

Pengelompokan Komentar Netizen pada Media Sosial Pemerintah Daerah Berdasarkan Frekuensi Kata Kunci

Nur Aini Rakhmawati* dan Nody Risky Pratomo

Departemen Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi

Institut Teknologi Sepuluh Nopember

Surabaya

*nur.aini@is.its.ac.id

Abstrak—Menurut survei internetworldstats yang dirilis pada Januari 2018, jumlah pengguna internet di Indonesia telah mencapai 132,7 juta jiwa. Dari jumlah tersebut, 40% di antaranya merupakan pengguna aktif media sosial. Hal ini mengakibatkan peningkatan penerapan konsep *e-government* pada pemerintah dengan media sosial. Melalui media sosial, masyarakat dapat melakukan interaksi kepada akun media sosial pemerintah kota dalam memberikan informasi atau kritik dan saran terkait kotanya. Melalui fenomena tersebut, diperlukan aplikasi untuk melakukan pengelompokan komentar masyarakat di media sosial berdasarkan Satuan Kerja Perangkat Daerah (SKPD). Proses kategorisasi dilakukan dengan melihat kesesuaian kata kunci SKPD dengan komentar pada media sosial Facebook, Twitter, dan Youtube akun resmi pemerintah kota dan *influencer*. Selain itu, aplikasi ini dapat memberikan rekomendasi kata kunci berdasarkan komentar yang masuk. Proses rekomendasi dilakukan dengan cara mengambil kata kunci yang memiliki frekuensi terendah. Dalam melakukan pengujian, digunakan pemerintah kota Depok sebagai contoh. Hasil pengujian menunjukkan bahwa dari 4.325 komentar yang telah dilakukan *preprocessing*, 10,35% komentar berhasil dikategorikan sesuai SKPD. Pengujian rekomendasi kata turut dilakukan dengan mengambil dinas dalam pemerintahan kota yang sama. Dari 19 dinas, terdapat 3 dinas yang mendapatkan rekomendasi kata. Dari 3 dinas, 1 dinas memiliki kata kunci yang relevan dan 2 sisanya tidak relevan dengan SKPD. Hal ini disebabkan tidak tepatnya kata kunci SKPD untuk melakukan kategorisasi serta jumlah komentar terkategorisasi yang masih sedikit.

Kata kunci: *egovernment*, kategorisasi, media sosial, *regular expression*

1. Pendahuluan

Dalam pemerintahan tingkat kota, terdapat Satuan Kerja Perangkat Daerah (SKPD) yang bertujuan untuk menjalankan kebijakan, program, dan kegiatan pembangunan berdasarkan tugas dan fungsinya[1]. Dasar hukum yang mengatur dalam pembentukan SKPD terdapat dalam pasal 120 UU no. 32 tahun 2004 tentang Pemerintahan Daerah. SKPD adalah pelaksana fungsi eksekutif yang harus berkoordinasi satu sama lain agar penyelenggaraan pemerintahan dapat berjalan dengan baik. SKPD sendiri terbagi menjadi sekretariat daerah, staf-staf ahli, dinas-dinas, badan-badan, inspektorat daerah, dan lembaga-lembaga lain yang bertanggung jawab langsung kepada kepala daerah. Dengan adanya SKPD, pemerintah kota dapat membagi program pembangunan berdasarkan tugas dan fungsinya masing-masing. Di lain sisi, pemerintah sedang mengencankan konsep *e-government* sebagai upaya pemanfaatan dan penggunaan teknologi informasi dan komunikasi untuk meningkatkan efisiensi dan efektivitas pemerintahan. Melalui *e-government*, pemerintah dapat memberikan berbagai jasa pelayanan kepada masyarakat, menyediakan akses informasi dan mengadakan proses pemerintahan yang bertanggung jawab dan transparan kepada masyarakat.

Penerapan konsep *e-government* tidak terlepas dari

kemajuan pengguna internet di Indonesia yang mengalami pengembangan yang pesat. Berdasarkan survei yang dilakukan oleh Internetworldstats yang dirilis pada Januari 2018, jumlah pengguna internet di Indonesia telah mencapai 132,7 juta dari jumlah penduduk Indonesia sebesar 256,2 juta jiwa[2]. Hal ini menandakan kenaikan sebesar 51,8% dibandingkan dengan jumlah pengguna internet di Indonesia pada tahun 2014 dengan jumlah 88 juta pengguna internet. Dari jumlah tersebut, 40% di antaranya merupakan pengguna aktif media sosial. Dengan jumlah sebesar 106 juta jiwa yang menggunakan media sosial, 85% di antaranya mengakses media sosial melalui perangkat seluler[3].

Melalui media sosial, masyarakat dapat melakukan interaksi kepada akun media sosial pemerintah kota dalam memberikan informasi atau kritik dan saran terkait kotanya. Berdasarkan penelitian yang dilakukan oleh Sultana Lubna Alam dan Aodah Diamah pada halaman akun Facebook Pariwisata Pemerintah Australia, masyarakat secara aktif melakukan interaksi dengan cara memberikan informasi terbaru di halaman Facebook Pariwisata Pemerintahan Australia[4]. Dengan melihat kesempatan ini, pemerintah kota dapat mengetahui informasi dari komentar masyarakat melalui media sosial terkait permasalahan yang sedang terjadi atau topik yang sedang ramai dibicarakan masyarakat. Dengan mengetahui informasi atau permasalahan yang

terjadi di suatu kota, pemerintah kota dapat melakukan aksi sesuai dengan SKPD terkait sesuai dengan tugas dan fungsinya masing-masing.

Melalui fenomena tersebut, diperlukanlah suatu aplikasi yang bertujuan untuk melihat komentar masyarakat di media sosial yang telah terkategori berdasarkan SKPD di masing-masing kota. Dengan aplikasi ini, pemerintah kota dapat mengetahui topik apa yang sedang ramai dibicarakan masyarakat atau masalah yang sedang terjadi. Proses kategorisasi disesuaikan dengan kata kunci yang terkait dengan SKPD di masing-masing kota agar pemerintah kota langsung memberikan arahan kepada SKPD terkait. Pada penelitian sebelumnya kami sudah melakukan kategorisasi media sosial pemda[5][6], namun belum secara khusus melakukan kategorisasi di tingkat SKPD. Furqon *et al* [5] menganalisis jenis posting Pemda di media sosial seperti video, gambar dilihat dari banyaknya *like* dan hanya media sosial Facebook. Izzati *et al* [6] menganalisis jenis interaksi Pemda dengan masyarakat via media sosial. Analisis interaksi Pemda di media sosial juga dilakukan oleh Suryadharma [7] dengan menggunakan metode wawancara. Pentingnya penggunaan media sosial pada humas Pemda juga dikaji oleh [8][9].

Kontribusi dari penelitian ini adalah:

- a. Pengembangan aplikasi yang memudahkan pemerintah daerah untuk melakukan pengelompokan komentar sosial media
- b. Pemberian rekomendasi kata kunci untuk pengelompokan komentar sosial media.

2. Metode

Tahapan metodologi menjelaskan tahapan yang dilakukan dalam penelitian dari perancangan aplikasi hingga pembuatan visualisasi untuk aplikasi.

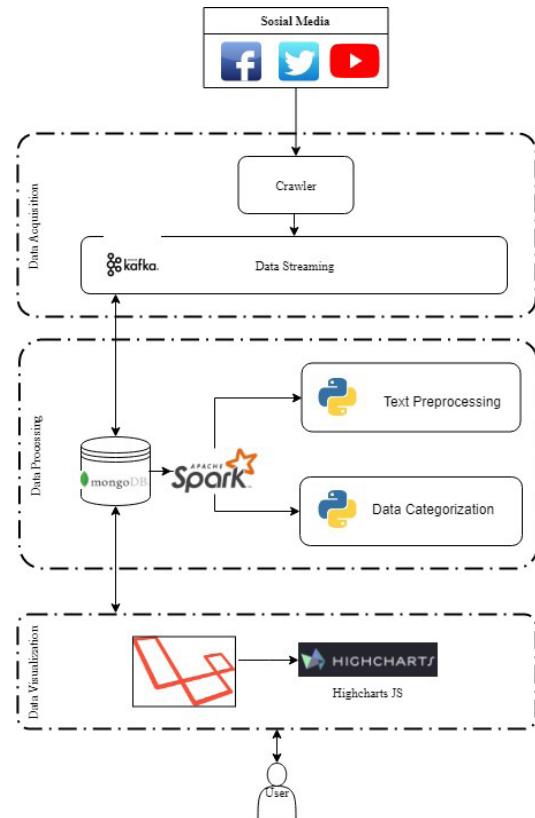
a. Perancangan Aplikasi

Tahap pertama yang dilakukan yaitu perancangan aplikasi. Perancangan aplikasi dilakukan dengan mendesain arsitektur sistem untuk aplikasi.

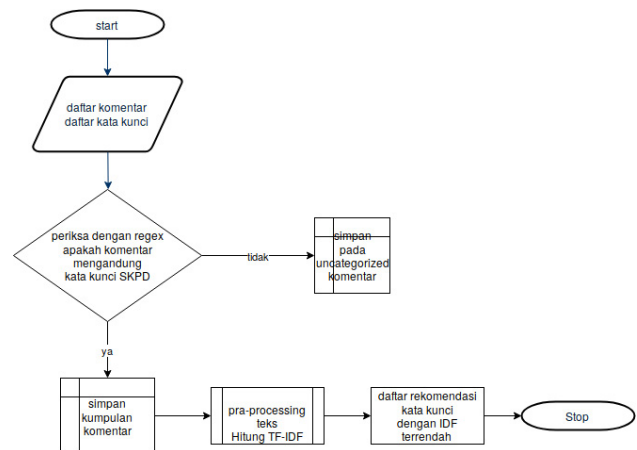
Arsitektur sistem (Gambar 1) dimulai dengan *crawler* yang melakukan pengambilan data dengan menggunakan API di masing-masing media sosial. Dalam melakukan pengambilan data, digunakan Apache Kafka sebagai *data pipeline* dalam *streaming data*. Data kemudian dilakukan *preprocessing* dan pengkategorian menggunakan Apache Spark. Data yang sudah dikategorikan dimasukkan ke dalam MongoDB dan ditampilkan kepada *user* melalui *website* dengan *framework* Laravel dan *library* highchart.

b. Akuisisi Data

Tahap ini dilakukan dengan pengambilan data pada akun resmi media sosial pemerintah kota di Indonesia dan akun *influencer* di masing-masing kota. Adapun media sosial yang digunakan yaitu Facebook, Twitter, dan Youtube. Proses pengambilan data dilakukan dengan teknik *crawling*. Proses *crawling* memanfaatkan *Application Programming Interface* (API) dari masing-masing media sosial. Adapun entitas yang diambil dalam proses *crawling* yaitu komentar yang masuk pada akun resmi media sosial pemerintah kota. Dalam melakukan akuisisi data digunakan aplikasi agregasi Apache Kafka untuk menyatukan data yang bersifat *streaming* dan digunakan *database* MongoDB sebagai tempat penyimpanan data.



Gambar 1 Arsitektur sistem



Gambar 2 Proses kategorisasi dan rekomendasi kata

c. Text Preprocessing

Tahapan *preprocessing data* dilakukan untuk mengubah setiap kata pada komentar di media sosial yang masuk menjadi sebuah *vector*. Dalam melakukan *text preprocessing*, terdapat tiga tahapan yang dilakukan yaitu *stemming*, *stop word removal*, dan *n-grams*. *N-grams* merupakan sebuah teknik untuk memecah kalimat menjadi potongan N kata. Dengan melakukan pemecahan kalimat menjadi kata, N grams dapat digunakan untuk memprediksi kata selanjutnya dalam sebuah kalimat. Hal ini membuat prediksi kata tidak hanya berupa satu kata namun lebih dari satu kata. N-grams terbagi menjadi empat jenis yaitu *Unigrams* dengan N=1, *Bigrams* untuk N=2, *Trigrams* dengan N=3, dan N grams untuk N>3. Dalam memberikan prediksi kata, N-grams bekerja

menggunakan Markov Assumption[10][11][12]. Hal ini bertujuan agar prediksi kata yang digunakan tidak melihat kata sebelumnya yang terlalu jauh. Perhitungan prediksi kata selanjutnya dihitung dengan menggunakan rumus (1):

$$P(W_n | W_{n-1}) = \text{Count}(W_n - 1, W_n) / \text{Count}(W_n - 1) \quad (1)$$

W_{n-1} = kata sebelumnya

W_n = prediksi kata

Dengan melihat kemungkinan tertinggi dari semua kata, prediksi pemilihan kata dapat dilakukan. Hal ini dinamakan dengan *Maximum Likelihood Estimation*.

Salah satu permasalahan yang terdapat pada *N-grams* yaitu bergantungnya struktur kalimat dalam pemilihan kata. Sebagai contoh, apabila kalimat yang diberikan memiliki keterangan tempat setelah subjek baru kemudian prediksi kata maka prediksi kata akan melihat kata terakhir dalam keterangan. Hal ini akan mengakibatkan bias karena prediksi kata akan lebih tepat jika melihat pada subjek yang diberikan.

d. Pengkategorian Komentar

Tahap ini bertujuan untuk melakukan kategorisasi dari komentar yang masuk. Dengan melakukan *pattern matching*, komentar yang memiliki kata kunci yang ditemukan dari *text preprocessing* akan disesuaikan dengan unit-unit dinas di pemerintah kota. Dalam melakukan *pattern matching*, terdapat langkah yang dilakukan yaitu *Regular Expressions*. *Regular expression* atau *Regex* merupakan *pattern* yang mengspesifikasikan kumpulan karakter yang ada dalam sebuah kalimat. *Regex* merupakan kalimat yang merupakan kombinasi dari *normal characters* dan *metacharacters* atau *metasequences*. *Metacharacters* atau *metasequences* merupakan sebuah karakter yang mewakili makna seperti kuantitas, lokasi, atau tipe karakter untuk melakukan *regex*[13]. Sekumpulan kata yang didapatkan dari *regular expression* dinamakan dengan *pattern matching*. Proses kategorisasi dan rekomendasi dapat dilihat pada Gambar 2.

e. Rekomendasi Kata

Pada tahap ini dilakukan perhitungan IDF untuk komentar yang telah memiliki label dinas. Metode TF-IDF merupakan suatu cara untuk memberikan bobot hubungan suatu kata (*term*) terhadap dokumen. Metode ini menggabungkan dua konsep perhitungan bobot yaitu frekuensi kemunculan kata dalam dokumen dan *inverse* frekuensi dokumen yang mengandung kata tersebut[14]. TF adalah algoritma pembobotan heuristik yang menentukan bobot dokumen berdasarkan kemunculan *term* (istilah). Semakin sering sebuah *term* (istilah) muncul maka semakin tinggi bobot dokumen untuk istilah tersebut[15]. IDF merupakan banyaknya istilah tertentu dalam keseluruhan dokumen yang dapat dihitung dengan persamaan (2):

$$\text{Idf } j = \log(n/n_j) \quad (2)$$

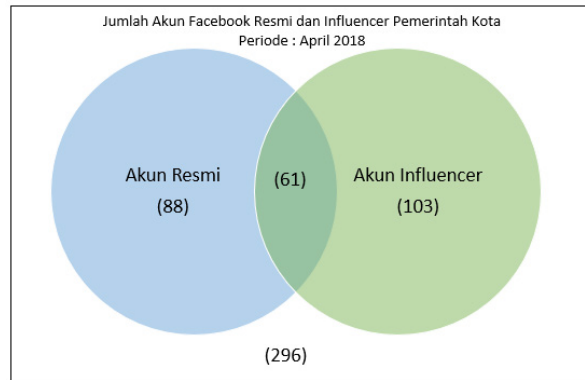
n = jumlah seluruh dokumen

n_j = jumlah dokumen yang mempunyai istilah j

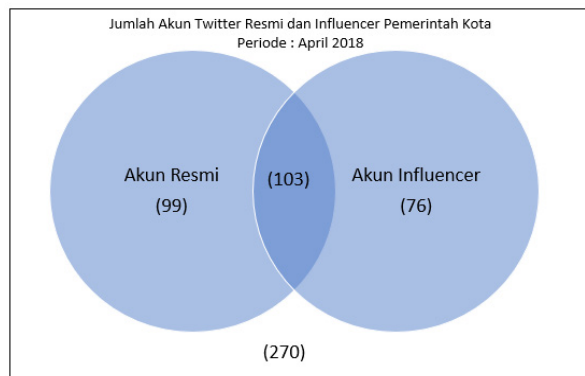
Kata yang memiliki nilai IDF paling rendah akan diambil dan digunakan sebagai rekomendasi kata kunci kepada dinas terkait.

Tabel 1. Jumlah akun media sosial dan *influencer* pemerintah kota April 2018

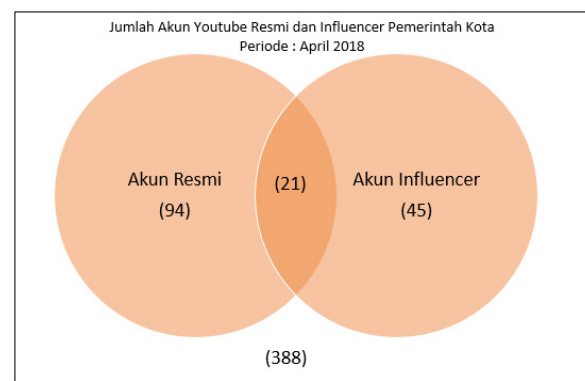
Media sosial	Jumlah Akun Resmi	Jumlah Akun <i>Influencer</i>
Facebook	260	164
Twitter	297	179
Youtube	223	40



Gambar 3. Jumlah akun facebook resmi dan *influencer*



Gambar 4. Jumlah akun twitter resmi dan *influencer*



Gambar 5. Jumlah akun youtube resmi dan *influencer*

f. Perancangan dan Pembuatan Visualisasi Aplikasi

Pada tahap ini dilakukan tahapan perancangan dan pembuatan visualisasi dengan *website*. Tahap ini bertujuan untuk menghasilkan rancangan dan hasil akhir aplikasi berupa *website* sebagai bentuk visual dan memudahkan pemerintah kota untuk mengetahui komentar yang masuk sesuai dengan unit dinas masing-masing. Tahapan ini terbagi menjadi empat bagian yaitu perencanaan, desain, *coding*, dan pengujian.

Tabel 2. Kata kunci untuk dinas di pemerintah Kota Depok

Nama Dinas	Kata Kunci
Badan Keuangan Daerah	Anggaran
Dinas Kearsipan dan Perpustakaan	Buku
Dinas Kependudukan dan Pencatatan Sipil	Penduduk
Dinas Kesehatan	Puskesmas
Dinas Ketahanan Pangan, Pertanian dan Perikanan	Pangan
Dinas Komunikasi dan Informatika	Teknologi
Dinas Koperasi dan Usaha Mikro	Usaha
Dinas Lingkungan Hidup dan Kebersihan	Bersih
Dinas Pekerjaan Umum dan Penataan Ruang	Infrastruktur
Dinas Pemadam Kebakaran dan Penyelamatan	Kebakaran
Dinas Pemuda, Olahraga, Kebudayaan, dan Pariwisata	Wisata
Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu	Portal
Dinas Pendidikan	Sekolah
Dinas Perdagangan dan Perindustrian	Perdagangan
Dinas Perhubungan	Jalan
Dinas Perlindungan Anak, Pemberdayaan Masyarakat dan Keluarga	Anak
Dinas Perumahan dan Permukiman	Perumahan
Dinas Sosial	Masyarakat
Dinas Tenaga Kerja	Pekerjaan

Tabel 3. Persentase jumlah komentar yang masuk

Label	Persen
Dinas	10,35
Duplicate	1,11
Uncategorized	88,54

Tabel 4. Jumlah komentar untuk setiap media sosial Kota Depok

Akun Kota Depok	Label Dinas	Label <i>Duplicate</i>	Label <i>Uncategorized</i>
Facebook Resmi	224	34	744
Facebook <i>Influencer</i>	209	13	3003
Twitter Resmi	7	0	2
Twitter <i>Influencer</i>	0	0	0
Youtube Resmi	8	1	80
Youtube <i>Influencer</i>	0	0	0

3. HASIL

a. Pengambilan Data

Dalam melakukan pengambilan data, objek yang diambil yaitu komentar pada media sosial Facebook, Twitter, dan Youtube pada 413 kabupaten, 97 kota, dan 32 provinsi. Kemudian dilanjutkan dengan verifikasi akun resmi pemerintah dan akun *influencer* di masing-masing pemerintah kota. Berikut merupakan jumlah akun resmi dan akun *influencer* pada periode April 2018

Pada media sosial Facebook, terdapat 88 provinsi/kota yang hanya memiliki akun resmi, 103 provinsi/kota yang hanya memiliki akun *influencer*, 61 provinsi/kota yang keduanya memiliki akun resmi dan *influencer*, dan 296

provinsi/kota yang tidak memiliki akun resmi dan *influencer* (Gambar 3).

Gambar 4 memperlihatkan statistik media sosial Twitter dimana terdapat 99 provinsi/kota yang hanya memiliki akun resmi, 76 provinsi/kota yang hanya memiliki akun *influencer*, 103 provinsi/kota yang keduanya memiliki akun resmi dan *influencer*, dan 270 provinsi/kota yang tidak memiliki akun resmi dan *influencer*.

Gambar 5 menggambarkan kondisi media sosial Youtube, di mana terdapat 94 provinsi/kota yang hanya memiliki akun resmi, 45 provinsi/kota yang hanya memiliki akun *influencer*, 21 provinsi/kota yang keduanya memiliki akun resmi dan *influencer*, dan 388 provinsi/kota yang tidak memiliki akun resmi dan *influencer*.

Tabel 5. Jumlah komentar untuk setiap dinas di Kota Depok

Nama Dinas	FR	FI	TR	YR
Badan Keuangan Daerah	0	0	0	0
Dinas Kearsipan dan Perpustakaan	0	4	1	0
Dinas Kependudukan dan Pencatatan Sipil	0	0	0	0
Dinas Kesehatan	8	1	1	0
Dinas Ketahanan Pangan, Pertanian dan Perikanan	1	0	0	0
Dinas Komunikasi dan Informatika	1	0	1	0
Dinas Koperasi dan Usaha Mikro	4	8	1	0
Dinas Lingkungan Hidup dan Kebersihan	7	9	0	1
Dinas Pekerjaan Umum dan Penataan Ruang	1	0	0	0
Dinas Pemadam Kebakaran dan Penyelamatan	0	0	0	0
Dinas Pemuda, Olahraga, Kebudayaan, dan Pariwisata	2	4	1	0
Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu	0	0	0	0
Dinas Pendidikan	1	8	0	2
Dinas Perdagangan dan Perindustrian	0	0	0	0
Dinas Perhubungan	143	79	0	5
Dinas Perlindungan Anak, Pemberdayaan Masyarakat dan Keluarga	44	81	1	0
Dinas Perumahan dan Permukiman	0	0	0	0
Dinas Sosial	12	15	1	0
Dinas Tenaga Kerja	0	0	0	0

FR = Facebook resmi, FI = Facebook *Influencer*, TR = Twitter Resmi, YR = Youtube Resmi

Tabel 6. Kata kunci rekomendasi untuk dinas di Kota Depok

Nama Dinas	Kata Kunci Rekomendasi	Relevan/Tidak Relevan	Nilai IDF
D i n a s Perhubungan	<i>Traffic light</i>	Relevan	1.0
D i n a s Perlindungan Anak, Pemberdayaan Masyarakat dan Keluarga	Baca	Tidak Relevan	1.0
Dinas Sosial	Nik	Tidak Relevan	1.728

b. Hasil Pengelompokan

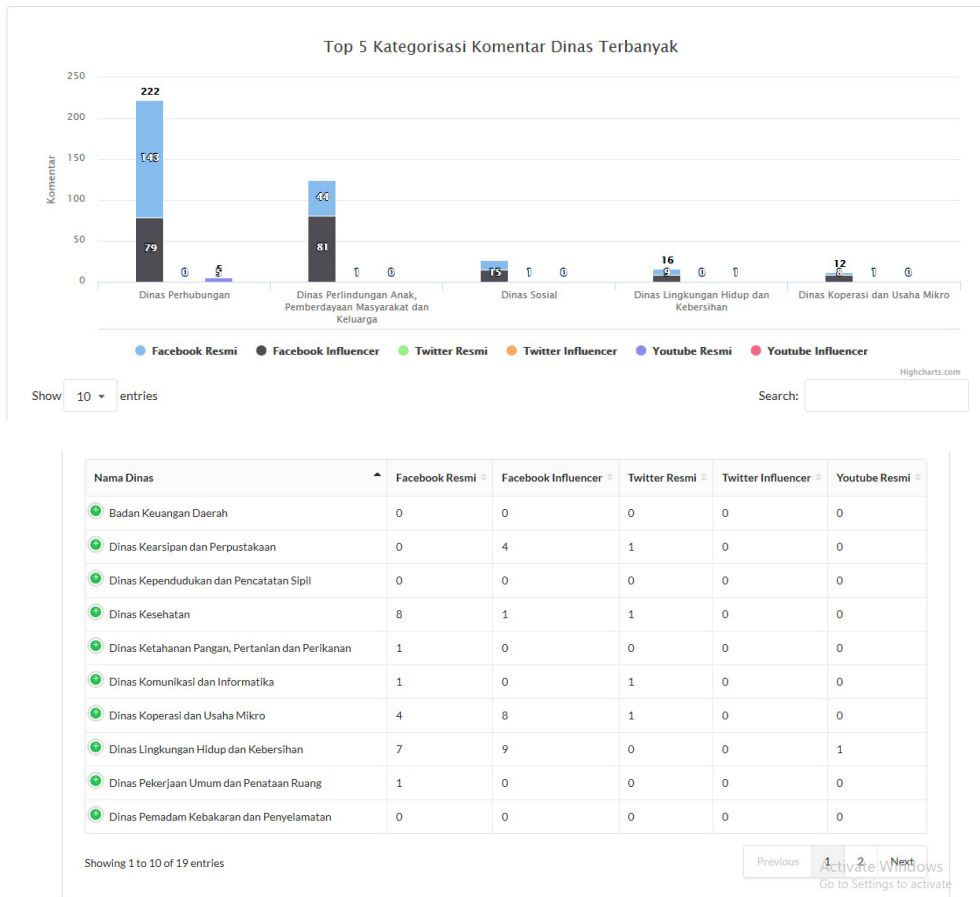
Pengujian hasil kategorisasi dilakukan pada akun media sosial Kota Depok. Pengumpulan komentar kota Depok untuk media sosial Facebook dan Youtube dilakukan pada periode 1 Januari 2016 hingga 29 Mei 2018 sementara untuk media sosial Twitter dilakukan pada periode 26 Juni 2018. Total komentar yang didapatkan berjumlah 4.609 komentar dengan rincian Facebook resmi berjumlah 1.038, Facebook *Influencer* berjumlah 3.473, Twitter resmi berjumlah 9 dan Youtube resmi berjumlah 89. Setelah didapatkan komentar, langkah selanjutnya dilakukan pemilihan kata kunci untuk masing-masing dinas pada pemerintah kota Depok. Tabel 2 menjelaskan kata kunci untuk dinas-dinas di kota Depok.

Komentar yang sudah didapatkan dilakukan *preprocessing* data. Dengan melakukan *preprocessing*, jumlah komentar berkurang menjadi 4.325 komentar. Selanjutnya komentar dilakukan kategorisasi berdasarkan kata kunci yang telah dipilih sebelumnya. Apabila kata kunci ditemukan maka akan diberi label sesuai dengan nama

dinas, apabila ditemukan lebih dari satu kata kunci akan diberi label "*duplicate*", dan apabila tidak ditemukan maka akan diberi label "*uncategorized*".

Berdasarkan Tabel 3, diketahui bahwa dari 4.325 komentar, 10,35% komentar berhasil diberi label dinas, 1,11% komentar diberi label "*duplicate*" dan 88,54% komentar diberi label "*uncategorized*". Berikut jumlah tabel untuk masing-masing komentar di setiap media sosial.

Berdasarkan tabel 3, diketahui bahwa jumlah *uncategorized* pada akun sosial media Facebook sangat signifikan. Hal ini disebabkan karena kata kunci yang dimasukkan ke dalam dinas belum bisa merepresentasikan data komentar yang masuk. Hal tersebut berlaku pada akun media sosial Youtube di mana jumlah komentar yang terlabel *uncategorized* masih dominan dalam pelabelan. Pada akun sosial media Twitter, komentar yang masuk sudah dominan dengan label dinas karena data komentar memiliki kata kunci yang sesuai dengan dinas. Berikut merupakan hasil pelabelan untuk komentar di setiap dinas.



Gambar 6. Tampilan hasil kategorisasi aplikasi

Berdasarkan tabel 5 diketahui bahwa Dinas Perhubungan memiliki jumlah komentar terbanyak yang disusun dengan Dinas Perlindungan Anak, Pemberdayaan Masyarakat, dan Keluarga. Hal ini menandakan bahwa *keyword* yang digunakan mewakili dari dinas terkait. Hasil kategorisasi kemudian ditampilkan melalui visualisasi pada aplikasi.

c. Rekomendasi Kata Kunci

Setelah didapatkan komentar yang telah terkategori maka langkah selanjutnya dilakukan pengujian terhadap rekomendasi kata kunci. Perhitungan kata kunci dilakukan dengan menghitung nilai IDF untuk setiap kata dalam komentar antar dinas. Perhitungan dilakukan terhadap dinas yang memiliki jumlah komentar di atas rata-rata komentar yang masuk di pemerintah kota tersebut. Sehingga tidak semua dinas memiliki kata kunci. Berikut merupakan kata kunci untuk masing-masing dinas. Pemilihan kata kunci dipilih berdasarkan kata kunci dengan nilai IDF terendah.

Berdasarkan table 6 diketahui bahwa terdapat tiga dinas yang memiliki rekomendasi kata dari komentar yang masuk. Adapun pada Dinas Perhubungan rekomendasi kata yang diberikan adalah “*traffic light*” dengan nilai IDF 1.0. Berdasarkan komentar yang didapatkan, kata “*traffic light*” relevan dengan SKPD yang ada karena komentar berupa saran masyarakat terhadap kondisi *traffic light* di kota Depok yang harus diubah. Pada Dinas Perlindungan Anak, Pemberdayaan Masyarakat, dan Keluarga rekomendasi kata yang diberikan yaitu “*baca*” dengan

nilai IDF 1.0. Kata “*baca*” dalam komentar tidak relevan karena komentar menyatakan keluhan kepada pemerintah kota Depok oleh masyarakat yang membaca *billboard* di kota Depok mengatakan kalau kota Depok adalah kota ramah anak, tetapi kenyataannya tidak. Terakhir pada Dinas Sosial memiliki rekomendasi kata “*nik*” dengan nilai IDF 3,6741. Kata kunci “*nik*” tidak relevan dengan SKPD karena kata ‘*nik*’ dalam komentar menyatakan keluhan masyarakat terhadap dibutuhkannya *nik* dalam proses registrasi sim *bandphone*.

Ketidakrelevanan kata kunci dipengaruhi oleh kata kunci dalam melakukan kategorisasi. Hal ini menyebabkan jumlah komentar yang terkategori menjadi sedikit dan mempengaruhi nilai IDF.

Gambar 6 menampilkan visualisasi yang terdapat pada *website*, di mana terdapat grafik batang kondisi setiap media sosial terhadap SKPD.

4. Kesimpulan

Penelitian ini menyajikan arsitektur untuk melakukan pengambilan data dari media sosial Facebook, Twitter, dan Youtube. Arsitektur ini mampu menjalankan proses pengumpulan data komentar pada akun resmi dan *influencer* di masing-masing media sosial. Pengujian kategorisasi komentar yang menggunakan 4.325 komentar pada akun media sosial Kota Depok yang telah dilakukan *preprocessing* menghasilkan 10,35% komentar dikategorikan dinas, 1,11% komentar dikategorikan *duplicate*, dan 88,54% komentar dikategorikan *uncategorized*

dengan kata kunci asumsi yang digunakan. Apabila kata kunci merepresentasikan komentar pada dinas, maka jumlah komentar yang terkategori dinas akan meningkat. Pengujian rekomendasi kata kunci yang menggunakan 19 dinas menghasilkan 1 kata kunci untuk setiap 3 dinas. Ketiga dinas tersebut memiliki jumlah komentar di atas rata-rata komentar yang telah terkategori. Adapun kata kunci yang dihasilkan tidak relevan untuk ketiga dinas. Hal ini disebabkan karena jumlah komentar di setiap dinas sangat sedikit dan nilai IDF untuk masing-masing kata bernilai sama.

5. Daftar Pustaka

- [1] Republik Indonesia, "UNDANG-UNDANG REPUBLIK INDONESIA NOMOR 32 TAHUN 2004 TENTANG PEMERINTAHAN DAERAH," no. 1, pp. 1–7, 2004.
- [2] D. I. N. Indonesia, "Digital in India a Snapshot of the Country ' S Key Digital Statistical Indicators," *Hootsuite*, pp. 199–213, 2018.
- [3] TETRA PAK, "the Tetra Pak Index 2017," no. 10, 2017.
- [4] S.L. Alam and A. Diamah. "Understanding user participation in Australian Government tourism Facebook page." ACIS 2012: Location, location, location: *Proceedings of the 23rd Australasian Conference on Information Systems 2012*. ACIS, 2012.
- [5] M. A. Furqon, et al. "Analisis Jenis Posting Media Sosial Pemerintah Daerah di Indonesia Berdasarkan Like dan Analisis Sentimental Masyarakat." *Jurnal Sosioteknologi* 17.2 (2018): 177-190, 2018
- [6] A. N. Izzati, et al. "Kategorisasi Jenis Interaksi antara Pemerintah dan Masyarakat dan Popularitas Media Sosial Pemerintah Daerah." *Jurnal Sistem Informasi* 14.1 (2018): 1-8, 2018
- [7] B. Suryadharma, and T. D. Susanto. "Faktor Penerimaan Media Sosial Instansi Pemerintah di Indonesia." *INTEGER: Journal of Information Technology* 2.2 2017
- [8] N. Kurniasih, S. Sos, and M. Hum. "Penggunaan Media Sosial Bagi Humas Di Lembaga Pemerintah." *Forum Kebumasan Kota Tangerang*, 2013.
- [9] D. Purworini. "Model informasi publik di era media sosial: Kajian grounded teori di pemda sukoharjo." *Komuniti: Jurnal Komunikasi dan Teknologi Informasi* 6.1 (2017): 3-14, 2017
- [10] C. Ramisch. "N-gram models for language detection." *M2R Informatique-Double dipl'ome ENSIMAG-UJF/UFRIMA* (2008).
- [11] E. Manjavacas, et al. "Synthetic literature: Writing science fiction in a co-creative process." *Proceedings of the Workshop on Computational Creativity in Natural Language Generation (CC-NLG 2017)*, 2017
- [12] D. Jurafsky and J. H. Martin, "N-Grams" *Speech Lang. Process.*, pp. 2–7, 2014.
- [13] D. R. Tobergte and S. Curtis, "Regular Expression Pocket Reference," *Journal of Chemical Information and Modeling*, vol. 53, no. 9. pp. 1689–1699, 2013.
- [14] O. Karmayasa and I. B. Mahendra, "Implementasi VSM dan Notasi Metode TF-IDF pada Sistem Temu Kembali Informasi," *J. Progr. Stud. Tek. Inform.*, p. 9, 2010.
- [15] M. Fitri. "Perancangan Sistem Temu Balik Informasi dengan Metode Pembobotan Kombinasi TF-IDF untuk Pencarian Dokumen Berbahasa Indonesia." *Jurnal Sistem dan Teknologi Informasi (JustIN)* 1.1 (2012): 80-85. 2013.