



Data Science and Moral Gap: Hoax and Sentiment

Dhomas Hatta Fudholi, Ph.D.

Data Science Talk & Workshop Series #2 2020

Sabtu, 14 Maret 2020



Dhomas Hatta Fudholi, Ph.D.

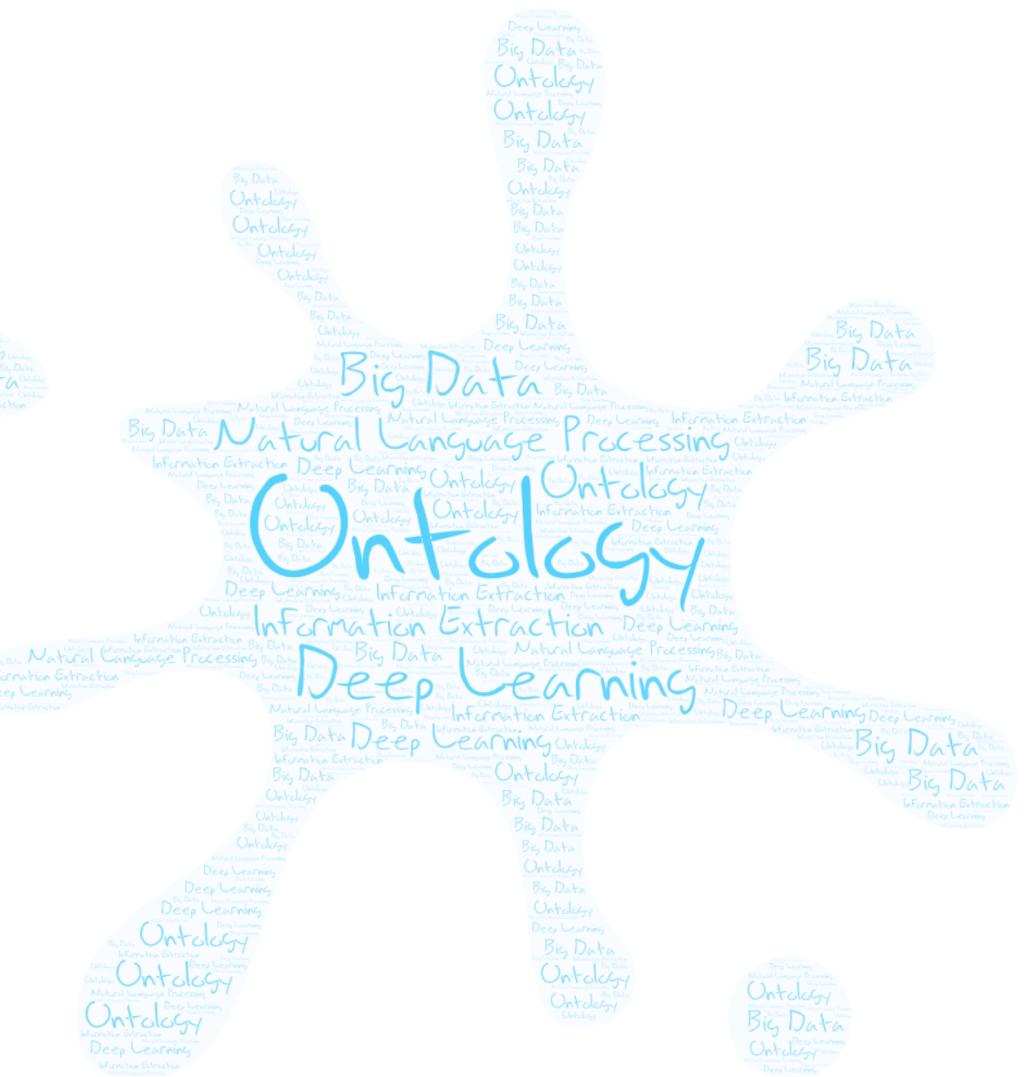
Assistant Professor

CDS Expert

Secretary of Undergraduate Program in Informatics

Department of Informatics, Universitas Islam Indonesia

hatta.fudholi@uii.ac.id





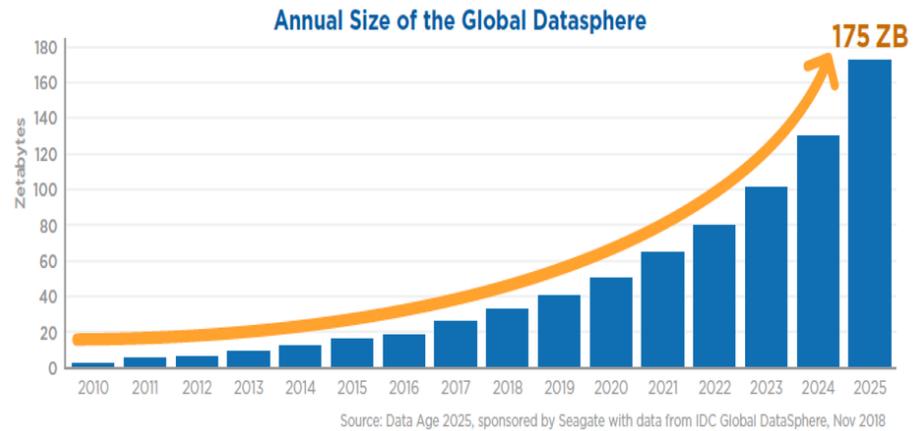
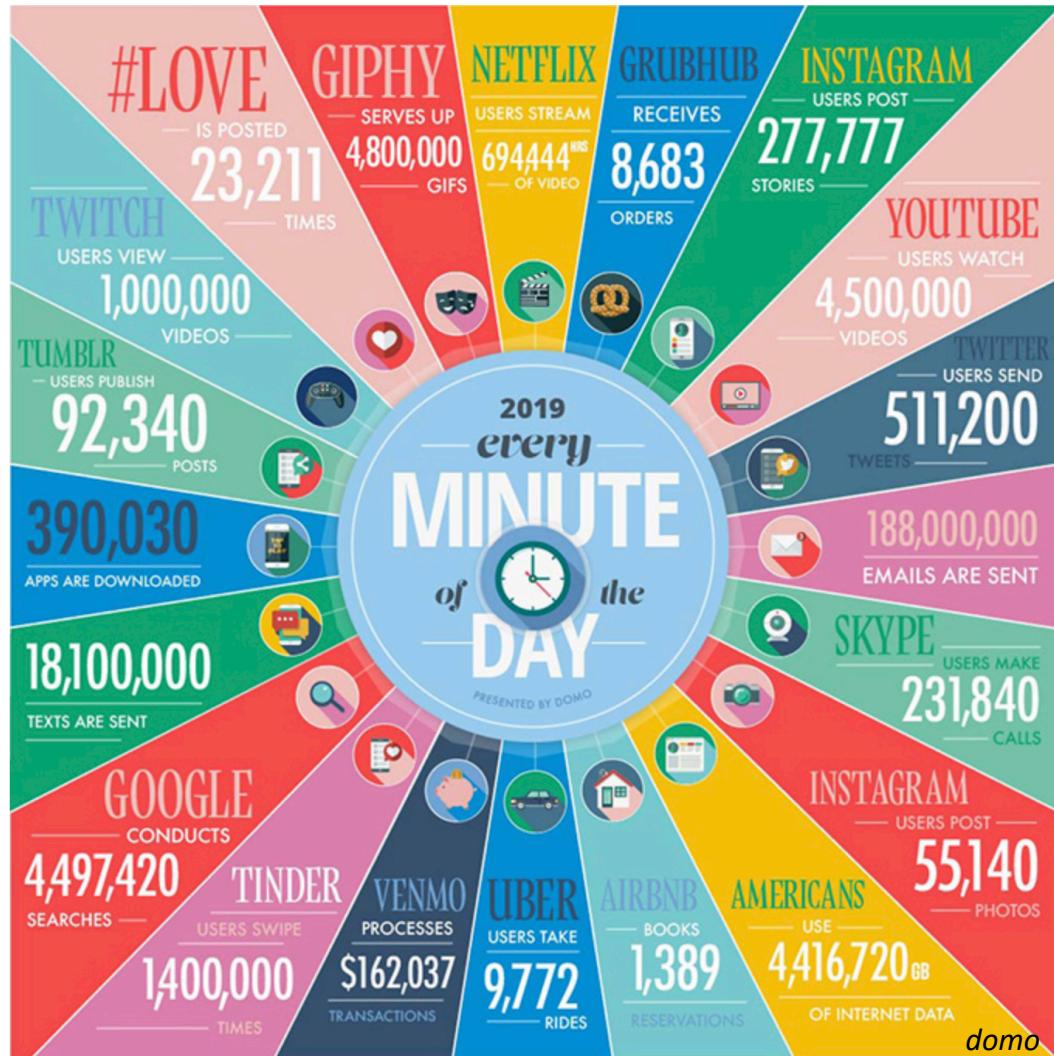
UNIVERSITAS
ISLAM
INDONESIA



e-zigurat.com

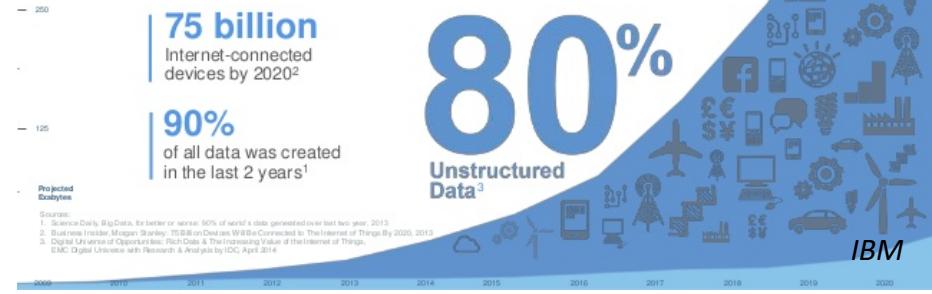


UNIVERSITAS
ISLAM
INDONESIA



With the rise of cloud, mobility, IoT, social and analytics, the data explosion is accelerating.

This confluence of technologies has amplified the data explosion, creating incredible growth-on-growth for unstructured data. New data sources are added daily, resulting in a valuable data ecosystem for every business.





UNIVERSITAS
ISLAM
INDONESIA

"DATA IS THE NEW OIL."

From the beginning of recorded time until 2003, we created 5 exabytes of data.

In 2011 the same amount was created every two days.

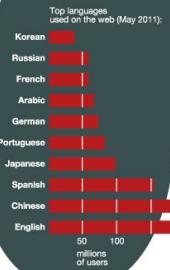
By 2013, it's expected that the time will shrink to 10 minutes.

Every hour, we create enough Internet traffic to fill 7 billion DVDs.

Side by side, that's seven times the height of Everest.

Coined in 2006 by Clive Humby, a British data commercialization entrepreneur, this now famous phrase was embraced by the World Economic Forum in a 2011 report, which considered data to be an economic asset, like oil.

English is the dominant language of the web. But by 2014 it will be Chinese, if its current rate of increase continues.



247 billion EMAILS are sent every day. (Up to 80% are spam.)

There are 133 million BLOGS on the web. As of August 2012, there were just over 4 million articles in the English Wikipedia.



Just as a study of activity on Twitter gave residents, family members, and journalists advance warning of details about the devastating earthquake and tsunami in Japan, high-frequency traders, with the help of computer algorithms, use Big Data to follow trends and to act quickly on their findings.

These specialized algorithms make split-second decisions to buy or sell a commodity.

New cable being laid under the Atlantic will shave

5 milliseconds

from the current 50 milliseconds it takes for trading instructions to travel between New York City and London.

With new fiber-optic cable, the round-trip time between New York and London will be 59.6 milliseconds.

This 5-millisecond saving is worth many millions of dollars to the trading firms who use the cable (and who will pay millions to do so).

How they save 5 milliseconds

The depth of the Atlantic Ocean varies. The new cable will lie on areas of the ocean floor that are up to 1,000 feet shallower than the current fastest cable. By taking a different route, the new cable is shorter, meaning that the time it takes for messages to travel along it is shortened.

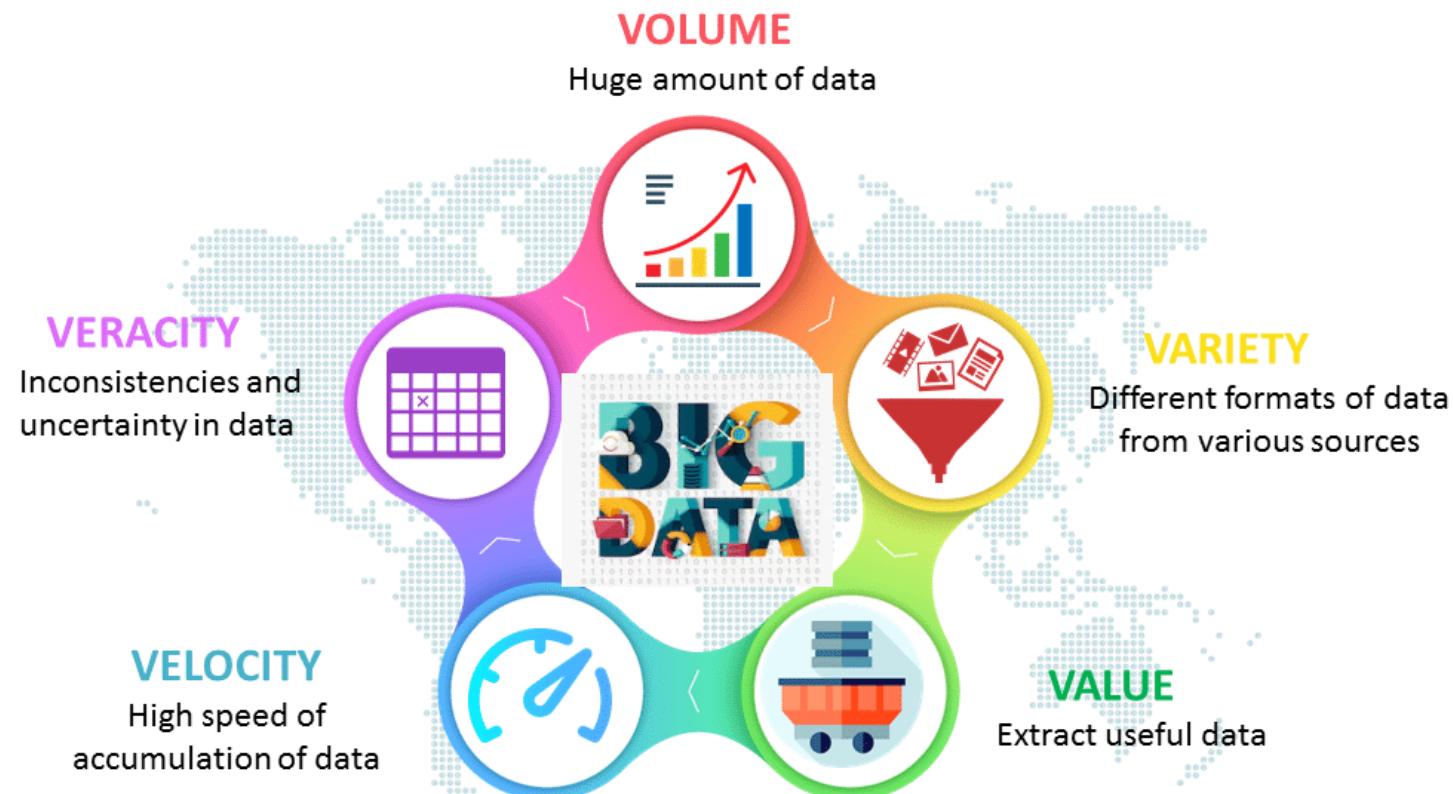


Information graphics by Neal Holden

pinterest.com



UNIVERSITAS
ISLAM
INDONESIA



if ($V == HF^2$) then E++

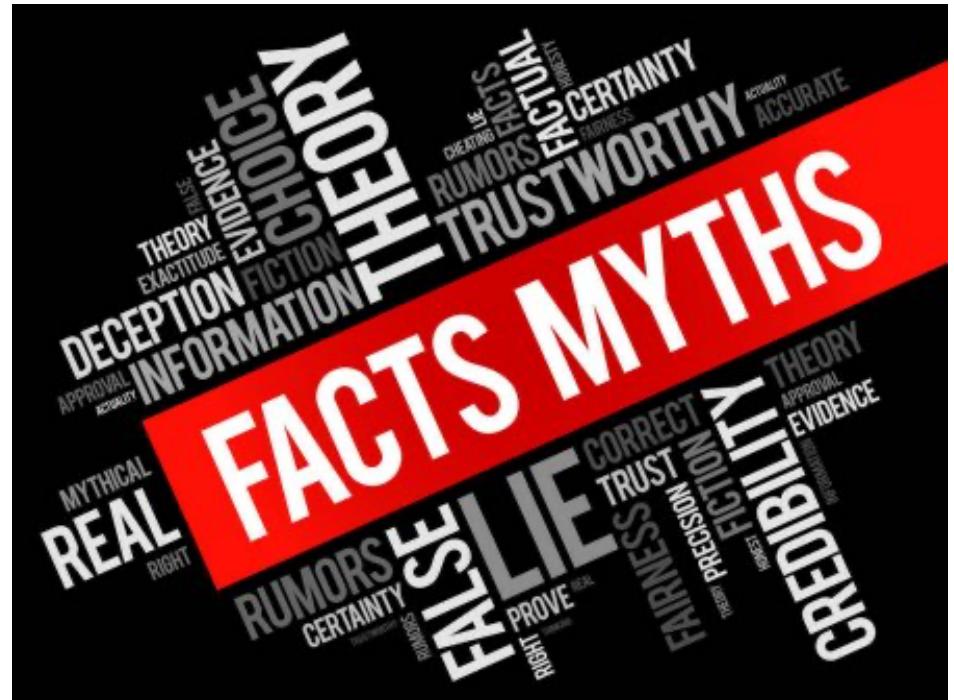
Data Science and Moral Gap: Hoax and Sentiment

The "V"

Data Science and Moral Gap: Hoax and Sentiment

Veracity

- lexico.com (Oxford)
 - conformity to facts; accuracy.
- dictionary.com
 - conformity to truth or fact; accuracy.
- dictionary.cambridge.org
 - the quality of being true, honest, or accurate.



depositphotos.com

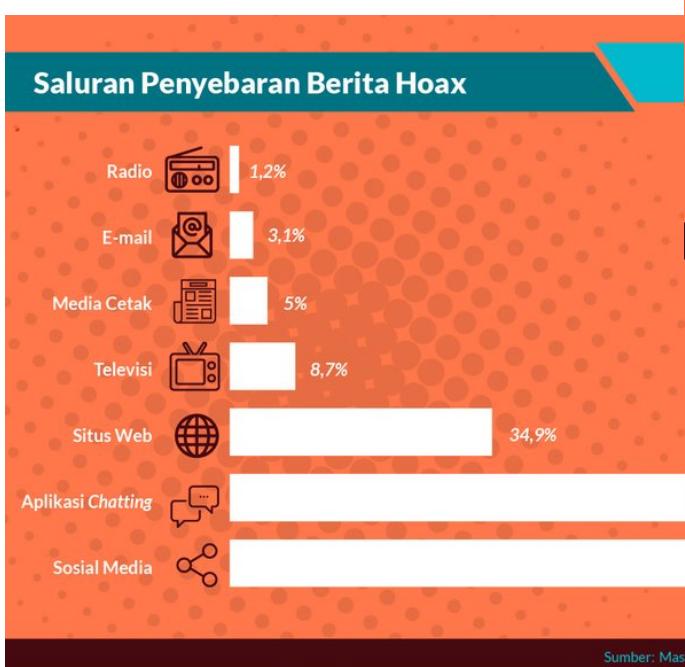
The “HF”

Data Science and Moral Gap: Hoax and Sentiment

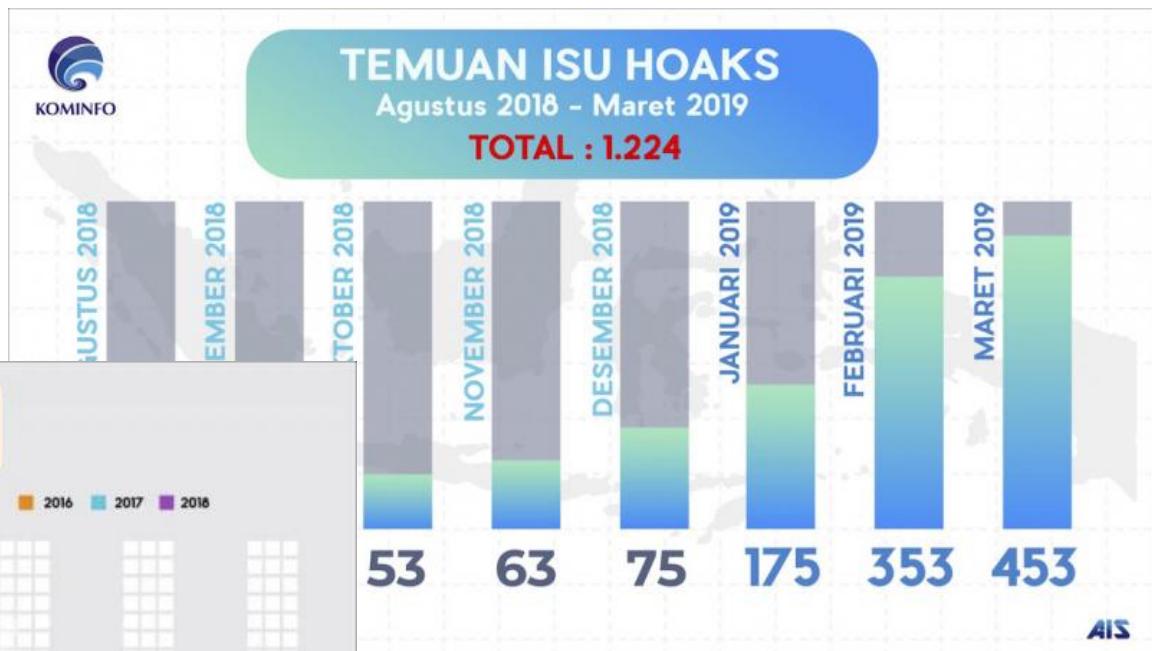
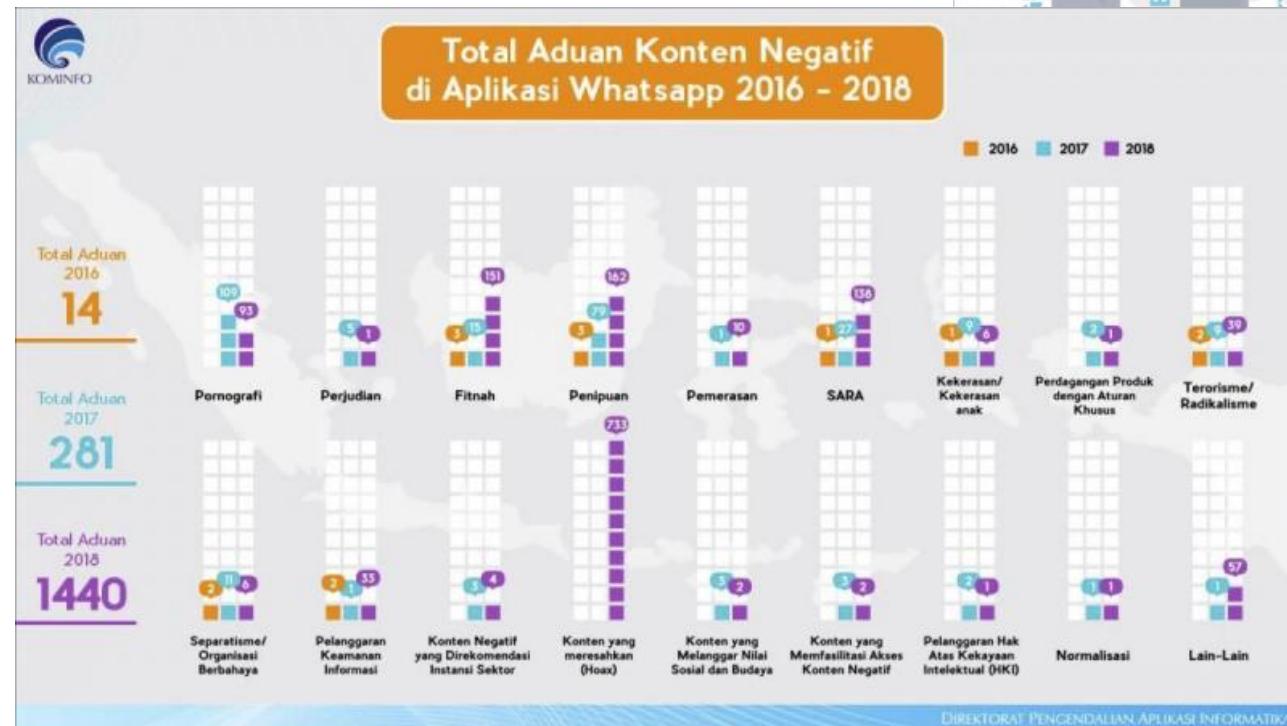


UNIVERSITAS
ISLAM
INDONESIA

Hate and Fake



Hate and Fake





Hate and Fake

13

03-2020

[HOAKS] Urin dan Kotoran Sapi Jadi Obat Virus Corona

Kategori Hoaks

13

03-2020

[HOAKS] Minum Alkohol Bisa Kurangi Risiko Terkena Corona

Kategori Hoaks

11

03-2020

[HOAKS] Ramalan Tahun 2016 Terkait Virus Corona

Kategori Hoaks

10

03-2020

[HOAKS] Pesan Berantai Berisi Pencegahan Virus Corona yang Mengatasnamakan UNICEF

Kategori Hoaks

<https://kominfo.go.id/>



Hate and Fake

Verify Your Account



This message was identified as junk. [Not junk](#) [Show images](#)



Jpmorgan Chase <H.J.Bongartz@t-online.de>

3/30/2019 12:13 PM



To: smrfs@emailonline.chase.com



The picture can't be displayed.

Chase : Important notification regarding your chase bank debit/credit/ATM/Prepaid Card.

Dear Customer

As part of our security commitment to safeguard your online banking transactions and activities, this notification is to confirm that you or an authorized party used your account information's for online transactions using your chase debit/credit card details.

To view and confirm this transaction, kindly click the button below

[View all Transactions](#)

Chase Card Services
Email Operations Team

Your privacy is important to us. See our online Security Center to learn how to protect your information. Chase Privacy Operations, P.O. Box 659752, San Antonio, TX 78265-9752.

© 2019 JP Morgan Chase Bank, N.A. Member FDIC

<https://kominfo.go.id/>



Hate and Fake

[PROPAGANDA MACHINE](#)

This Site Uses AI to Generate Fake News Articles

It's hard to mince words about how much this sucks.

DAN ROBITZSKI | FEBRUARY 13TH 2020

A new website, “[NotRealNews.net](#),” uses artificial intelligence to populate what resembles a news site’s home page, complete with AI-written [fake news stories](#).

The website, a project by the AI [development company Big Bird](#), is supposed to be a showcase of how the company’s algorithms can help journalists quickly write compelling news, according to the website’s “[about](#)” page.



UNIVERSITAS
ISLAM
INDONESIA

Pix2Story Tell stories like never before.



Browse gallery

Learn more on [Al.lab](#)



Adventure



Thriller



Sci-fi



Share

An adventure story

By AI Pix2Story bot

Once the motorcycle was taken by its owner , it was a difficult trek . Another night , so many of them were more cautious about what happened to be the perfect motorcycle . Her mind raced back to the road , but the path that smoothed down the road became a dirt road . His bike must be bound . He ' d never seen someone like that , unless he ' s caught up with his thoughts , he said to lessen the dread of human activity .

<https://pix2story.azurewebsites.net>



Adventure



Thriller



Sci-fi



Share

An adventure story

By AI Pix2Story bot

Once the motorcycle was taken by its owner , it was a difficult trek . Another night , so many of them were more cautious about what happened to be the perfect motorcycle . Her mind raced back to the road , but the path that smoothed down the road became a dirt road . His bike must be bound . He ' d never seen someone like that , unless he ' s caught up with his thoughts , he said to lessen the dread of human activity .



Thriller



Sci-fi



Share

A thriller story

By AI Pix2Story bot

The GPS disappeared again, no time to get into the dirt lane. Elspeth rushed toward her side. "Mr. Leonard dashed to his aid. He was still standing in the doorway of the truck, his eyes wide open, and the smell. He wondered if he was going to die for a pulse. Wasn't cold. "You must be dead. You're going to die here, Mr. Johnson and the camera would show us the depths of the High Command after the outer missile was shot.

The other “HF”

Data Science and Moral Gap: Hoax and Sentiment



Human's Faith

- **Don't rush to spread the word**

"It is enough for a man to prove himself a liar when he goes on narrating whatever he hears." [HR. Muslim no.7]

- **Check the truth of a news carefully**

"O you who believe, if you come to the wicked who bring a message, then examine it carefully, so that you do not inflict a disaster upon a people without knowing the circumstances that cause you to regret your actions." [QS. Al-Hujuraat [49]: 6]

- **Is there any value in the news?**

"Whoever believes in Allah and the Last Day, should talk what is good or keep quiet." (HR. Bukhari no. 6018 dan Muslim no. 74)

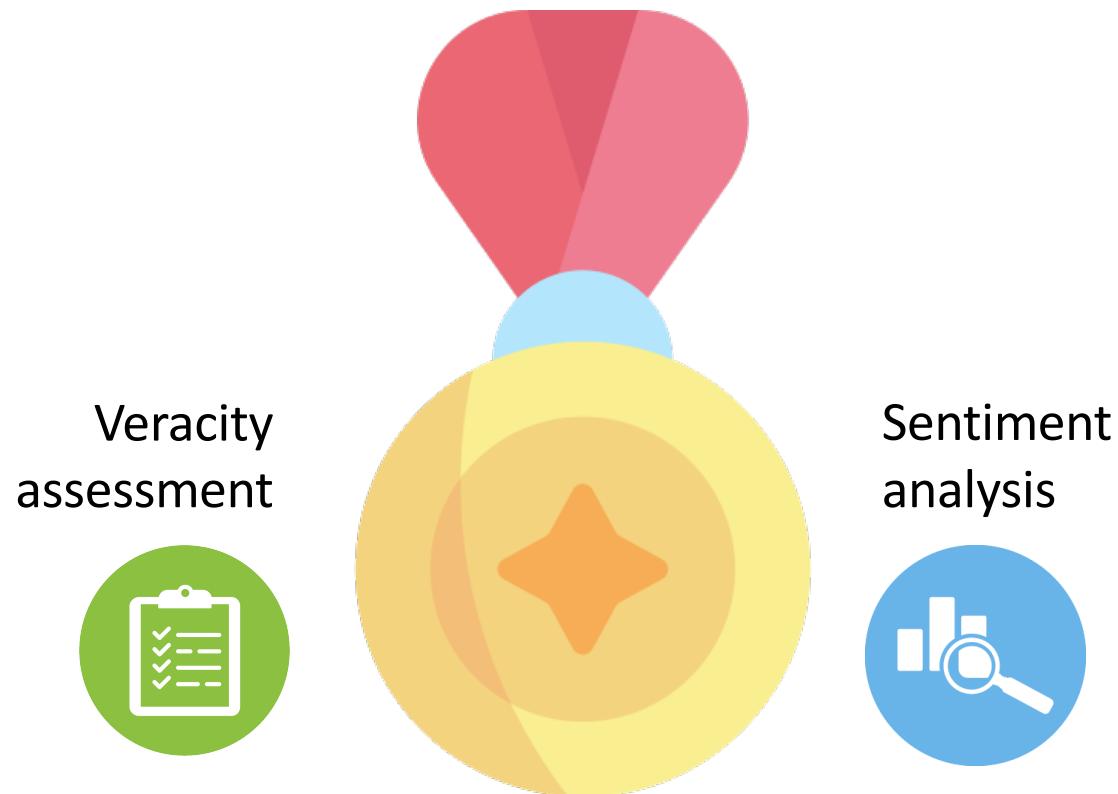
The “E”

Data Science and Moral Gap: Hoax and Sentiment



UNIVERSITAS
ISLAM
INDONESIA

Excellence in Data Science





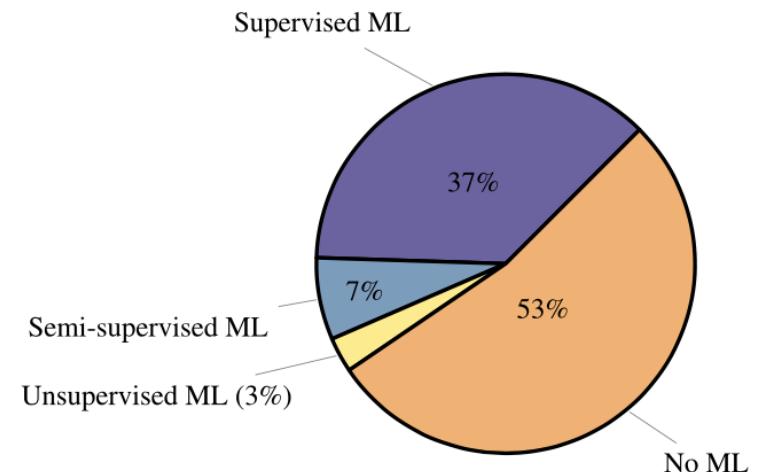
Veracity Assessment

- Why?
 - Fake news, malicious rumors, fabricated reviews.
 - Incomplete, ambiguous, contradicting, biased, or wrong information.
 - The need to identify the spreading of false information.

Veracity Assessment

Three main approaches

- The implicit features approach hypothesizes that non-veridical statements differ from veridical statements not only concerning the actual claim but also in other aspects that can be used for assessment.
- The explicit fact checking approach makes use of external data to evaluate a claim in relation to existing knowledge.
- The appeal to authority approach stipulates that a claim can be trusted if it is also claimed or can be verified by an authoritative source.



reference: <https://doi.org/10.1016/j.dss.2019.113132>



Veracity Assessment [example]

- **Input**

- A headline and a body text - either from the same news article or from two different articles.

- **Output**

- Classify the stance of the body text relative to the claim made in the headline into one of four categories:
 - **Agrees:** The body text agrees with the headline.
 - **Disagrees:** The body text disagrees with the headline.
 - **Discusses:** The body text discuss the same topic as the headline, but does not take a position
 - **Unrelated:** The body text discusses a different topic than the headline

<http://www.fakenewschallenge.org>

Pix2Story Tell stories like never before.

REVISITED



Browse gallery

Learn more on [Al.lab](#)



Adventure



Thriller



Sci-fi



Share

An adventure story

By AI Pix2Story bot

Once the motorcycle was taken by its owner , it was a difficult trek . Another night , so many of them were more cautious about what happened to be the perfect motorcycle . Her mind raced back to the road , but the path that smoothed down the road became a dirt road . His bike must be bound . He ' d never seen someone like that , unless he ' s caught up with his thoughts , he said to lessen the dread of human activity .

<https://pix2story.azurewebsites.net>



Adventure



Thriller



Sci-fi



Share

An adventure story

By AI Pix2Story bot

Once the motorcycle was taken by its owner , it was a difficult trek . Another night , so many of them were more cautious about what happened to be the perfect motorcycle . Her mind raced back to the road , but the path that smoothed down the road became a dirt road . His bike must be bound . He ' d never seen someone like that , unless he ' s caught up with his thoughts , he said to lessen the dread of human activity .



Thriller



Sci-fi



Share

A thriller story

By AI Pix2Story bot

The GPS disappeared again, no time to get into the dirt lane. Elspeth rushed toward her side. “Mr. Leonard dashed to his aid. He was still standing in the doorway of the truck, his eyes wide open, and the smell. He wondered if he was going to die for a pulse. Wasn’t cold. “You must be dead. You’re going to die here, Mr. Johnson and the camera would show us the depths of the High Command after the outer missile was shot.

Sentiment Analysis

- Sentiment analysis (SA) is a process of studying public opinion about an entity.
- Opinion is a judgment of a person regarding an entity that varies from one another and tells about the choice of opinion holder.
- Sentiment analysis plays an important role in decision-making and the recommender system



reference: https://doi.org/10.1007/978-981-15-0029-9_7



Sentiment Analysis

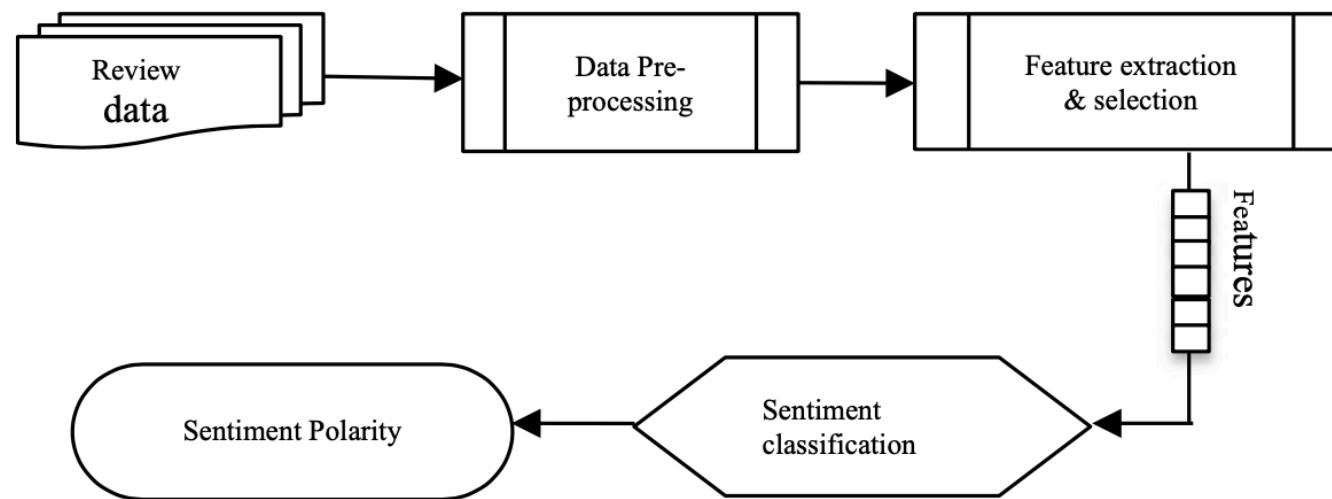
- Document level
- Sentence level
- Feature level or Aspect level





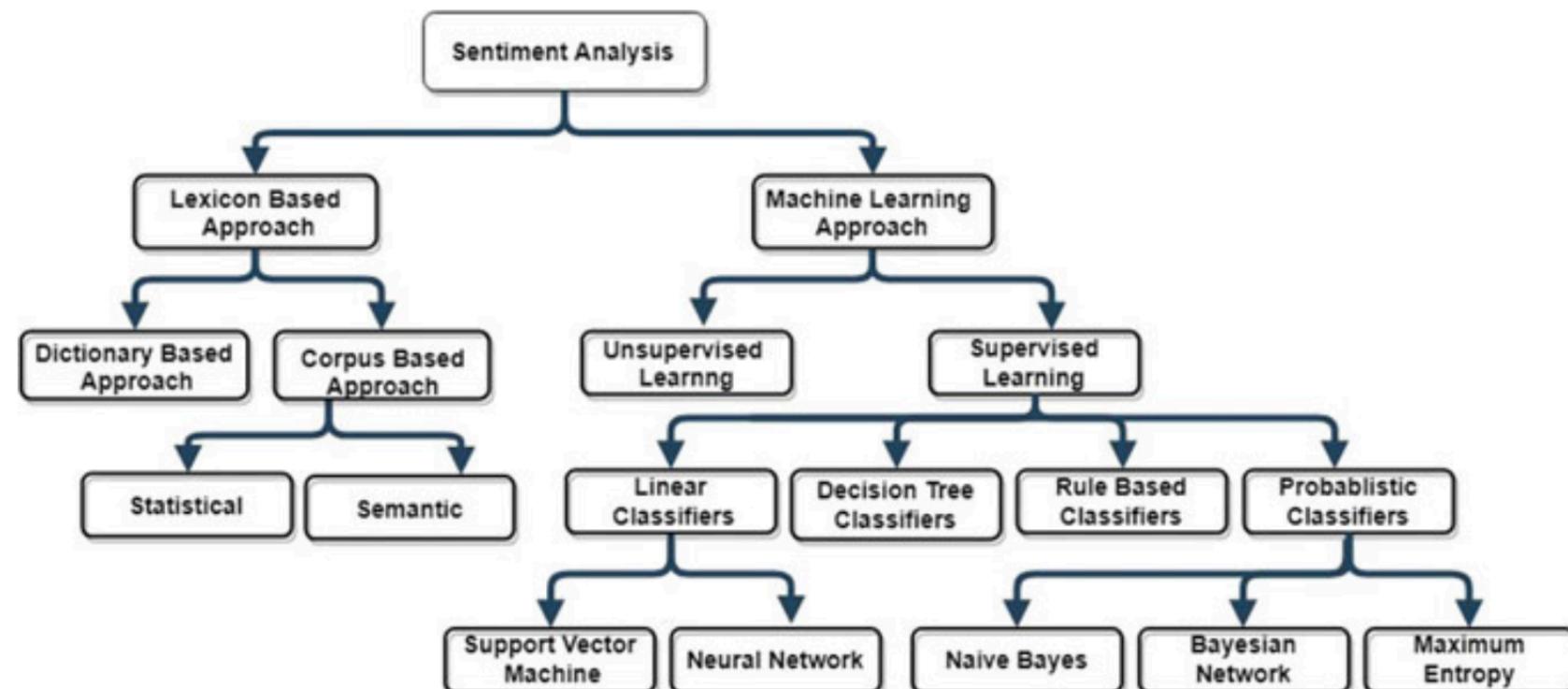
UNIVERSITAS
ISLAM
INDONESIA

Sentiment Analysis





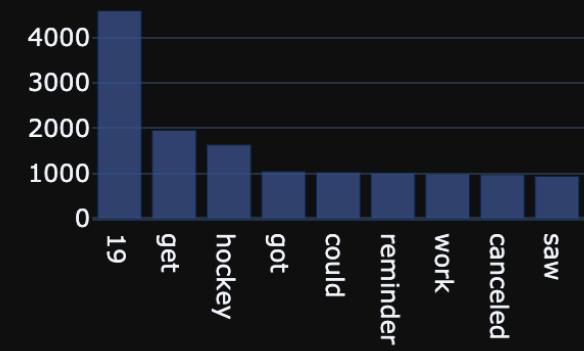
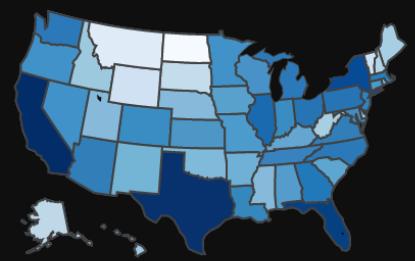
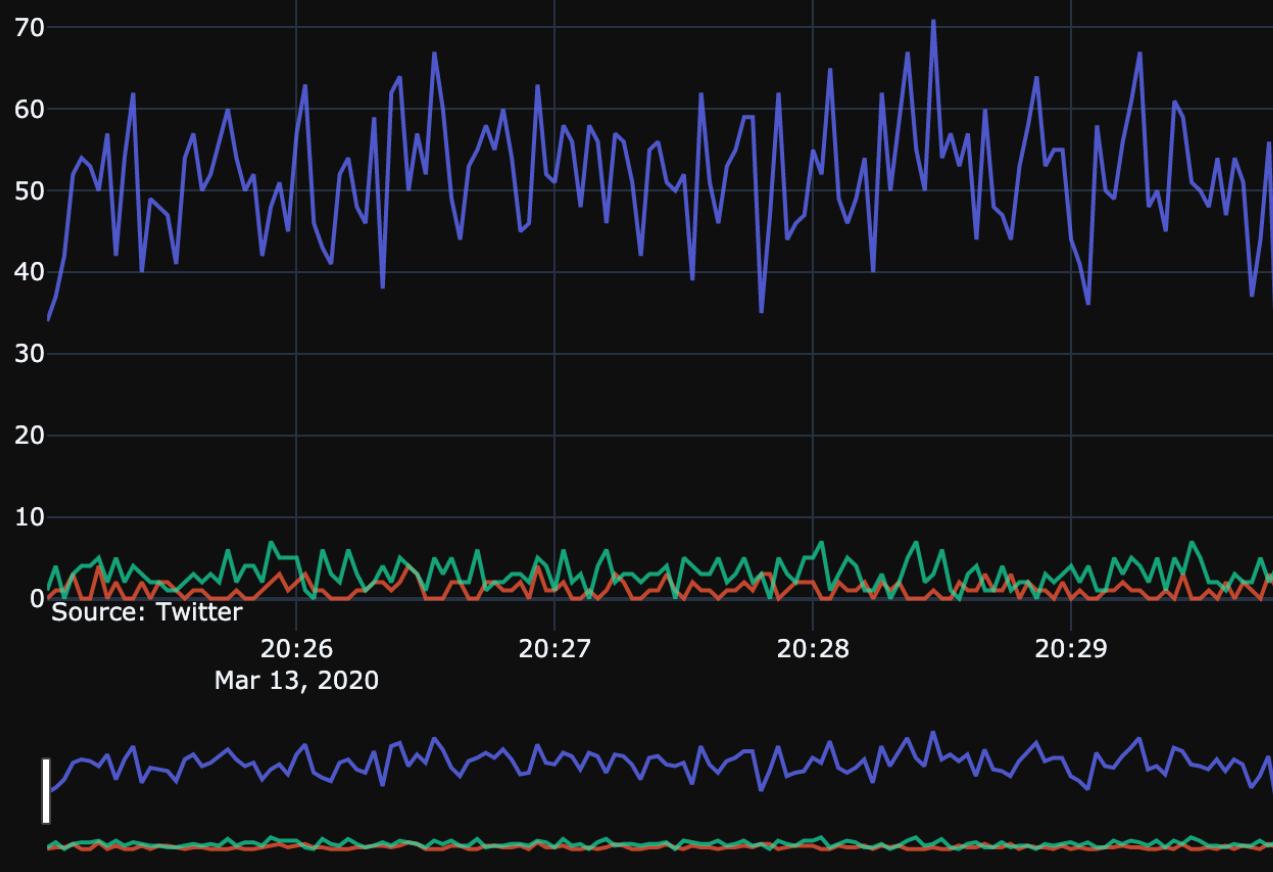
Sentiment Analysis





UNIVERSITAS
ISLAM
INDONESIA

Real-time tracking 'COVID-19' mentions on Twitter 03-13 20:30 UTC



E++

The positive excellences

Data Science and Moral Gap: Hoax and Sentiment

Dhomas Hatta Fudholi, Ph.D. <hatta.fudholi@uii.ac.id>

Data Science Talk & Workshop Series #2 2020

Sabtu, 14 Maret 2020