

**ANALISIS DAN PENGELOMPOKAN LAPORAN PENGADUAN MASYARAKAT
PADA SISTEM LAPOR.GO.ID DENGAN ALGORITMA K-MEANS
PERIODE 2013-2015**

Umi Oktovia Lia Kartini¹⁾ dan Sigit Birowo²⁾

1) Alumni Program Studi Sistem Informasi

2) Staf Pengajar Program Studi Sistem Informasi

Institut Bisnis dan Informatika Kwik Kian Gie

Jl. Yos Sudarso Kav.87 Sunter Jakarta Utara 14350

<http://www.kwikkiangie.ac.id>

e-mail: ²sigit.birowo@kwikkiangie.ac.id

ABSTRACT

Indonesian people has a lot of access to share information about events that occur in everyday life that so the reports data that can be obtained are very numerous. In this study the authors will carry out the implementation and grouping with applying text mining method on report data from LAPOR.go.id to find and classify complaint topics that is most frequently reported with the k-means algorithm. Within the applications, text mining is an advanced technique to analyze and process the textual data that is part of data mining. The application of text mining is used to process the report data prior to the clustering and cluster modeling based on k-means algorithm.

Keywords : Reports, Complaints, People, Report, Trending, Grouping, Cluster, Text Mining, Data Mining, K-Means

1. PENDAHULUAN

Teknologi informasi merupakan suatu hal yang tidak dapat dipisahkan dari perkembangan dunia pada era globalisasi. Semakin banyaknya penggunaan teknologi informasi pada kehidupan manusia membuat semakin cepatnya perkembangan yang dirasakan di berbagai bidang seperti dalam bidang ekonomi, sosial, politik, budaya, pendidikan, dan lain-lain. Perkembangan yang pesat ini menimbulkan berbagai macam bentuk teknologi yang dirancang untuk membantu manusia menghadapi segala masalah yang muncul dalam kehidupan sehari-hari yang tidak mampu diselesaikan tanpa bantuan sistem komputasi pada komputer. Masalah-masalah tersebut muncul karena semakin banyaknya penggunaan teknologi yang digunakan oleh manusia, sehingga banyaknya problem yang berkaitan dengan dunia digital dan dunia maya semakin bertambah jumlahnya.

Manusia secara tidak sadar telah hidup berdampingan dengan teknologi sehingga tidak mampu dipisahkan dalam berbagai bidang, bahkan banyak orang yang tergolong

berlebihan dan menggunakan teknologi sehingga menjadi sebuah bagian dari gaya hidup yang dilakukan tiap hari. Kemudahan akses informasi yang didapat dan dapat ditelusuri pada era digital dalam internet telah memberikan banyak sekali manfaat yang tidak terhitung seperti kemudahan memperoleh dan berbagi informasi yang dibutuhkan atau dimiliki dari sumber-sumber yang tersedia, memperlancar hubungan atau komunikasi dari satu tempat ke tempat lain dengan mudah, meningkatkan produktivitas kerja atau bisnis, dan lain-lain.

Perkembangan teknologi seperti desktop pc, netbook, smartphone dan tablet membuat semakin mudahnya mendapatkan akses untuk mencari atau membagikan informasi dimana saja dan kapan saja. Masyarakat Indonesia bisa dibidang cukup memiliki banyak akses untuk membagikan informasi tentang kejadian yang terjadi di kehidupan sehari-hari sehingga data laporan yang dapat diperoleh dari berbagai kejadian tersebut sangat banyak jumlahnya. Berbagai macam laporan masyarakat yang didapat diseluruh Indonesia dapat terdiri dari berbagai masalah seperti ekonomi, sosial, kesehatan, infrastruktur, dll dapat dilaporkan

setiap saat sehingga data laporan masyarakat yang terkumpul juga semakin besar dan perlu penanganan secara digital oleh pihak pemerintah agar seluruh keluhan masyarakat dapat tersampaikan.

Oleh karena itu pada penelitian ini penulis akan membahas tentang sebuah aplikasi mobile yang bernama LAPOR! (Layanan Aspirasi dan Pengaduan Online Rakyat). Aplikasi ini adalah sebuah sarana pengaduan berbasis media sosial yang terhubung dengan media sosial seperti facebook dan twitter untuk melaporkan berbagai keluhan, kejadian, masalah, atau aspirasi kepada instansi pemerintah secara nasional di seluruh Indonesia. Melalui aplikasi ini masyarakat dapat melakukan pengiriman laporan atau aspirasi tentang suatu topik melalui sms, website (<https://www.lapor.go.id>), atau pun melalui aplikasi mobile yang tersedia pada platform android. LAPOR! dikembangkan oleh Kantor Staf Presiden dalam rangka meningkatkan partisipasi masyarakat untuk pengawasan program dan kinerja pemerintah dalam penyelenggaraan pembangunan dan pelayanan publik. LAPOR! telah digunakan oleh lebih dari 290.000 pengguna sejak April, 2015 dan menerima rata-rata lebih dari 800 laporan masyarakat per harinya. LAPOR! menjadi cikal-bakal sistem aspirasi dan pengaduan masyarakat yang terpadu secara nasional.

Data laporan yang diambil dari aplikasi LAPOR! tersebut akan digunakan oleh penulis untuk dianalisa dan dilakukan pengelompokkan untuk mencari tahu tren topik laporan yang paling banyak dilaporkan oleh masyarakat sehingga mempermudah pencarian topik yang diinginkan dan membantu memfokuskan isu yang paling banyak dibahas dengan menggunakan metode text mining dengan algoritma clustering k-means, untuk menemukan trend dan pola topik berdasarkan data laporan yang diperoleh pada <http://data.go.id>. Website <http://data.go.id> merupakan sebuah portal data publik yang dapat diakses oleh publik secara gratis untuk mendapatkan informasi secara bebas tentang berbagai instansi pemerintah dan berbagai bidang di Indonesia.

Dengan banyaknya jumlah data yang didapatkan maka diperlukan sebuah metode

pengolahan data yang mampu mengolah data dalam jumlah besar untuk menemukan pola yang tersembunyi didalamnya sehingga informasi yang terdapat dalam pola tersebut dapat diambil dan diolah ke dalam bentuk yang berguna untuk pengambilan keputusan. Yang disebut dengan data mining. Pada data mining sendiri ada sebuah metode lanjutan yang digunakan untuk mengolah data yang berupa dokumen atau teks secara khusus yang disebut dengan text mining. Pada dasarnya pengolahan data dalam text mining cukup mirip dengan data mining tetapi karena text mining merupakan metode lanjutan yang terfokus pada pengolahan data tekstual maka modul-modul yang digunakan untuk proses clustering cukup berbeda. Proses clustering akan dilakukan pada aplikasi rapidminer dimana pada aplikasi tersebut akan dilakukan proses dengan menggunakan modul text mining untuk mengolah data laporan sebelum dilakukan proses pengelompokkan dengan algoritma k-means. Algoritma k-means digunakan untuk memperoleh model kluster untuk pengelompokkan data laporan yang akan dilakukan sehingga dapat dibagi menjadi kelompok-kelompok yang sesuai dengan tren topik pencarian yang paling banyak pada tiap kelompok kluster yang akan dibuat.

Karena data laporan masih berbentuk data mentah yang tidak terstruktur dan berasal dari postingan sehingga untuk melakukan pengolahan lebih lanjut harus dilakukan proses struktursisasi teks. Sedangkan dalam pengelompokkan judul laporan masih belum diterapkan pengelompokkan judul laporan berdasarkan topik yang paling sering dilaporkan sehingga sulit untuk melakukan sorting menurut trend judul. Ini menyebabkan sulitnya mencari topik yang penting dibahas oleh masyarakat sehingga laporan yang dihasilkan kurang efektif. Karena terbatasnya kategori maka akan dilakukan juga analisa dari laporan tersebut agar pengelompokkan komentar dapat dilakukan secara lebih efektif mengingat belum dilakukan pengelompokkan laporan berdasarkan isi atau komentar laporan sehingga tidak mampu melakukan pengelompokkan berdasarkan trend pendapat pelapor.

Jika model pengelompokkan sudah ditemukan maka kategorisasi yang dilakukan dapat dikembangkan sebab belum ada

kategorisasi judul dan isi laporan berdasarkan kelompok tertentu atau klaster dari trend judul yang diinginkan. Maka penelitian yang akan dilakukan ini akan difokuskan kepada pengolahan data laporan menggunakan metode text mining dengan algoritma k-means untuk menemukan pola pengelompokan yang sesuai dalam pengolahan data tersebut.

2. LANDASAN TEORI

2.1 Data

Data adalah sebuah fakta atau figur yang terekam [6] atau adalah aliran fakta mentah yang merepresentasikan kejadian yang terjadi pada suatu organisasi atau lingkungan fisik sebelum diorganisir dan disusun menjadi bentuk yang dapat dimengerti dan digunakan oleh seseorang [7]. Data juga bisa disebut sebagai aliran fakta mentah [8].

2.2 Informasi

Informasi dapat didefinisikan sebagai data yang direpresentasikan dalam sebuah bentuk yang bermakna [6] atau data yang telah dibentuk menjadi bentuk yang berarti dan berguna untuk manusia [7]. Pendapat lain mengatakan informasi adalah kumpulan fakta terorganisir dan terolah sehingga mereka memiliki nilai tambahan di luar nilai fakta individu [8].

2.3 Sistem

Sistem merupakan sebuah kumpulan set dari prosedur bisnis (atau komponen) yang digunakan dalam satu unit bisnis, yang bekerja untuk mencapai suatu tujuan [12] atau seperangkat komponen yang saling terkait, dengan batas yang jelas, yang saling bekerja sama untuk mencapai seperangkat tujuan dengan menerima input dan menghasilkan output dalam proses transformasi yang terorganisir [9].

2.4 Data Mining

Menurut [9] *data mining* adalah penggunaan utama dari *database data warehouse* dan data statis yang dikandungnya. Dalam *data mining*, data dalam *data warehouse* dianalisis untuk mengungkapkan pola tersembunyi dan tren dalam kegiatan bisnis sejarah. *Data mining* juga memiliki definisi alat analisis-informasi yang melibatkan penemuan otomatis pola dan

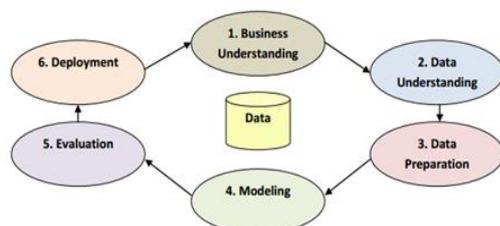
hubungan dalam *data warehouse* [8] atau sebuah aplikasi teknik statistik untuk menemukan pola dan hubungan diantara data untuk klasifikasi dan prediksi [6].

2.5 Tahap-tahap Data Mining

Tahapan dalam *data mining* terbagi dalam beberapa langkah yang disebut CRISP-DM (CRoss-Industry Standard Process for Data Mining) [1] yaitu antara lain adalah :

1. Business Understanding/Organizational Understanding (Pemahaman bisnis/organisasi) : Tahap pemahaman sistem yang berjalan dan kebutuhan apa yang dibutuhkan dalam menyelesaikan masalah yang timbul didalamnya.
2. Data Understanding (Pemahaman data) : Tahap pemahaman dan pengumpulan data yang dibutuhkan untuk sebelum dilakukan persiapan untuk analisa. Pada tahap ini data yang dikumpulkan harus merupakan data yang tepat digunakan untuk proses penelitian dan mewakili masalah yang akan dipecahkan serta sesuai dengan kebutuhan dan kepentingan.
3. Data Preparation (Persiapan data) : Tahap persiapan dan seleksi data yang telah dikumpulkan dan diubah menjadi bentuk yang dapat diolah dalam model yang ditentukan selanjutnya.
4. Modeling (Pemodelan) : Proses analisa dan pemodelan data yang telah disiapkan dimana dalam ini dilakukan penerapan atau penghitungan berdasarkan algoritma atau metode yang ditentukan untuk mendapatkan hasil yang diinginkan sesuai dengan kebutuhan pengguna dan melakukan representasi pemecahan masalah.
5. Evaluation (Evaluasi) : Melakukan analisa dan evaluasi dari hasil model yang telah dibuat apakah sudah sesuai standar dan telah memecahkan masalah atau memenuhi kebutuhan dari pengguna.
6. Deployment (Penerapan) : Tahap penerapan hasil dari model yang telah

dievaluasi dan dianalisa untuk kemudian dijadikan bentuk yang dapat diolah kembali.



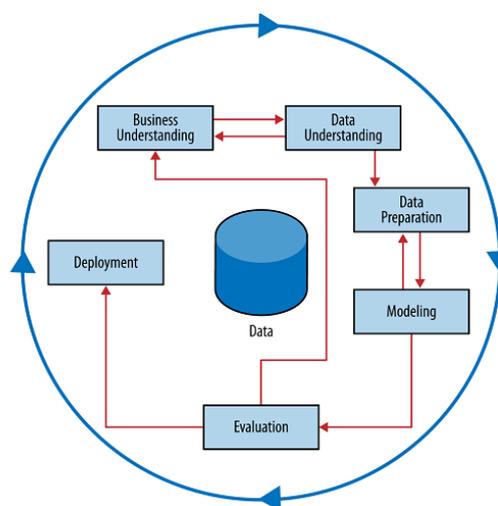
Gambar 1. Tahapan data mining (CRISP-DM)

Menurut [10] tahapan dalam data mining dalam CRISP-DM dibagi dalam 6 proses besar yaitu :

1. *Business Understanding* : Pada tahap pertama ini tim desain harus berpikir hati-hati tentang scenario penggunaan (*use case*). Awalnya, sangat penting untuk memahami masalah yang akan dipecahkan. Hal ini mungkin tampak jelas, tetapi proyek bisnis jarang datang sebagai sesuatu yang bersifat “pradikemas” dan mudah untuk dipecahkan atau tidak ambigu.
2. *Data Understanding* : Dalah tahap pemahaman data yang kita perlu lakukan adalah menggali di bawa permukaan untuk mengungkap struktur masalah bisnis dan data yang tersedia, dan kemudian mencocokkan mereka untuk satu atau lebih banyak dari data dimiliki agar memiliki ilmu yang cukup besar dan teknologi untuk diterapkan.
3. *Data Preparation* : Dalam teknologi analisa kita bahwa ada beberapa persyaratan untuk bagaimana data tersebut dapat digunakan. Terkadang sering dibutuhkan data dalam bentuk yang berbeda dari yang disediakan secara alami, maka beberapa proses konversi akan diperlukan.
4. *Modelling* : Tahap pemodelan adalah tahap utama di mana teknik *data mining* yang dipilih diterapkan untuk data. Pada tahap ini penting untuk memiliki beberapa pemahaman tentang ide-ide dasar *data mining*, termasuk berbagai teknik dan algoritma yang ada, karena ini adalah bagian dari kerajinan dimana

ilmu pengetahuan dan teknologi dapat dibawa untuk digunakan secara tepat.

5. *Evaluation* : Tujuan dari tahap evaluasi adalah untuk menilai hasil *data mining* secara ketat dan untuk mendapatkan kepercayaan diri bahwa mereka adalah valid dan reliabel sebelum dipindahkan. Jika kita melihat cukup keras di setiap dataset kita akan menemukan pola, tetapi mereka mungkin tidak bertahan tanpa pengawasan yang hati-hati.
6. *Deployment* : Dalam tahap penerapan, hasil data mining dan teknik data mining sendiri mulai digunakan nyata dalam rangka mewujudkan beberapa pengembalian investasi dalam tahap penelitian sebelumnya.



Gambar 2. Tahapan data mining (CRISP-DM)

2.6 Clustering

Menurut [13] Analisis Cluster merupakan analisa yang memberikan wawasan ke dalam data dengan membagi obyek ke dalam kelompok (cluster) dari objek, sehingga objek dalam sebuah cluster lebih mirip satu sama lain daripada objek dalam cluster lainnya.

2.7 K-means

K-means adalah algoritma berbasis prototipe, algoritma clustering partitional sederhana yang

mencoba untuk menemukan K yang tidak tumpang tindih pada cluster. Kelompok ini diwakili oleh centroid mereka (centroid cluster biasanya rata-rata poin dalam cluster itu). Proses pengelompokan K-cara sebagai berikut. Pertama, K centroid awal yang dipilih, di mana K ditentukan oleh pengguna dan menunjukkan jumlah yang diinginkan dari cluster. Setiap titik dalam data ini kemudian ditugaskan ke centroid terdekat, dan setiap koleksi poin ditugaskan untuk pusat massa suatu bentuk cluster. Centroid dari setiap cluster kemudian diperbarui berdasarkan poin ditugaskan untuk klaster itu. Proses ini diulang sampai tidak ada titik perubahan cluster [13].

3. METODOLOGI PENELITIAN

3.1 Teknik Pengumpulan Data

Penelitian ini menggunakan data sekunder yang berisi data laporan dari website LAPOR! (<https://www.lapor.go.id>) dan <http://data.go.id/> yang merupakan sebuah portal data publik yang dapat diakses oleh publik secara gratis untuk mendapatkan informasi secara bebas tentang berbagai instansi pemerintah dan berbagai bidang di Indonesia. Data tersebut dikumpulkan melalui data survey dan pendataan langsung dari instansi pemerintah sehingga sumber data dan keabsahannya dapat dipertanggung jawabkan.

Untuk metode pengumpulan data nya penelitian ini akan menggunakan metode kuantitatif dengan tujuan untuk mengembangkan model matematis atau pengembangan sebuah pola dan mencari hubungan diantara pola tersebut untuk dijadikan sebuah informasi yang lebih berguna.

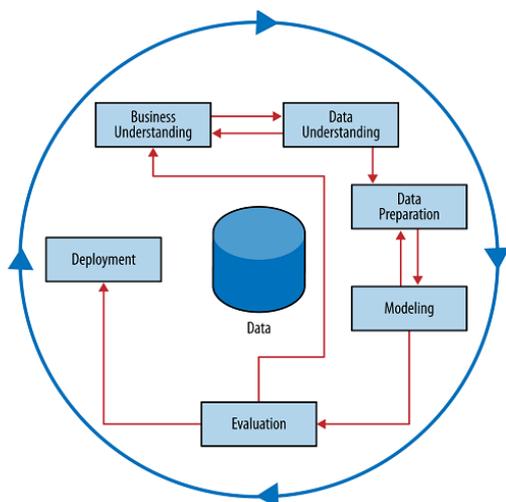
3.2 Teknik Pemrosesan Data

Dalam penelitian ini penulis menggunakan metode (CRISP-DM) CRISP-DM (CROSS-Industry Standard Process for Data Mining) untuk melakukan pemrosesan data secara keseluruhan yang terbagi dalam 6 proses besar yaitu :

1. *Business Understanding* : Pada tahap pertama ini tim desain harus berpikir hati-hati tentang scenario penggunaan (*use case*). Awalnya, sangat penting untuk memahami masalah yang akan dipecahkan. Hal ini mungkin tampak

jelas, tetapi proyek bisnis jarang datang sebagai sesuatu yang bersifat “pradikemas” dan mudah untuk dipecahkan atau tidak ambigu.

2. *Data Understanding* : Dalam tahap pemahaman data yang kita perlu lakukan adalah menggali di bawah permukaan untuk mengungkap struktur masalah bisnis dan data yang tersedia, dan kemudian mencocokkan mereka untuk satu atau lebih banyak dari data yang dimiliki agar memiliki ilmu yang cukup besar dan teknologi untuk diterapkan.
3. *Data Preparation* : Dalam teknologi analisa kita bahwa ada beberapa persyaratan untuk bagaimana data tersebut dapat digunakan. Terkadang sering dibutuhkan data dalam bentuk yang berbeda dari yang disediakan secara alami, maka beberapa proses konversi akan diperlukan.
4. *Modelling* : Tahap pemodelan adalah tahap utama di mana teknik *data mining* yang dipilih diterapkan untuk data. Pada tahap ini penting untuk memiliki beberapa pemahaman tentang ide-ide dasar *data mining*, termasuk berbagai teknik dan algoritma yang ada, karena ini adalah bagian dari kerajinan dimana ilmu pengetahuan dan teknologi dapat dibawa untuk digunakan secara tepat.
5. *Evaluation* : Tujuan dari tahap evaluasi adalah untuk menilai hasil *data mining* secara ketat dan untuk mendapatkan kepercayaan diri bahwa mereka adalah valid dan reliabel sebelum dipindahkan. Jika kita melihat cukup keras di setiap dataset kita akan menemukan pola, tetapi mereka mungkin tidak bertahan tanpa pengawasan yang hati-hati.
6. *Deployment* : Dalam tahap penerapan, hasil data mining dan teknik data mining sendiri mulai digunakan nyata dalam rangka mewujudkan beberapa pengembalian investasi dalam tahap penelitian sebelumnya.



Gambar 3. Tahapan data mining (CRISP-DM)

HASIL DAN PEMBAHASAN

4.1 Rancangan Sistem

Dalam penelitian ini penulis akan melakukan proses pemodelan kluster dari data isi dan judul laporan yang dijadikan sebagai sampel penelitian. Proses tersebut terdiri dari tahap *document preparation* (persiapan dokumen), yaitu proses untuk mempersiapkan data sampel yang digunakan dengan aplikasi *Kemangi* dan modul *text processing* dari aplikasi *Rapidminer*. Setelah proses pemodelan kluster selesai dilakukan dan model telah diperoleh selanjutnya akan dilakukan perancangan GUI (*Graphical User Interface*) dengan bantuan aplikasi *Microsoft Excel* dan *Microsoft Access* agar hasil dari pengelompokkan dan pemodelan kluster dapat ditampilkan kepada para pembaca dan dapat dengan mudah didokumentasikan sebagai hasil penelitian yang telah berhasil dilakukan.

4.2 Rancangan Basis Data

Sumber data yang digunakan sebagai sampel dalam penelitian ini merupakan data laporan masyarakat yang dikumpulkan dan diperoleh dari aplikasi LAPOR! dan LAPOR.go.id. Sumber data tersebut diperoleh secara langsung dari website portal data resmi Indonesia yaitu data.go.id. Dari data laporan tersebut penulis akan melakukan pengelompokkan dan pemodelan kluster dengan algoritma k-means agar pola dari trend judul dan isi laporan dapat terlihat dan dikelompokkan dengan mudah. Tentu saja

seperti yang dijelaskan pada tahap rancangan sistem diatas sebelum data siap untuk digunakan sebagai sampel pemodelan akan dilakukan tahap persiapan dokumen (*document preparation*) dengan bantuan aplikasi *Kemangi* dan modul *Text Processing* dari aplikasi *Rapidminer* untuk data sampel yang akan digunakan. Model kluster dan hasil pengelompokkan berdasarkan trend dari judul dan isi laporan kemudian akan disajikan dalam aplikasi yang dirancang dalam *Microsoft Excel* dan *Access* untuk para pembaca. Proses ini terjadi melalui beberapa proses :

1. *Document Preparation* dan *Pre-Processing*
2. *Cluster Modelling* (Pemodelan Kluster)

4.3 Perancangan dan Implementasi GUI

Dari semua hasil pemodelan yang didapat dari tahap-tahap diatas hasilnya akan di dipindahkan ke dalam *Microsoft Access* untuk disajikan dalam GUI sederhana untuk para pembaca. Pada tahap ini semua hasil penelitian (pemodelan) kluster akan di pindahkan terlebih dahulu ke dalam *Microsoft Excel* terlebih dahulu sebelum dipindahkan ke *Microsoft Access* menggunakan aplikasi *Microsoft Access 2013* untuk menghasilkan tampilan GUI yang simpel dan mudah dilihat oleh pembaca.

Tabel 1
Hasil Wordlist Judul Laporan (2013)

Topik	Total Muncul	Persentase Keseluruhan
jakarta	658	23.89
jalan	351	12.75
banjir	306	11.11
kelurahan	172	6.25
kartu	149	5.41
barat	140	5.08
warga	120	4.36
logistik	119	4.32
liar	112	4.07
guru	110	3.99
bantuan	92	3.34
spbu	92	3.34
utara	90	3.27
pluit	86	3.12
sampah	85	3.09
timur	84	3.05
mohon	79	2.87
rusak	77	2.80
biaya	76	2.76
pembuatan	76	2.76

Tabel 2
Hasil Wordlist Judul Laporan (2014)

Topik	Total Muncul	Persentase Keseluruhan
raskin	2225	81.12
desa	1020	37.19
blsm	419	15.28
aduan	360	13.12
jawa	225	8.20
harga	209	7.62
sesuai	208	7.58
kota	205	7.47
diterima	183	6.67
sasaran	166	6.05
timur	161	5.87
barat	129	4.70
informasi	128	4.67
utara	121	4.41
warga	120	4.37
pemotongan	117	4.27
selatan	90	3.28
menerima	81	2.95
lampung	75	2.73
sumatera	67	2.44

Tabel 3
Hasil Wordlist Judul Laporan (2015)

Topik	Total Muncul	Persentase Keseluruhan
penerima	457	21.73
peserta	236	11.22
informasi	229	10.89
menerima	214	10.18
jawa	210	9.99
barat	174	8.27
kuningan	174	8.27
info	172	8.18
kartu	162	7.70
anak	138	6.56
saldo	134	6.37
jalan	132	6.28
jabar	121	5.75
sasaran	112	5.33
utara	109	5.18
psks	93	4.42
warga	92	4.37
keluarga	91	4.33
pemilik	91	4.33
jakarta	85	4.04

Tabel 4
Hasil Wordlist Isi Laporan (2013)

Topik	Total Muncul	Persentase Keseluruhan
jakarta	1662	21.36
mohon	943	12.12
jalan	625	8.03
warga	470	6.04
kelurahan	435	5.59
pemprov	431	5.54
pemerintah	404	5.19
terima	403	5.18
kasih	382	4.91
provinsi	361	4.64
tolong	347	4.46
bantuan	284	3.65
barat	267	3.43
kartu	267	3.43
banjir	246	3.16
rumah	245	3.15
daerah	242	3.11
kecamatan	242	3.11
melaporkan	236	3.03
sekolah	236	3.03

Tabel 5
Hasil Wordlist Isi Laporan (2014)

Topik	Total Muncul	Persentase Keseluruhan
raskin	2186	31.97
desa	1654	24.19
menerima	1378	20.15
rumah	1186	17.34
tangga	1134	16.58
penerima	664	9.71
blsm	586	8.57
beras	476	6.96
harga	311	4.55
kartu	298	4.36
mohon	224	3.28
orang	224	3.28
miskin	204	2.98
didesa	179	2.62
warga	175	2.56
uang	170	2.49
bantuan	163	2.38
dusun	163	2.38
terima	151	2.21
jawa	140	2.05

Tabel 6
Hasil Wordlist Isi Laporan (2015)

Topik	Total Muncul	Persentase Keseluruhan
mohon	1145	8.86
kartu	1107	8.56
jakarta	973	7.53
anak	922	7.13
jalan	629	4.87
menerima	626	4.84
rumah	586	4.53
bandung	585	4.52
terima	534	4.13
kasih	505	3.91
bpjs	499	3.86
keluarga	487	3.77
sekolah	451	3.49
desa	425	3.29
kota	425	3.29
pemprov	404	3.12
tangga	372	2.88
nama	363	2.81
warga	343	2.65
orang	320	2.48

Tabel 7
Hasil Cluster Data Judul Laporan 2013

Hasil Klaster Judul Laporan 2013					
Atribut Cluster 0	Cluster 0	Atribut Cluster 1	Cluster 1	Atribut Cluster 2	Cluster 2
Jalan	0.0207	informasi	0.3139	banjir	0.1649
kelurahan	0.0164	manfaat	0.2652	jakarta	0.1412
Jakarta	0.0163	penggunaan	0.2585	logistik	0.1046
Guru	0.0143	kartu	0.2081	kartu	0.0734
Liar	0.0129	blsm	0.1522	pluit	0.0662
Warga	0.0124	pencairan	0.1303	pintar	0.0577
Spbu	0.0123	jadual	0.0909	penjaringan	0.0564
Raskin	0.0119	sasaran	0.0345	evakuasi	0.0467
Barat	0.0109	sehat	0.0220	sehat	0.0460
transjakarta	0.0105	sekolah	0.0167	jalan	0.0408
Sampah	0.0103	kriteria	0.0159	bantuan	0.0348
Rusak	0.0098	sesuai	0.0132	korban	0.0298
Pungutan	0.0098	pembuatan	0.0122	pekojan	0.0287
Biaya	0.0095	peruntukan	0.0108	kapuk	0.0277
Harga	0.0091	saran	0.0107	muara	0.0245
Mohon	0.0089	jakarta	0.0106	barat	0.0236
Sekolah	0.0088	fungsi	0.0105	utara	0.0199
bangunan	0.0088	penggunakan	0.0097	bandengan	0.0192
pembuatan	0.0085	raskin	0.0082	wilayah	0.0171

Tabel 8
Hasil Cluster Data Judul Laporan 2014

Hasil Klaster Judul Laporan 2014					
Atribut Cluster 0	Cluster 0	Atribut Cluster 1	Cluster 1	Atribut Cluster 2	Cluster 2
Blsm	0.043	raskin	1.000	jawa	0.418
Aduan	0.038	aada	0.000	timur	0.144
Harga	0.037	abad	0.000	barat	0.087
Desa	0.035	abang	0.000	purworejo	0.059
Sasaran	0.029	abian	0.000	raskin	0.047
Sesuai	0.028	abung	0.000	kebumen	0.035
Raskin	0.026	aceh	0.000	semarang	0.032
Diterima	0.025	adibarang	0.000	lumajang	0.028
menerima	0.023	adikarso	0.000	banyumas	0.028
informasi	0.021	admin	0.000	kediri	0.025
pemotongan	0.021	administrasi	0.000	jombang	0.024
Kota	0.017	aduan	0.000	pemalang	0.024
Warga	0.016	agam	0.000	magelang	0.020
Utara	0.016	agung	0.000	sukabumi	0.020
Sumatera	0.012	agustus	0.000	bandung	0.020
Liter	0.012	aikmel	0.000	brebes	0.020
Lampung	0.012	ajam	0.000	blitar	0.018
Info	0.012	akses	0.000	mojokerto	0.018
Selatan	0.012	akta	0.000	bekasi	0.018

Tabel 9
Hasil Cluster Data Judul Laporan 2015

Hasil Klaster Judul Laporan 2015					
Atribut Cluster 0	Cluster 0	Atribut Cluster 1	Cluster 1	Atribut Cluster 2	Cluster 2
Blsm	0.043	raskin	1.000	jawa	0.418
Aduan	0.038	aada	0.000	timur	0.144
Harga	0.037	abad	0.000	barat	0.087
Desa	0.035	abang	0.000	purworejo	0.059
Sasaran	0.029	abian	0.000	raskin	0.047
Sesuai	0.028	abung	0.000	kebumen	0.035
Raskin	0.026	aceh	0.000	semarang	0.032
Diterima	0.025	adibarang	0.000	lumajang	0.028
menerima	0.023	adikarso	0.000	banyumas	0.028
informasi	0.021	admin	0.000	kediri	0.025
pemotongan	0.021	administrasi	0.000	jombang	0.024
Kota	0.017	aduan	0.000	pemalang	0.024
Warga	0.016	agam	0.000	magelang	0.020
Utara	0.016	agung	0.000	sukabumi	0.020
Sumatera	0.012	agustus	0.000	bandung	0.020
Liter	0.012	aikmel	0.000	brebes	0.020
Lampung	0.012	ajam	0.000	blitar	0.018
Info	0.012	akses	0.000	mojokerto	0.018
Selatan	0.012	akta	0.000	bekasi	0.018

Tabel 10
Hasil Cluster Data Isi Laporan 2013

Hasil Klaster Isi Laporan 2013					
Atribut Cluster 0	Cluster 0	Atribut Cluster 1	Cluster 1	Atribut Cluster 2	Cluster 2
Bantuan	0.022	kartu	0.084	jalan	0.065
Warga	0.020	guru	0.068	jakarta	0.038
Banjir	0.016	sekolah	0.056	pemprov	0.035
kelurahan	0.015	jakarta	0.040	mohon	0.034
Rumah	0.015	anak	0.038	pemerintah	0.027
Spbu	0.014	pintar	0.029	provinsi	0.027
Mohon	0.013	sehat	0.029	tindak	0.027
Jakarta	0.013	biaya	0.027	terimakasih	0.026
Daerah	0.013	sertifikasi	0.025	melaporkan	0.023
Tolong	0.013	mohon	0.025	kasih	0.023
Raskin	0.011	terima	0.024	terima	0.022
Utara	0.011	penerima	0.023	sampah	0.022
Terima	0.011	dana	0.023	pasar	0.021
Pluit	0.011	murid	0.023	terkait	0.021
Kasih	0.010	siswa	0.022	raya	0.020
Orang	0.010	kasih	0.022	jembatan	0.020
Mobil	0.009	uang	0.021	busway	0.020
Barat	0.009	orang	0.020	rusak	0.019
Bensin	0.009	keluarga	0.019	transjakarta	0.019

Tabel 11
Hasil Cluster Data Isi Laporan 2014

Hasil Klaster Isi Laporan 2014					
Atribut Cluster 0	Cluster 0	Atribut Cluster 1	Cluster 1	Atribut Cluster 2	Cluster 2
Blsm	0.030	tangga	0.117	tangga	0.448
Beras	0.022	rumah	0.112	menerima	0.435
Kartu	0.021	menerima	0.106	rumah	0.428
Desa	0.020	penerima	0.103	desa	0.321
Raskin	0.018	desa	0.072	raskin	0.222
Orang	0.016	harga	0.070	penerima	0.205
Penerima	0.016	didesa	0.069	harga	0.033
menerima	0.015	raskin	0.069	beras	0.019
Mohon	0.015	beras	0.051	uang	0.003
Miskin	0.014	liter	0.049	kepala	0.002
Bantuan	0.013	tebus	0.032	kasih	0.002
Terima	0.013	perbulan	0.013	blsm	0.002
Uang	0.012	uang	0.013	tolong	0.002
Dana	0.012	bulannya	0.011	didesa	0.002
Rumah	0.012	bulan	0.010	aaqzh	0.000
Warga	0.012	blsm	0.010	aathq	0.000
Harga	0.011	haya	0.010	abadi	0.000
Tangga	0.011	harganya	0.009	abaiqn	0.000
Dusun	0.011	jatah	0.009	abang	0.000

Tabel 12
Hasil Cluster Data Isi Laporan 2015

Hasil Klaster Isi Laporan 2015					
Atribut Cluster 0	Cluster 0	Atribut Cluster 1	Cluster 1	Atribut Cluster 2	Cluster 2
Kartu	0.193	jakarta	0.020	tangga	0.341
Anak	0.185	mohon	0.019	menerima	0.327
Pintar	0.115	saldo	0.016	rumah	0.284
Sekolah	0.100	bpjs	0.015	keluarga	0.172
Indonesia	0.097	jalan	0.015	desa	0.163
menerima	0.076	bandung	0.014	penerima	0.086
Penerima	0.073	terima	0.013	kartu	0.041
Keluarga	0.053	kasih	0.013	card	0.035
Orang	0.032	kartu	0.011	anak	0.034
Sehat	0.024	pemprov	0.011	sekolah	0.030
Tangga	0.022	kesehatan	0.011	didesa	0.027
Rumah	0.019	kota	0.011	anggota	0.022
Terima	0.016	anak	0.010	pintar	0.013
Nama	0.015	warga	0.009	indonesia	0.009
bersekolah	0.015	kelurahan	0.009	siswa	0.008
Card	0.012	tolong	0.009	memiliki	0.008
Bantuan	0.012	nama	0.009	nama	0.008
Kasih	0.012	pemerintah	0.009	bersekolah	0.008
Anggota	0.011	kecamatan	0.008	bywvl	0.007

**Analisis Pengelompokan Laporan
Pengaduan Masyarakat Pada Sistem
Lapor.go.id Dengan Algoritma K-Means
Periode 2013-2015**



Gambar 4
Tampilan Menu Utama

4.4 Implikasi Penelitian

Penelitian ini memberikan implikasi bagi institusi yaitu dalam penelitian ini adalah pihak sistem LAPOR!, antara lain :

1. Menyediakan hasil penelitian penerapan text mining untuk melakukan pengelompokan data laporan dengan algoritma k-means.
2. Membuat sebuah model pengelompokan yang dapat membantu kemajuan sistem LAPOR! di masa depan agar lebih efektif.

5 SIMPULAN

Dari hasil penelitian yang dilakukan untuk mengelompokkan data laporan masyarakat pada sistem lapor.go.id dengan algoritma *k-means* clustering penulis dapat menyimpulkan beberapa hal berikut :

1. Diperolehnya data laporan masyarakat yang telah terstruktur menurut topik yang paling sering dilaporkan melalui data wordlist.

2. Diterapkannya pengelompokan judul laporan berdasarkan algoritma *k-means* clustering untuk menemukan topik atau tema yang paling sering muncul pada tiap kluster untuk memudahkan pengelompokan data berdasarkan trend judul atau kategori.
3. Diterapkannya pengelompokan isi laporan berdasarkan algoritma *k-means clustering* untuk menemukan topik atau tema yang paling sering muncul pada tiap kluster untuk memudahkan pengelompokan data berdasarkan trend pendapat pelapor atau kategori.
4. Diperolehnya data yang berisi kategori dari judul dan isi laporan berdasarkan data kluster pada masing-masing data tersebut yang dapat diolah kembali sesuai dengan keinginan.

6 REKOMENDASI

Dalam pembuatan penelitian karya akhir ini penulis sekiranya memberi saran atau masukan lanjutan untuk para pembaca dan

peneliti seiring dengan penyelesaian penelitian yang dilakukan yaitu antara lain:

1. Sulitnya menerapkan algoritma k-means secara manual tanpa bantuan aplikasi data mining seperti Rapidminer sehingga penelitian masih tergantung dengan komputasi sistem tanpa penghitungan secara manual dalam jumlah besar. Pada penelitian selanjutnya penulis berharap adanya penerapan penghitungan secara manual yang dicantumkan untuk mendukung keakuratan penelitian menggunakan algoritma k-means.
2. Data laporan yang diperoleh didapat dari situs penyedia data publik gratis sehingga data yang ada belum dapat diteliti secara aktual karena harus menunggu pengumpulan data dari pihak yang mengelola. Ini dikarenakan karena keterbatasan aplikasi rapidminer yang digunakan oleh penulis yang tidak mampu untuk melakukan web crawling untuk mengumpulkan data secara aktual. Pada penelitian berikutnya diharapkan dapat dilakukan secara aktual dengan menerapkan pengumpulan data berupa web crawling.
3. Dalam penelitian ini jumlah data yang dijadikan sampel dikurangi dalam jumlah yang cukup banyak oleh penulis karena keterbatasan spesifikasi perangkat lunak dan perangkat keras sehingga waktu yang dibutuhkan untuk memproses data dengan jumlah yang banyak memakan cukup waktu yang lama. Penulis berharap dengan spesifikasi yang lebih mendukung dan aplikasi data mining dan text mining yang lebih memadai, peneliti berikutnya dapat memaksimalkan penggunaan algoritma k-means untuk memproses data yang lebih banyak dengan hasil yang lebih akurat.

7 DAFTAR REFERENSI

- [1] A.North, Matthew (2012), *Data Mining for the Masses*, Edisi ke-1, Georgia : Global Text Project.
- [2] Anggrawal, Charu C (2015), *Data Mining The Textbook*, New York :

- Springer International Publishing. Brown, Meta S. (2014), *Data Mining For Dummies*, New Jersey : John Wiley & Sons, Inc.
- [3] Brown, Meta S. (2014), *Data Mining For Dummies*, New Jersey : John Wiley & Sons, Inc.
 - [4] Ediyanto, et al (2013), *Jurnal : Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis*, Buletin Ilmiah Mat. Stat. dan Terapannya, Jakarta.
 - [5] Han, Jiawei et al (2012), *Data Mining Concepts and Techniques*, Edisi ke-3 Waltham : Elsevier Inc.
 - [6] Kroenke David M. (2014), *Experienceing MIS*, Edisi ke-5, New Jersey : Pearson Education, Inc
 - [7] Laudon, Kenneth C. dan Jane P. Laudon, (2014), *Management Information Systems: Managing the Digital Firm*, Edisi ke-13, New Jersey : Pearson Prentice Hall.
 - [8] M.Stair, Ralph dan George W. Reynolds (2012), *Fundamentals of Information Systems*, Edisi ke-6, Boston : Cengage Learning.
 - [9] O'Brien James A. dan George M Marakas (2012), *Introduction to Information Systems*, Edisi ke-16, New York: Mc Graw Hill Irwin.
 - [10] Provott, Foster dan Tom Fawcett (2013), *Data Science for Business*, Sebastopol : O'Reilly Media, Inc.
 - [11] Ramadhani, Rima Dias (2012), *Jurnal : Data Mining Menggunakan Algoritma K-Means Clustering untuk Menentukan Strategi Promosi Universitas Dian Nuswantoro*, Universitas Dian Nuswantoro, Semarang.
 - [12] Valacich C Joseph dan Joey F George (2015), *Essential of Systems Analysis and Design*, Edisi ke-6 Edinburgh Gate : Pearson Education Limited.
 - [13] Wu Junjie (2012), *Advances in K-Means Clustering A Data Mining Thinking*, New York : Springer Publications.