

ANALISIS SENTIMENT MASYARAKAT SELAMA BULAN RAMADHAN DALAM MENGHADAPI PANDEMI COVID-19

Ardan Fahmi¹⁾, Irfan Ramadhan²⁾, Agussalim³⁾

E-mail: ¹⁾ardanfahmi19@gmail.com, ²⁾iromadhon15@outlook.com,

³⁾agussalim.si@upnjatim.ac.id

^{1,2,3)}Program Studi Sistem Informasi, Fakultas Ilmu Komputer, UPN "Veteran" Jawa Timur

Abstrak

Pada awal tahun 2020 suatu wabah penyakit berkembang dengan luar biasa hampir menginfeksi seluruh dunia. Penyakit ini adalah virus corona atau secara internasional disebut Coronavirus disease (COVID-19). Penyakit tersebut dapat mengakibatkan masalah lain diluar dari masalah medis yaitu masalah pada bidang sosial media dimana jutaan masyarakat Indonesia beraktivitas hanya dirumah dan menimbulkan perasaan bosan. Dari perasaan bosan ini masyarakat pengguna internet dapat menuangkan perasaan ke sosial media sehingga berpotensi menimbulkan tulisan yang tidak baik. Sehingga dilakukannya penelitian ini berfokus pada masalah yang terjadi pada sosial media khususnya Twitter untuk melihat tulisan yang dibuat oleh masyarakat Indonesia selama pandemi. Metode yang digunakan dalam penelitian ini tahapannya dimulai dari memperoleh data menggunakan teknik web scraping python, langkah sampling data exploration memilih 10 kota sebagai target analisis, langkah sentiment analysis melakukan analisis pendapat orang-orang yang diungkapkan oleh teks dan dapat divisualisasikan menjadi *Wordcloud*, langkah topik modeling untuk memilih kegiatan terkait yang dilakukan dari setiap topik tersebut, dan langkah terakhir yaitu Suggestions adalah dari berbagai topik yang telah diperoleh aktivitas yang sering dilakukan, dapat dijadikan saran untuk aktivitas yang dilakukan pada masa pandemi. Hasil akhir yang diperoleh yaitu terdapat kata yang sering digunakan dalam tweet yang memperoleh nilai Polaritas yang tinggi mencerminkan bahwa kata tersebut adalah kata yang positif dan dapat dijadikan saran aktivitas selama pandemi yaitu 'ramadhan', 'semangat', 'sabar', dan 'promo'.

Kata kunci: COVID-19, Twitter, Sentimen analisis, Big Data, Python, NLP

1. PENDAHULUAN

Pada awal tahun 2020 dunia dikejutkan dengan suatu wabah penyakit, penyakit ini berkembang dengan luar biasa hampir menginfeksi pada seluruh dunia. Penyakit ini adalah virus corona atau secara internasional disebut Coronavirus disease (COVID-19). Menurut lembaga kesehatan dunia, WHO semenjak Januari 2020 telah menyatakan dunia masuk ke dalam darurat global terkait virus ini [1]. Indonesia adalah termasuk salah satu Negara yang terinfeksi virus ini ditandai semenjak pada bulan Februari 2020 terdapat salah satu Warga Negara Indonesia yang terinfeksi virus tersebut. Hingga pada bulan Mei 2020 dikutip dari laman resmi Gugus Tugas Percepatan Penanganan COVID-19 data orang yang terinfeksi positif virus corona sebesar 24.538 orang.

Dari kondisi tersebut pemerintah Indonesia telah mengambil beberapa kebijakan sebagai upaya untuk mencegah penyebaran virus corona. Dimulai pada akhir bulan Februari 2020 pemerintah Indonesia telah menetapkan status darurat bencana terkait pandemi virus ini, dan terus berlangsung serta diikuti dengan penyebaran kebijakan yang dilakukan pada beberapa wilayah di Indonesia karena mengingat virus telah menyebar pada beberapa daerah hingga terjadi sampai pada bulan Mei 2020 saat ini. Kebijakan lain juga diterapkan oleh pemerintah Indonesia yaitu penerapan Physical Distancing suatu konsep untuk dapat mengurangi bahkan memutus mata rantai infeksi Covid-19 seseorang harus menjaga jarak aman dengan manusia lainnya minimal 2 meter, tidak melakukan kontak langsung dengan orang lain, dan menghindari pertemuan massal [1]. Penerapan Physical Distancing dinilai sebagai pondasi terkait munculnya beberapa kebijakan lain yang diambil oleh pemerintah Indonesia sebagai upaya untuk mencegah penyebaran virus corona, diantaranya yaitu kebijakan Work From Home untuk para pekerja yang sebelumnya bekerja pada tempat kerja atau kantor menjadi bekerja pada rumah masing-

masing, kebijakan menutup seluruh tempat lembaga pendidikan, kebijakan menutup tempat peribadatan, kebijakan larangan untuk mudik lebaran, dan lain-lain.

Tetapi dengan kebijakan yang diambil oleh pemerintah banyak masyarakat yang tidak menyikapi dengan baik, seperti contohnya pemerintah telah meliburkan siswa dan mahasiswa untuk tidak berkuliah atau bersekolah ataupun memberlakukan bekerja di dalam rumah, namun kondisi ini dimanfaatkan oleh banyak masyarakat untuk berlibur atau berada pada tempat yang menimbulkan kerumunan [1].

Tidak hanya masalah yang dijumpai pada kehidupan nyata tetapi masalah lain juga timbul pada dunia maya atau pada bidang sosial media. Menurut data statistika pada keseluruhan tahun 2019 total pengguna internet di Indonesia diproyeksikan menjadi 107,2 juta pengguna. Hal ini menimbulkan permasalahan baru mengingat kondisi pada saat pandemi virus corona masyarakat Indonesia beraktivitas hanya dirumah dan menimbulkan perasaan bosan. Dari perasaan bosan ini masyarakat pengguna internet dapat menuangkan perasaan ke sosial media mereka masing-masing sehingga berpotensi menimbulkan tulisan yang tidak baik seperti kalimat yang isinya menjatuhkan pemerintah, perkelahian antar individu, menjatuhkan suku, ras, agama orang lain, dan lain-lain.

Untuk mengatasi masalah-masalah yang terjadi maka penelitian ini dibuat berfokus pada masalah yang terjadi pada sosial media khususnya Twitter untuk melihat tulisan yang dibuat oleh masyarakat Indonesia selama pandemi virus corona, menganalisis sentimen dari tulisan tersebut hasilnya akan diperoleh berapa banyak tulisan yang memiliki sentimen positif, netral, maupun negatif dan dapat divisualisasikan menjadi *Wordcloud*, menganalisis topik apa yang disenangi oleh masyarakat selama masa pandemi, serta dapat memberikan saran hal-hal positif yang dapat dilakukan selama masa pandemi.

2. STUDI LITERATUR

2.1 Big Data

Big Data secara singkat adalah kemampuan untuk mengelola sejumlah data besar yang berbeda, pada kecepatan yang tepat, dan dalam kerangka waktu yang tepat untuk memungkinkan analisis dan reaksi waktu nyata. Big Data juga dapat didefinisikan sebagai kombinasi teknologi yang bisa memanajemen data yang bervariasi dalam jumlah besar, kecepatan yang tepat, dan memiliki ketepatan saat melakukan analisis dan reaksi. Tiga karakteristik yang dimiliki Big Data, yaitu volume, velocity, dan variety [2].

2.2 Python

Python adalah bahasa yang cocok untuk pembelajaran dan pemrograman dunia nyata. Python adalah bahasa pemrograman berorientasi objek tingkat tinggi yang kuat yang dibuat oleh Guido van Rossum [3].

Konstruksi bahasa memungkinkan pengguna untuk menulis program yang jelas pada skala kecil dan besar [4]. Fitur terpenting dalam Python adalah ia mendukung banyak paradigma pemrograman, termasuk pemrograman berorientasi objek, imperatif dan fungsional atau gaya prosedural.

2.3 Google Colaboratory

Google Colaboratory (juga dikenal sebagai Colab) adalah layanan *cloud* yang didasarkan pada Jupyter Notebooks untuk menyebarkan pendidikan dan penelitian *machine learning*. Ini memberikan *runtime* yang sepenuhnya dikonfigurasi untuk *deep learning* dan akses gratis ke GPU yang tangguh.

Google telah menciptakan Colaboratory (biasa disebut Colab), layanan *cloud* untuk menyebarkan pendidikan dan penelitian *machine learning*. Runtime yang disediakan oleh layanan *cloud* ini sepenuhnya dikonfigurasi dengan pustaka kecerdasan buatan (AI) terkemuka dan juga menawarkan GPU yang kuat. Layanan Google ini ditautkan ke akun Google Drive, dan gratis [5].

2.4 NLP

Natural Language Processing adalah kemampuan teknologi komputasi dan atau linguistik komputasi untuk memproses bahasa alami manusia. Natural language processing adalah bidang ilmu komputer, kecerdasan buatan, dan linguistik komputasi yang dipadukan dengan interaksi antara komputer dan bahasa alami manusia.

Natural language processing dapat didefinisikan sebagai pemrosesan otomatis bahasa alami manusia. Dalam industri, orang membutuhkan NLP untuk analisis pasar, pengembangan perangkat lunak web, dan lain-lain [6].

2.5 Web Scraping

Pemrograman web scraping didesain untuk mengambil semua data penting dari berbagai toko online dan mengumpulkannya ke situs web baru. Alat *scraping* untuk web digunakan untuk mendapatkan informasi yang berasal dari host web, dan sebagai bagian dari penggunaan yang digunakan untuk pesanan web, penambahan web dan penambahan data, pengamatan perubahan harga online dan korelasi nilai, area perubahan halaman web, memeriksa, reputasi web, penggabungan web dan, penggabungan data web [7].

2.6 Sampling Data Exploration

Data Exploration terdiri dari beberapa upaya untuk menemukan wawasan yang menarik terkait dengan topik sumber data yang menarik. Hasilnya dapat berupa grafik yang diperoleh tingkat signifikansi yang tampak pada sumber data [8].

2.7 Sentiment Analysis

Analisis Sentimen dapat dikategorikan ke dalam dua pendekatan: pendekatan berbasis Machine Learning dan pendekatan berbasis Leksikon. Pendekatan berbasis Machine Learning mengklasifikasikan teks menggunakan algoritma klasifikasi, sedangkan pendekatan berbasis Leksikon menggunakan kamus sentimen dengan kata-kata pendapat dan mencocokkannya dengan data untuk menentukan polaritas. Nilai-nilai sentimen diberikan pada kata-kata yang menggambarkan sikap positif, negatif dan netral pembicara [9].

2.8 Wordcloud

Wordcloud adalah sejenis daftar berbobot untuk memvisualisasikan data bahasa atau teks, yang mendapatkan perhatian yang meningkat dan lebih banyak peluang aplikasi saat pendekatan big data time. Ini adalah cabang penting dari Data mining, yang telah mendapatkan perhatian yang meningkat dan lebih banyak peluang aplikasi. Saat ini, ada beberapa generator cloud kata online yang tersedia untuk pengguna dengan permintaan sederhana, seperti mengulangi frasa yang tepat, atau mengumpulkan data teks dari halaman web [10].

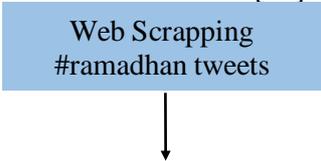
2.9 Topic Modeling

Pemodelan topik adalah metode Data mining yang digunakan untuk memahami dan mengelompokkan korporasi data yang besar. Dengan demikian, ini adalah alat yang dapat digunakan pustakawan teologis dalam alur kerja profesional.

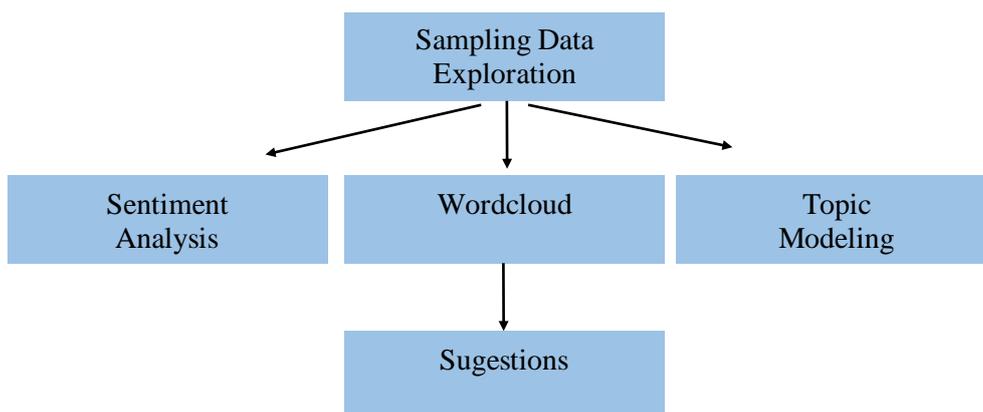
Model topik memodelkan kumpulan dokumen, bukan dengan menurunkannya, tetapi dengan menghasilkan topik yang mewakili konten kumpulan dokumen. Untuk lebih spesifik, model topik adalah model probabilistik yang digunakan untuk menemukan topik, atau struktur laten, di kumpulan dokumen [11].

3. METODOLOGI

Metodologi penelitian yang digunakan merujuk pada praktikum yang telah dilakukan oleh seorang mahasiswa dari University of California bidang Business Analytics yang diperoleh dari laman web Github [12].



Web Scraping
#ramadhan tweets



Gambar 1. Bagan Metodologi

Berdasarkan bagan metodologi tersebut pada langkah pertama yaitu melakukan Web Scraping untuk mendapatkan seluruh tweet yang menggunakan tagar #ramadhan. Pada penelitian ini mencari kata ramadhan karena pada saat pandemi bertepatan dibulan ramadhan dan kami ingin melihat bagaimana sentimen masyarakat menanggapi pandemi pada saat bulan ramadhan. Pada langkah sampling data exploration kami memilih 10 kota sebagai target kami, karena orang-orang di kota tersebut lebih mungkin mengalami depresi karena terdampak oleh pandemi. Pada langkah sentiment analysis adalah kami mulai melakukan analisis dari data yang telah diperoleh, kami menganalisis tweet dan mencoba memahami pendapat orang-orang yang diungkapkan oleh teks, dan hasilnya akan dapat mengukur sentimen dibalik setiap tweet dengan nilai positif, netral, atau negatif. Setelah membagi tweet menjadi kategori positif, negatif, dan netral, kami menghitung banyak frekuensi kata yang paling sering banyak muncul dari setiap tweet dan dari setiap kategori serta menjadikannya Wordcloud. Pada langkah topik modeling kami melakukan *sorting* dari ribuan tweet menjadi 30 topik karena topik tersebut cocok untuk data yang telah kami miliki, selanjutnya kami menggunakan kata kunci, yang merupakan kata umum dalam setiap topik, untuk memilih kegiatan terkait yang dilakukan dari setiap topik tersebut. Dan langkah terakhir yaitu Sugestions adalah dari berbagai topik yang telah diperoleh aktivitas yang sering dilakukan, dapat dijadikan saran untuk aktivitas yang dilakukan pada masa pandemi.

4. HASIL DAN PEMBAHASAN

3.1 Web Scraping

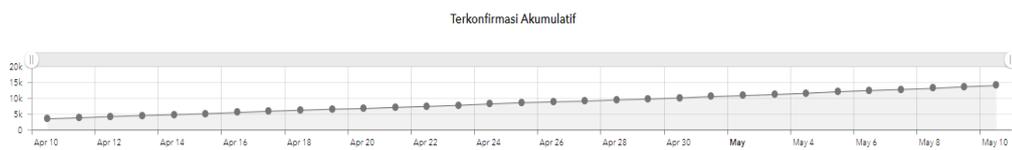
Proses web scraping dilakukan menggunakan bahasa pemrograman Python untuk mendapatkan seluruh tweet yang menggunakan tagar #ramadhan. Pada langkah ini juga menentukan kriteria tweet apa saja yang diambil, dan hasilnya kami mengambil data username, text(tweet), date, retweets, favorites, dan hashtags.

3.2 Sampling Data Exploration

Data Exploration dilakukan dengan memilih 10 kota besar yang terdampak virus corona dengan total kasus lebih dari sekitar 10.000 kasus positif, kota tersebut yaitu Bandung, Denpasar, Jakarta, Makassar, Palembang, Pekanbaru, Semarang, Sidoarjo, Solo, dan Surabaya. Sementara itu, mereka memiliki cukup pengguna Twitter untuk memenuhi kebutuhan data. Lalu menggunakan Python untuk mengambil tweet dari pengguna yang tinggal di kota-kota ini. Secara khusus, melihat tagar '#ramadhan', karena pada saat pandemi bertepatan dibulan ramadhan dan kami ingin melihat bagaimana sentimen masyarakat menanggapi pandemi pada saat bulan ramadhan. Hasilnya diperoleh sekitar 1.500 tweet antara 10 April sampai 10 Mei 2020.



Gambar 2. Peta sebaran virus corona di Indonesia

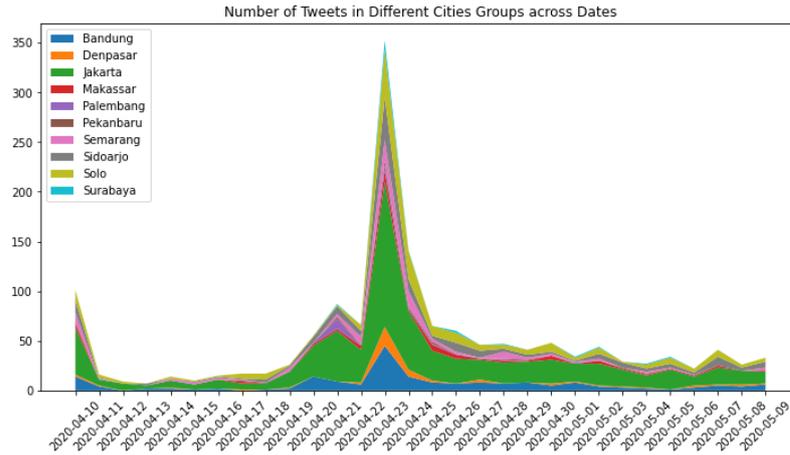


Gambar 3. Jumlah kasus positif virus corona di Indonesia



Gambar 4. Tweets dengan tagar ramadhan

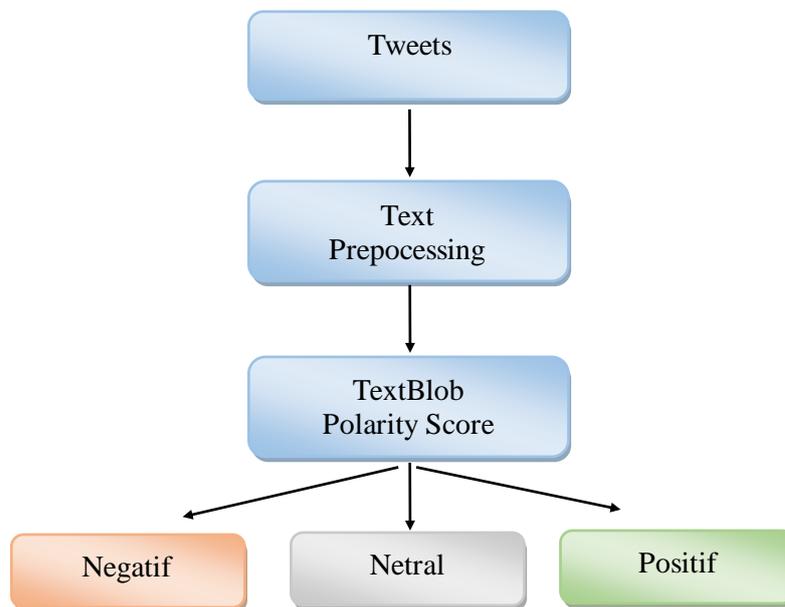
Selanjutnya pada gambar dibawah ini menjelaskan plot area menunjukkan dari tweet di 10 kota, terdapat peningkatan yang signifikan dalam jumlah tweet sekitar akhir april. Hal tersebut dapat di kaitkan bahwa beberapa kota tersebut mulai mem berlakukan kebijakan saat pandemi, jadi orang-orang mulai lebih banyak beraktivitas dari rumah dan lebih aktif di Twitter. Sebagai contoh pada kota Jakarta telah diberlakukan kebijakan PSBB yang dimulai pada hari Jumat 10 April 2020, hal tersebut dapat dilihat pada data terjadi peningkatan yang signifikan dalam jumlah Tweet mulai pada tanggal tersebut.



Gambar 5. Jumlah tweets per-tanggal dari berbagai kota

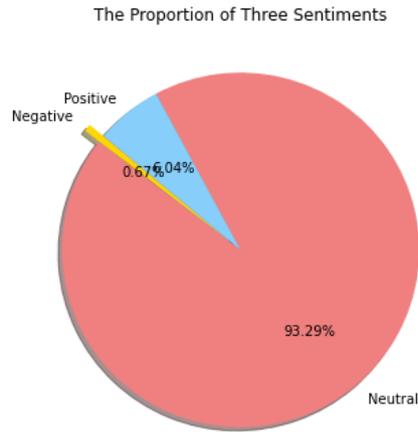
3.3 Sentiment Analysis

Sentiment Analysis dilakukan dengan menganalisis tweet dan mencoba memahami pendapat orang-orang yang diungkapkan oleh teks. Menggunakan bahasa Python pertama dilakukan langkah Text Preprocessing untuk menghapus karakter yang tidak diinginkan dari data tweet, lalu dengan menggunakan library TextBlob di Python dapat mengukur sentimen di balik setiap tweet dengan nilai positif, netral, atau negatif, yang disebut Polarity.

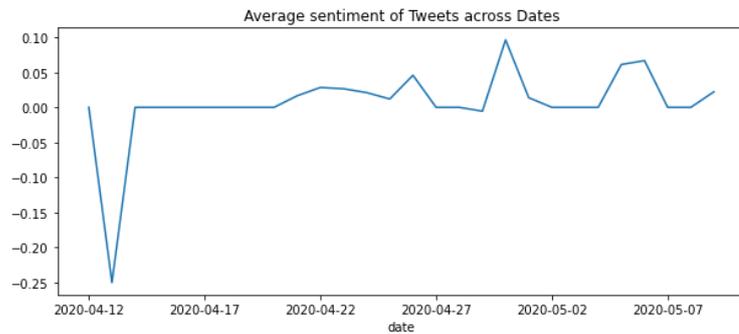


Gambar 6. Langkah Sentiment Analysis

Dalam diagram lingkaran dibawah ini telah diperoleh hasil sentiment analysis, 93% dari tweet menghasilkan nilai netral, 6% menghasilkan positif, dan negatif sebesar 0,67%. Terdapat juga hasil skor sentimen per-tanggal dan dapat terlihat bahwa tweet orang-orang menjadi lebih positif dilihat dari nilai Polarity yang relatif naik.



Gambar 7. Dia gram lingkaran sentiment analysis



Gambar 8. Grafik nilai skor sentiment analysis

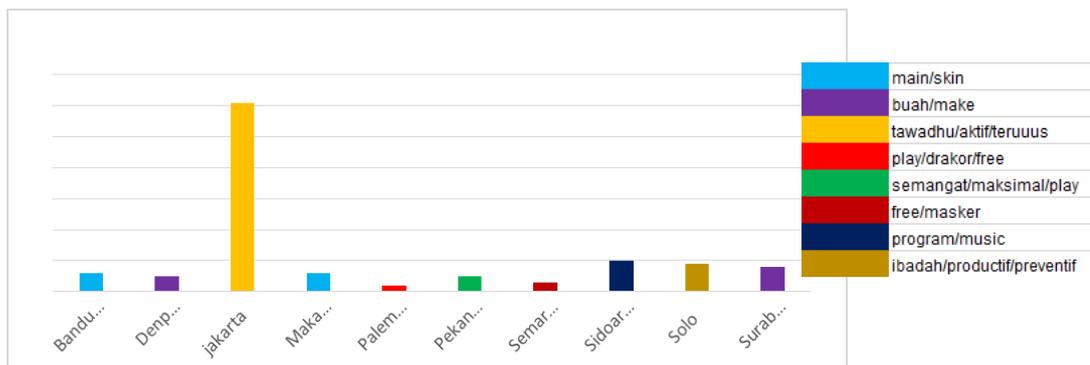
3.4 Wordcloud

Setelah mendapat kategori dari data tweet berdasarkan nilai Polaritas, Wordcloud dilakukan dengan menghitung frekuensi kata dari setiap kategori dan menggambarkan Wordcloud untuk masing-masing kategori. Hasilnya dari kategori positif dapat dilihat bahwa tweet sering mengandung kata seperti “ramadhan”, “sale”, “puasa”, “kuat”. Dari kategori negatif tweet yang sering terlihat seperti “gaada”, “pandemi”, “usaha”. Dan dari kategori netral tweet yang sering terlihat seperti “ramadhan”, “ibadah”, “puasa”.

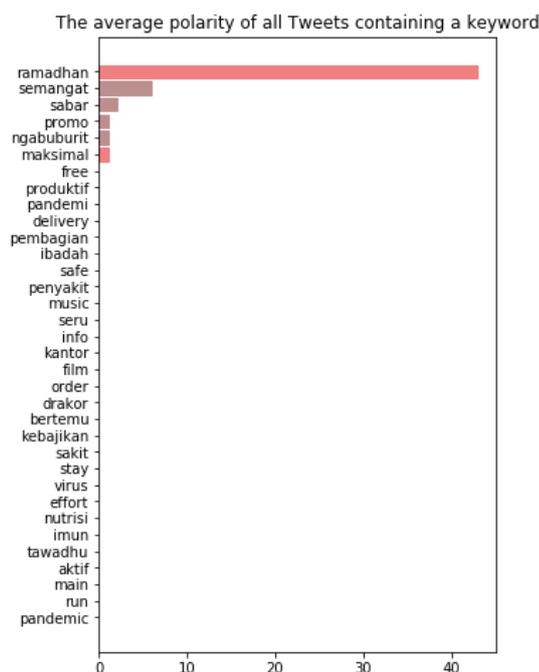
Word Clouds with Positive Polarity



Gambar 9. Wordcloud kategori positif



Gambar 13. Kegiatan yang sering dilakukan dari berbagai kota



Gambar 14. Ranking nilai Polaritas dari setiap kata

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan hasil dan pembahasan yang diperoleh didapatkan kesimpulan yaitu, data yang kami gunakan dalam penelitian ini diperoleh menggunakan teknik web scraping python dengan mengambil data tweet yang menggunakan tagar #ramadhan pada tanggal 10 April sampai 10 Mei 2020. Dari data tersebut digunakan sentimen analisis yaitu 93% menghasilkan nilai netral, 6% menghasilkan positif, dan negatif sebesar 0,67%, serta didapatkan grafik yang menunjukkan per-tanggal menjadi lebih positif. Dari data tersebut juga mendapatkan hasil kegiatan yang sering dilakukan berdasarkan topik dari berbagai kota yaitu di kota Bandung terdapat kata 'main/skin', di kota Denpasar terdapat kata 'buah/make', dan lain-lain. Dan yang terakhir topik tersebut dapat diperoleh nilai Polaritas yang mencerminkan bahwa nilai yang tinggi adalah kata yang positif dan dapat dijadikan saran aktivitas selama pandemi yaitu 'ramadhan', 'semangat', 'sabar', dan 'promo'.

4.2 Saran

Pada penelitian ini data yang digunakan menggunakan teknik web scraping python untuk mengambil data tweet dari twitter. Dengan melihat data tweet yang diperoleh

terdapat banyak kata yang tidak merupakan kata baku, maka saran untuk penelitian selanjutnya tidak hanya mengambil data dari twitter melainkan dapat menggunakan data dari situs berita, situs komunitas, dan lain-lain.

5. DAFTAR RUJUKAN

- [1] B. S. Mózo, “*濟無*No Title No Title,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2017, doi: 10.1017/CBO9781107415324.004.
- [2] *www.it-ebooks.info.*
- [3] K. R. Srinath, “Python – The Fastest Growing Programming Language,” *Int. Reasearch J. Eng. Technol.*, vol. 4, no. 12, pp. 354–357, 2017.
- [4] D. KUHLMAN, “A Python Book,” *A Python B.*, pp. 1–227, 2013.
- [5] T. Carneiro, R. V. M. Da Nobrega, T. Nepomuceno, G. Bin Bian, V. H. C. De Albuquerque, and P. P. R. Filho, “Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications,” *IEEE Access*, vol. 6, pp. 61677–61685, 2018, doi: 10.1109/ACCESS.2018.2874767.
- [6] W. Wagner, “Steven Bird, Ewan Klein and Edward Loper: Natural Language Processing with Python, Analyzing Text with the Natural Language Toolkit,” *Lang. Resour. Eval.*, vol. 44, no. 4, pp. 421–424, 2010, doi: 10.1007/s10579-010-9124-x.
- [7] D. M. Thomas and S. Mathur, “Data Analysis by Web Scraping using Python,” *Proc. 3rd Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2019*, pp. 450–454, 2019, doi: 10.1109/ICECA.2019.8822022.
- [8] Z. Zhao, L. De Stefani, E. Zraggen, C. Binnig, E. Upfal, and T. Kraska, “Controlling false discoveries during interactive data exploration,” *Proc. ACM SIGMOD Int. Conf. Manag. Data*, vol. Part F1277, pp. 527–540, 2017, doi: 10.1145/3035918.3064019.
- [9] F. Nausheen and S. H. Begum, “Using Python,” *2018 2nd Int. Conf. Inven. Syst. Control*, no. Icisc, pp. 1259–1262, 2018, doi: 10.1109/ICISC.2018.8399007.
- [10] Y. Jin, “Development of Word Cloud Generator Software Based on Python,” *Procedia Eng.*, vol. 174, pp. 788–792, 2017, doi: 10.1016/j.proeng.2017.01.223.
- [11] F. Torabi Asr and M. Taboada, “Big Data and quality data for fake news and misinformation detection,” *Big Data Soc.*, vol. 6, no. 1, pp. 1–14, 2019, doi: 10.1177/2053951719843310.
- [12] “GitHub - DaiLinyan/4g-covid19-challenge: Twitter Sentiment Analysis.” [Online]. Available: <https://github.com/DaiLinyan/4g-covid19-challenge>. [Accessed: 29-May-2020].