan trend yang positif pada variabel penyebab perubahan iklim (*global warming*), yang artinya banyak aktor (pemerintah, lembaga, swasta, dll) yang berkontribusi dalam mengurangi penyebab *global warming* (pemanasan global) yang merupakan faktor utama dalam fenomena perubahan iklim.

F. KESIMPULAN

1. Faktor yang paling berpengaruh terhadap fenomena perubahan iklim di Indonesia adalah factor-faktor yang berhubungan dengan *Global Warming* ("*pollution*", "*deforestation*", "*greenhouse*", "*plastic waste*", "*ilegal logging*", "*carbon dioxide*", "*methane*", "*emission*", "*emissions*", "*global warming*")

2. Fenomena perubahan iklim berpengaruh terhadap penyakit, namun korelasi antar kedua variable tersebut kecil.
3. Faktor yang paling berpengaruh dalam perubahan iklim yang memengaruhi timbulnya suatu penyakit adalah *global warming* atau pemanasan global yang disebabkan oleh emisi gas di udara yang membentuk efek rumah kaca.
4. Penyakit yang sering muncul akibat perubahan iklim adalah demam berdarah dan malaria.
5. Daerah di Indonesia rawan dengan timbulnya penyakit akibat perubahan iklim adalah Kabupaten Sorong (Irian Jaya Barat).
6. Adanya trend positif menunjukkan banyak aktor (pemerintah, lembaga, swasta, dll) yang berkontribusi dalam mengurangi penyebab *global warming* (pemanasan global).
7. Hal penting yang dilakukan adalah kegiatan penyadaran masyarakat tentang pola hidup sehat, baik kebersihan lingkungan maupun fisik.