

ANALISIS SENTIMEN PENGGUNA TWITTER PADA FINAL PIALA DUNIA QATAR 2022

Kelompok G Data Mining B

Aisyah Gefira	2006572945
Aldinda Albanna	2006533295
Faza Muthmainnah	2006523672
Inez Aurellia Rosyana	2006571204
Nurrochmi Andina	2006571103
Valery Ongso Putri	2006570965

TABLE OF CONTENTS

01

**Tujuan dan
Dataset**

04

**Sentiment
Analysis**

02

**Data
Preprocessing**

05

**Social Network
Analysis**

03

**Exploratory Data
Analysis**

06

**Kesimpulan dan
Rekomendasi**

01

Tujuan dan Dataset

Tujuan

- Melakukan klasifikasi apakah Final Piala Dunia Qatar 2022 memiliki sentimen positif, negatif, atau netral dari pengguna Twitter.
- Menemukan informasi atau *insight* dari data hasil *scraping* Twitter.
- Mengetahui akun Twitter yang berpengaruh.

Dataset

- *Scraping* data tweet menggunakan module twint.
- Rentang waktu: 18 Desember 2022 - 20 Desember 2022.
- Kata kunci : “pildun”

Cuplikan data *scraping*:

Unnamed: 0		id	conversation_id	created_at	date	timezone	place	tweet	language	hashtags	...	geo	source	user_rt_id	user_rt	retweet_id	reply_to	retweet_date	translate	trans_src	trans_dest
0	0	1605352334573256704	1605324825089945601	1.671581e+12	2022-12-20 23:59:57	0	NaN	@SamFerkyFers10 @dextratime hebat daimana ur...	in		...	NaN	NaN	NaN	NaN	NaN	[{"screen_name": "SamFerkyFers10", "name": "Sa...	NaN	NaN	NaN	NaN
1	1	1605351975775719425	1605282617234165761	1.671581e+12	2022-12-20 23:58:31	0	NaN	@Fasa77342812 @rakkaustarina1 @dextratime Lah...	in		...	NaN	NaN	NaN	NaN	NaN	[{"screen_name": "Fasa77342812", "name": "Fasa...	NaN	NaN	NaN	NaN
2	2	1605351517304721408	1605351517304721408	1.671581e+12	2022-12-20 23:56:42	0	NaN	messi sengaja menang pildun biar bisa keluar d...	in		...	NaN	NaN	NaN	NaN	NaN		NaN	NaN	NaN	NaN
3	3	1605350758920386008	1605228813805244416	1.671580e+12	2022-12-20 23:53:41	0	NaN	@dextratime Bayangin lah sebelumnya 2014 kala...	in		...	NaN	NaN	NaN	NaN	NaN	[{"screen_name": "dextratime", "name": "Extra...	NaN	NaN	NaN	NaN
4	4	1605350019648737280	1605324177959194625	1.671580e+12	2022-12-20 23:50:45	0	NaN	@UTDwb @whathevuck @dextratime nah itu, mung...	in		...	NaN	NaN	NaN	NaN	NaN	[{"screen_name": "UTDwb", "name": "Mindro", ...	NaN	NaN	NaN	NaN
...
10505	10505	1604540611100807168	1604540611100807168	1.671387e+12	2022-12-18 18:14:27	0	NaN	Juara pildun senam jantung dulu 🏆	in		...	NaN	NaN	NaN	NaN	NaN		NaN	NaN	NaN	NaN
10506	10506	1604540606042501121	1604517116996354049	1.671387e+12	2022-12-18 18:14:26	0	NaN	@Rianyoga _memeriahkan pildun niate	in		...	NaN	NaN	NaN	NaN	NaN	[{"screen_name": "Rianyoga_", "name": "Hooom...	NaN	NaN	NaN	NaN
10507	10507	1604540605480440001	1604539816032759809	1.671387e+12	2022-12-18 18:14:25	0	NaN	@babyjaerulo abis nnton pildun k	in		...	NaN	NaN	NaN	NaN	NaN	[{"screen_name": "babyjaerulo", "name": "sar", ...	NaN	NaN	NaN	NaN
10508	10508	160454060379852577	1604536870813446146	1.671387e+12	2022-12-18 18:14:25	0	NaN	@mahpudfour7 @The_Redsindo Ya mau lu ayahnya r...	in		...	NaN	NaN	NaN	NaN	NaN	[{"screen_name": "mahpudfour7", "name": "mahpu...	NaN	NaN	NaN	NaN
10509	10509	1604540597754548225	1604540597754548225	1.671387e+12	2022-12-18 18:14:24	0	NaN	Berkah sakit, dari jem 4-9 tidur. Jadinya bisa...	in		...	NaN	NaN	NaN	NaN	NaN		NaN	NaN	NaN	NaN

10510 rows × 39 columns

02

Data Preprocessing



Data Preprocessing

Cleaning Tweet

- Memeriksa missing value
- Menghapus baris yang duplikat berdasarkan teks tweet
- Menghapus tweet dari akun @jogjamnfs karena tidak memuat kata “pildun”
- Menghilangkan '@user' (mention) dari semua teks tweet
- Menghilangkan emoji
- Mengganti teks tweet dengan huruf kecil
- Menghilangkan link yang terdapat pada tweet
- Menghilangkan simbol
- Menghilangkan kata-kata yang tidak diperlukan (Custom StopWords)
- Mengganti kata slang dengan kata baku
- Melakukan *stemming*
- Melakukan lematisasi

Hasil Data Preprocessing

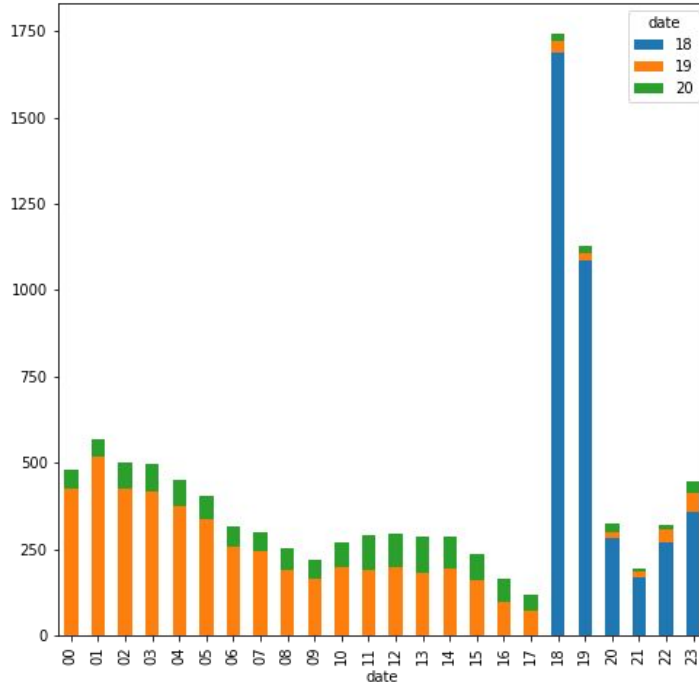
	date	username	link	clean_tweet	label	tweet
5629	2022-12-19 01:22:28	rantas_romy	https://twitter.com/rantas_romy/status/1604648...	ngerasa akhir tahun banyak banget peluang untu...	NaN	Gue ngerasa akhir tahun banyak bgt peluang keb...
5055	2022-12-19 02:35:44	txtdaricah	https://twitter.com/txtdaricah/status/16046667...	minsoc nonton pildun full ngantuk banget asli	NaN	minsoc + nonton pildun = full ngantuk bgt asli
5982	2022-12-19 00:37:34	pjmine99	https://twitter.com/pjmine99/status/1604637027...	pagi pagi buat tertawa sama meme pildun jancoqq	NaN	pagi pagi dibuat ngakak ama meme pildun jancoq...
6315	2022-12-18 23:49:53	bbytriceratops	https://twitter.com/bbytriceratops/status/1604...	pa semalem stop nonton keinget dulu argentina ...	NaN	Pas 2-2 semalem tuh gue stop nonton karena kei...
3715	2022-12-19 06:06:46	adwibasuki	https://twitter.com/adwibasuki/status/16047198...	dulu awal jelang selenggara biasa saja tidak s...	NaN	@bintangemon Dulu awal menjelang penyelenggara...
849	2022-12-20 08:42:47	iHugMez	https://twitter.com/iHugMez/status/16051215229...	wkwkw nape memang kamu takut kesaingin apa beg...	-1.0	lah wkwkw, nape emg? lu takut kesaingin apa be...
3797	2022-12-19 05:50:13	meimorable	https://twitter.com/meimorable/status/16047157...	allah baru sadar typo pildun	NaN	Ya Allah baru sadar ini pun masih typo. PILDU...
3650	2022-12-19 06:20:44	chwerryjubilee	https://twitter.com/chwerryjubilee/status/1604...	min kemarin menang pildun alias piala dunia	NaN	@wonhoom Ini min kmrn menang pildun alias pial...
1804	2022-12-19 15:44:48	otewenih	https://twitter.com/otewenih/status/1604865337...	ada ribut2 pildun	-1.0	masih ada yang ribut* tentang pildun?
1474	2022-12-19 22:54:42	denihasan89	https://twitter.com/denihasan89/status/1604973...	karena adek kamu tidak bisa ngapa2 in pildun	0.0	@KompasTV Ya karna adek lu ga bisa ngapa2 in d...

03

Exploratory Data Analysis



Jumlah *Tweet* Berdasarkan Hari dan Jam.

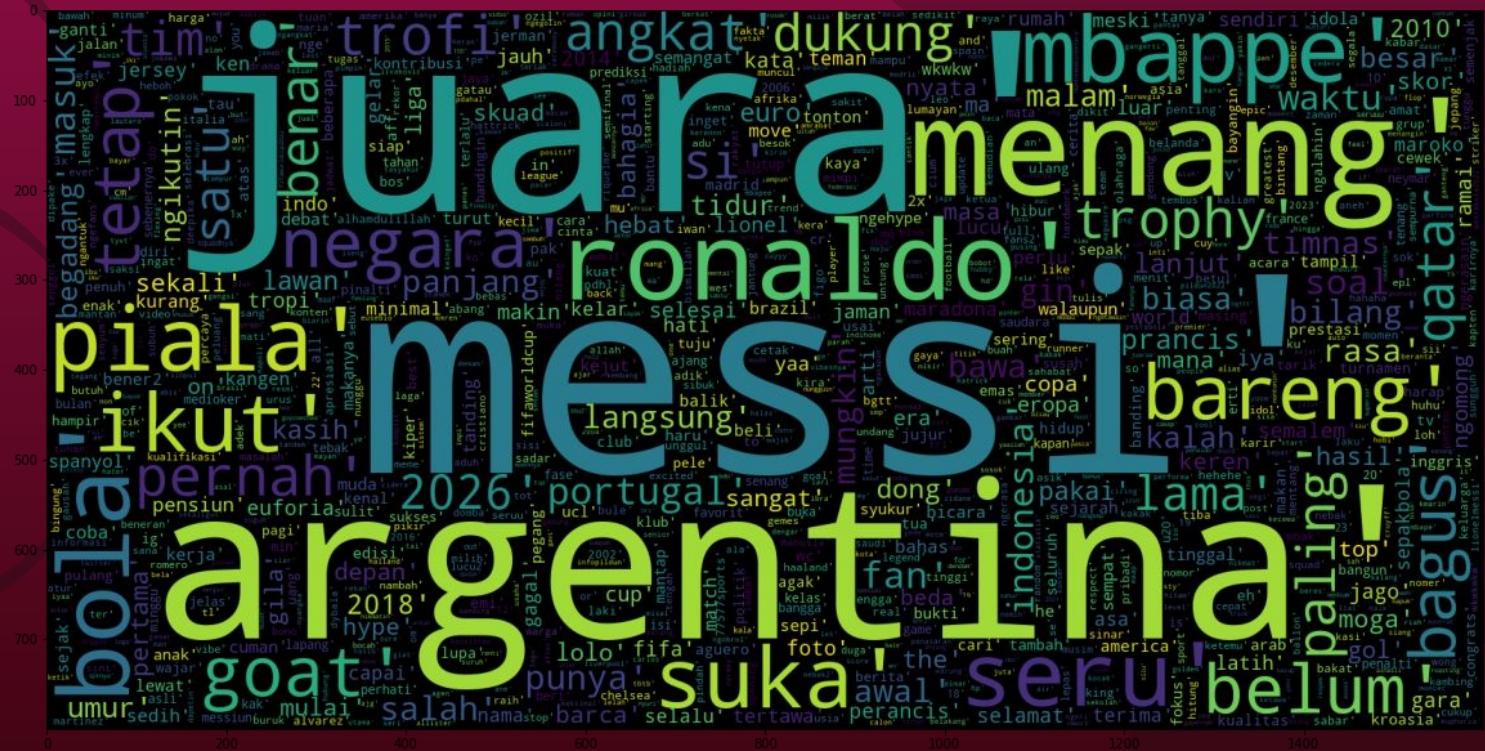


- Pengguna Twitter paling banyak menulis *tweet* mengenai “pildun” pada tanggal 18 Desember 2022 pukul 6 sore, beberapa jam sebelum final piala dunia dimulai.
- Pada tanggal 19 Desember 2022, pengguna Twitter paling banyak menulis *tweet* pada pukul 1 pagi, yaitu saat final piala dunia berlangsung.
- Pada tanggal 20 Desember 2022, beberapa hari setelah final selesai, *tweet* mengenai “pildun” mengalami pengurangan dibandingkan hari sebelumnya.

[illegible]

11

Positive Word Cloud



Positive tweet : Juara, Argentina, Messi, Menang, Mbappe

13

[illegible]

14

04

Sentiment Analysis

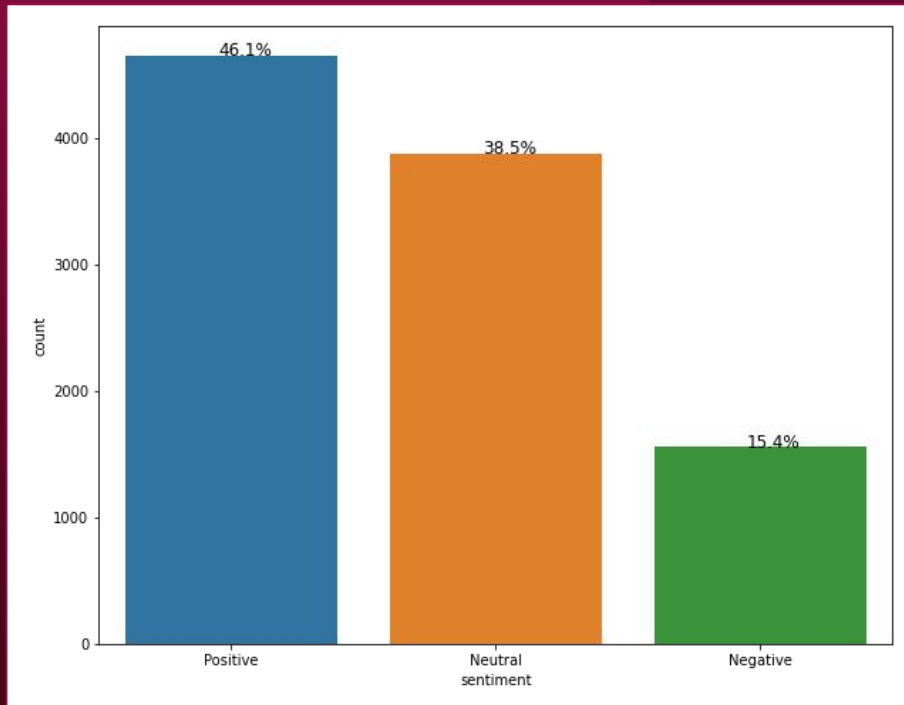


Subjektivitas dan Polaritas

	time_hour	subjectivity	polarity
count	10090.000000	10090.000000	10090.000000
mean	11.980575	0.395931	0.096895
std	7.402695	0.332971	0.242798
min	0.000000	0.000000	-1.000000
25%	5.000000	0.000000	0.000000
50%	13.000000	0.400000	0.000000
75%	18.000000	0.650000	0.230417
max	23.000000	1.000000	1.000000

- Berdasarkan subjektivitas, rata-rata mendapatkan nilai 0,39 yang berarti sentiment cenderung bersifat objektif.
- Berdasarkan polaritas, rata-rata mendapat nilai 0,09 yang berarti sentiment cenderung netral hingga positif.

Sentimen pada Data



Sebagian besar data merupakan sentimen positif.

Model Building

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Logistic Regression	42.12	41.98	42.12	42
Naive Bayes	41.96	41.75	41.97	41.8
SVM	42.12	41.84	41.12	41.86

05

Social Network Analysis

Social Network Analysis

Centrality by Degree



Social Network Analysis

Centrality by Degree

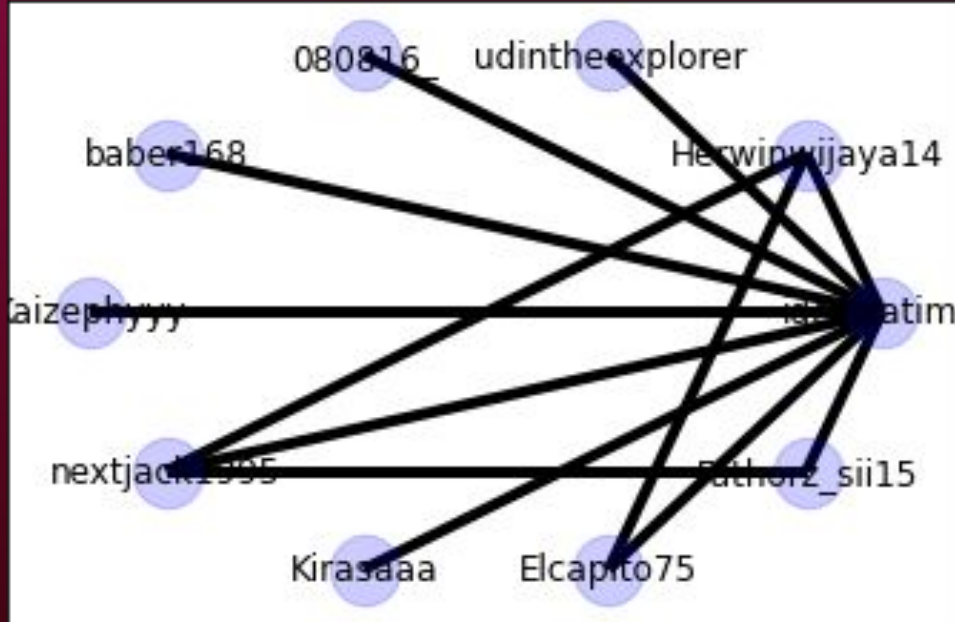
No.	Nama Pengguna	Centrality by Degree
1.	@idextratime	0.07233
2.	@convomfs	0.00162
3.	@Askrlfess	0.01325
4.	@FaktaSepakBola	0.00893
5.	@The	0.00703

No.	Nama Pengguna	Centrality by Degree
6.	@SiaranBolaLive	0.00667
7.	@tanyakanrl	0.00586
8.	@realmadridindo1	0.00451
9.	@utdfocusid	0.00415
10.	@bintangemon	0.00325

- Tabel di atas menunjukkan sepuluh akun yang memiliki *centrality by degree* tertinggi.
- Akun @idextratime memiliki nilai paling tinggi.
- Selain pengguna biasa, akun *autobase* dan *fanbase* juga memberikan pengaruh yang cukup besar.
- Salah satu *comedian*, aktor, dan selebriti internet, yaitu Bintang Emon juga masuk ke dalam 10 akun berpengaruh dalam penyebaran informasi mengenai Final Piala Dunia 2022.

Social Network Analysis

Closeness Centrality



Social Network Analysis

Closeness Centrality

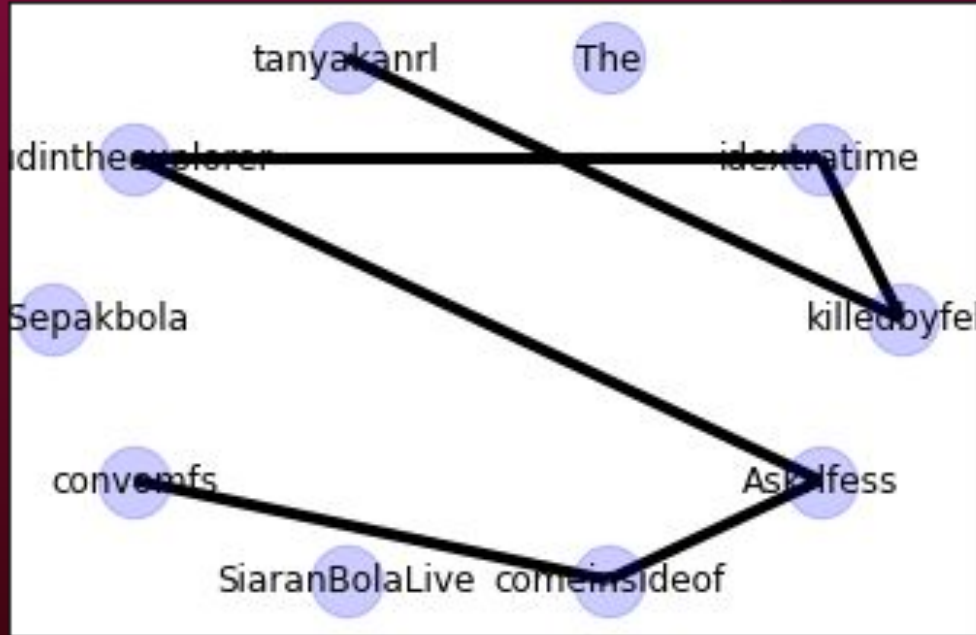
No.	Nama Pengguna	Closeness Centrality
1.	@idextratime	0.10713
2.	@Elcapito75	0.08468
3.	@Kirasaaa	0.08459
4.	@udintheexplorer	0.08451
5.	@babber168	0.08446

No.	Nama Pengguna	Closeness Centrality
6.	@Kaizephyyy	0.08358
7.	@080816_	0.08356
8.	@nextjack1995	0.08314
9.	@Herwinwijaya14	0.08311
10.	@Fathorz_sii15	0.08294

- Tabel di atas menunjukkan sepuluh akun yang memiliki *closeness centrality* tertinggi.
- Akun @idextratime memiliki nilai paling tinggi.
- Akun-akun lain yang menduduki 10 akun teratas adalah akun-akun pengguna biasa.

Social Network Analysis

Betweenness Centrality



Social Network Analysis

Beetweenness Centrality

No.	Nama Pengguna	Beetweenness Centrality
1.	@idextratime	0.07222
2.	@Askrlfess	0.01568
3.	@convomfs	0.01017
4.	@FaktaSepakBola	0.00794
5.	@udintheexplorer	0.00675

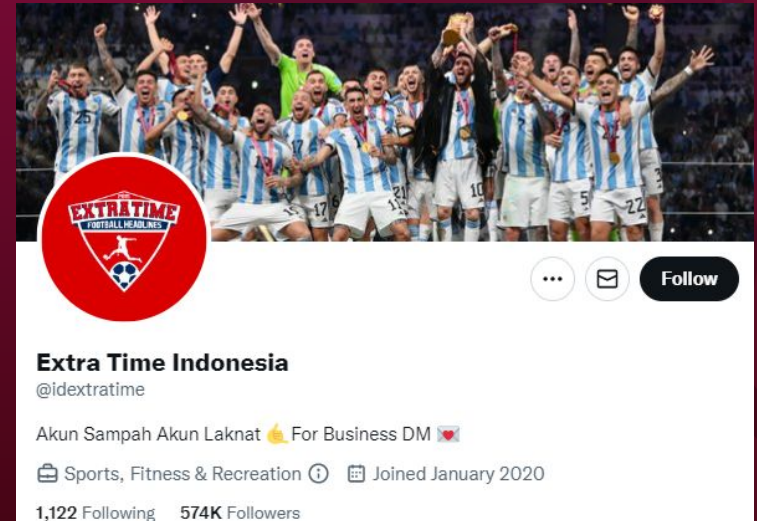
No.	Nama Pengguna	Beetweenness Centrality
6.	@SiaranBolaLive	0.00659
7.	@The	0.00571
8.	@tanyakanrl	0.00464
9.	@killedbyfeb	0.00422
10.	@comeinsideof	0.00402

- Tabel di atas menunjukkan sepuluh akun yang memiliki nilai *beetweenness centrality* tertinggi.
- Seperti halnya pada *centrality by degree* dan *closeness centrality*, akun @idextratime memiliki nilai paling tinggi.
- Selain akun pengguna biasa, akun-akun yang menduduki 10 teratas lainnya adalah akun *fanbase* dan *autobase*.

Social Network Analysis

Kesimpulan

- Dilihat dari *centrality by degree*, *closeness centrality*, dan *betweenness centrality*, akun @idextratime memiliki nilai tertinggi, yaitu akun yang memiliki pengaruh besar dalam penyebaran informasi mengenai Final Piala Dunia 2022 di Twitter.
- Akun @idextratime memiliki nilai paling tinggi, artinya akun ini memiliki jumlah *retweet* terbanyak, lebih cepat dalam menyampaikan informasi karena memiliki kedekatan dengan pengguna lainnya, dan unggul dalam hubungannya dengan pengguna berpengaruh lain di dalam jaringan Twitter.
- Akun @idextratime merupakan akun media berita olahraga yang aktif dalam membahas informasi terbaru terkait sepak bola dan memiliki pengikut sebanyak 574 ribu.



06

Kesimpulan dan Rekomendasi

KESIMPULAN

- Sebagian besar sentimen merupakan sentimen positif, yaitu sebesar 46,1%, dengan traffic tweet pada pukul 18.00
- Kata 'Messi', 'Juara', 'Argentina', 'Menang', 'Mbappe' menjadi kata yang paling banyak dibicarakan
- Rata-rata tweet cenderung bersifat objektif dengan sentimen cenderung netral hingga positif
- Model terbaik yang digunakan yaitu regresi logistik karena memiliki akurasi tertinggi sebesar 42,12% dan presisi, recall, f1 score terbesar.
- Akun twitter yang berpengaruh adalah @idextratime karena memiliki jumlah *retweet* terbanyak, lebih cepat dalam menyampaikan informasi karena memiliki kedekatan dengan pengguna lainnya, dan unggul dalam hubungannya dengan pengguna berpengaruh lain di dalam jaringan Twitter.

INSIGHT & REKOMENDASI

- Kata 'Messi' dan 'Argentina' memiliki pengaruh yang besar pasca piala dunia, dapat dipertimbangkan untuk menjadi headline
- Waktu traffic pukul 18.00 dapat dipertimbangkan dalam social media *analyst* untuk meng-upload konten ataupun berita
- Akun-akun lain yang berpengaruh dalam penyebaran informasi terkait sepak bola adalah akun *autobase*, yaitu akun dimana pengguna dapat mengirimkan *tweet* secara anonim melalui *direct message*. Akun *autobase* kini berkembang pesat dan menjadi wadah masyarakat untuk beropini.
- Akun *fanbase* juga merupakan akun yang berpengaruh, ditandai dengan akun @FaktaSepakBola, @SiaranBolaLive, @utdfocusid, dan @realmadridindo1 yang menduduki 10 teratas akun paling berpengaruh. Dapat dikatakan terdapat tendensi bahwa sepakbola merupakan salah satu olahraga yang diminati oleh masyarakat Indonesia dan memiliki kekuatan penggemar yang besar sehingga dapat dipertimbangkan untuk media pengiklanan dengan target audiens penggemar sepak bola.

Terima Kasih

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**



***Link* Google Colab dan Video Presentasi**

Link Google Colab

https://colab.research.google.com/drive/1LaufV_cyj7QTY24ermxyfX3SqYsDGMaR?usp=sharing

Link Video Presentasi

<https://bit.ly/VideoKelompokGDatMin>