

## APLIKASI PENGUKUR TINGKAT SENTIMEN PELANGGAN BERDASARKAN KOMPLAIN PELANGGAN PLN MENGGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR*

Rahmiati<sup>1</sup>, Dedy Irfan<sup>2</sup>, Agustin<sup>3</sup>, Siska Hedyati<sup>4</sup>

STMIK Amik Riau, Jl.Purwodadi Indah Km.10 Panam Pekanbaru

Universitas Negeri Padang, Jl. Prof. DR. Hamka Air Tawar Padang

STMIK Amik Riau, Jl.Purwodadi Indah Km.10 Panam Pekanbaru

STMIK Amik Riau, Jl.Purwodadi Indah Km.10 Panam Pekanbaru

*rahmiati@sar.ac.id*<sup>1</sup>, *irfankumango@yahoo.com*<sup>2</sup>, *agustin@sar.ac.id*<sup>3</sup>, *siskahedyati@stmik-amik-riau.ac.id*<sup>4</sup>

**Abstrack** - PLN is a state electricity company that has a role in providing services to the public regarding electricity. Problems that often occur due to PLN electricity are power outages caused by several factors such as weather, damage or because of savings. These problems result in losses in business, health, education and traffic flow. These various obstacles have an impact on the society, causing the community to comment both directly and through social media. Comments on the PLN service can be used as input for PLN in providing better services to the public. These comments are a source of data in text mining research for the classification of comment types. Users can give good and bad opinions. By using data processing techniques in text mining, this study utilizes the KNN (*K-Nearest Neighbor*) algorithm. In this study used a dataset of 100 comments and composed of 3 classes, namely positive, negative, and neutral. The purpose of this classifier is to find out the level of public sentiment towards PLN so that it can be an input for PLN to provide better services From the results of the training data and test data, it is found that the *K-Nearest Neighbor* algorithm is able to group these comments into positive, negative and neutral classes.

**Keywords** - sentiment, PLN, monitoring, android

**Intisari** - PLN merupakan perusahaan milik negara yang mempunyai peran dalam menyediakan pelayanan kepada masyarakat mengenai listrik. Permasalahan yang sering terjadi mengenai listrik PLN adalah pemadaman listrik yang disebabkan oleh beberapa faktor seperti cuaca, kerusakan maupun karena penghematan. Masalah tersebut mengakibatkan kerugian pada lini bisnis, kesehatan, pendidikan dan arus lalu lintas. Berbagai kendala tersebut memberikan dampak kepada masyarakat sehingga menyebabkan masyarakat berkomentar baik secara langsung maupun melalui sosial media. Adanya komentar pada layanan PLN tersebut, dapat dijadikan masukan untuk PLN dalam memberikan layanan yang lebih baik lagi kepada masyarakat. Komentar tersebut menjadi sumber data dalam penelitian *text mining* untuk klasifikasi jenis komentar. Pengguna dapat memberikan opini yang baik maupun tidak baik. Dengan menggunakan teknik pengolahan data dalam *text mining*, penelitian ini memanfaatkan hal tersebut dengan menggunakan Algoritma KNN (*K-Nearest Neighbor*). Pada penelitian ini menggunakan *dataset* 100 komentar dan terdiri dari 3 kelas, yaitu positif, negatif, dan netral. Tujuan dari pengklasifikasi ini adalah untuk mengetahui tingkat sentimen masyarakat terhadap PLN sehingga dapat menjadi masukan untuk PLN memberikan layanan yang lebih baik. Dari hasil data latih dan data uji, didapatkan Algoritma KNN (*K-Nearest Neighbor*) mampu mengelompokkan komentar tersebut ke dalam kelas positif, negatif dan netral.

**Kata Kunci** - sentimen, PLN, monitoring, android

## I. PENDAHULUAN

PT. PLN Persero merupakan perusahaan BUMN yang memiliki bidang pekerjaan mengenai kelistrikan. Pada akhir abad ke 19 PT. PLN dibentuk untuk dapat mengorganisir masalah kelistrikan yang ada di Indonesia. Semakin berkembangnya masyarakat berpengaruh pada tingkat pasokan listrik yang harus dipenuhi, sehingga PT. PLN membuat kantor pelayanan hampir diseluruh Indonesia sekaligus berimbas pada kelengkapan tenaga kerja atau penambahan karyawan yang besar dan dilakukan setiap tahunnya. PT. PLN memberikan layanan mengenai listrik kepada masyarakat. Informasi layanan yang diberikan oleh PLN sangat penting karena apabila listrik mati dalam beberapa jam maka akan mengakibatkan kerugian pada lini bisnis, kesehatan, pendidikan maupun arus lalu lintas.

Banyak kendala dan masalah yang terjadi hampir setiap tahunnya membuat konsumen ada yang merasa puas atau tidak puas dengan layanan. Saat ini PT. PLN memberikan *hotline* untuk layanan kepada konsumen namun itu dirasa masih kurang efektif bagi beberapa pengguna yang akan melaporkan atau komplain karena biaya telepon yang tidak murah dan juga sering sibuk jika masyarakat banyak yang menghubungi. Menurut bagian *Supervisor Sumber Daya Manusia dan Administrasi PLN*, banyaknya komentar, komplain dan aduan yang tersebar di media sosial yang ditujukan kepada PLN, dapat di jadikan pertimbangan dan menjadi referensi bagi PLN dalam mengambil tindakan kedepannya. Hal ini mengharuskan PLN mempunyai satu wadah baru untuk melayani komplain maupun aduan dari masyarakat. Dari banyaknya komentar/komplain/aduan dapat dihasilkan data mengenai tingkat sentimen masyarakat terhadap PLN sehingga dapat menjadi masukan untuk PLN dalam memberikan layanan yang lebih baik lagi kepada masyarakat.

Salah satu solusi yang ditawarkan adalah membangun aplikasi pengukur tingkat sentimen pelanggan sehingga masyarakat dapat dengan mudah memberikan komplain, komentar atau aduan dengan cepat, dan PLN dapat memberikan layanan maksimal dengan melihat hasil dari aplikasi tersebut. Aplikasi yang dibangun akan menggunakan platform *android*. Menganalisa komentar pelanggan, digunakan algoritma *K-Nearest Neighbor*, kelebihan dari algoritma ini yaitu terbukti mencapai hasil akurasi yang baik dan sesuai dengan perhitungan yang diterapkan dalam sebuah aplikasi[1], sehingga dalam penilaian tingkat sentimen dapat bekerja dengan baik. Aplikasi yang dibangun juga dilengkapi dengan informasi-informasi dari PLN untuk pelayanan masyarakat.

## II. SIGNIFIKANSI STUDI

### A. *Sentiment Analysis*

*Sentiment Analysis* merupakan perpaduan dari data mining dan *text mining*, yaitu sebuah cara yang digunakan untuk mengolah berbagai opini yang diberikan oleh konsumen melalui berbagai media, mengenai sebuah produk, jasa ataupun sebuah instansi. *Sentiment analysis* merupakan sebuah metode yang digunakan untuk memahami, mengekstrak data opini, dan mengolah data tekstual secara otomatis untuk mendapatkan sebuah *sentiment* yang terkandung dalam sebuah opini. Pada *Sentiment analysis* terdiri dari 3 jenis opini, yaitu opini positif, opini negatif dan opini netral. Adanya *sentiment analysis* membuat perusahaan atau instansi yang terkait dapat mengetahui respon masyarakat terhadap suatu pelayanan atau produk, melalui umpan balik dari masyarakat [2].

Analisis sentimen diartikan sebagai studi komputasional dari opini-opini orang, *appraisal* dan emosi melalui entitas, *event* dan atribut yang dimiliki. Tugas dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau fitur. Pendapat yang dikemukakan dalam dokumen dikelompokkan menjadi positif, negatif atau netral. [3].

### B. *Algoritma K-Nearest Neighbor (K-NN)*

Algoritma *K-Nearest Neighbor (K-NN)* adalah algoritma yang digunakan untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut [4]. Algoritma K-NN merupakan algoritma *supervised learning*. Algoritma *supervised learning* bertujuan untuk mendapatkan pola baru sedangkan *unsupervised learning* tujuannya untuk mendapatkan pola dalam sebuah data[5].

Prinsip kerja K-NN adalah mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga (*neighbor*) terdekatnya dalam data pelatihan. Data training diproyeksikan keruang berdimensi banyak, yang mana pada masing-masing dimensi menjelaskan fitur dari data yang ada. Masing-masing dimensi dibagi pada ruang-ruang berdasarkan klasifikasi data training. Sebuah titik pada ruang ditandai dengan kelas c, jika kelas tersebut merupakan klasifikasi yang banyak ditemui pada k buah tetangga terdekat titik tersebut[6].

Ketepatan algoritma KNN sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Algoritma KNN memiliki kelebihan yaitu ketangguhan terhadap data latih yang memiliki banyak noise dan efektif apabila data latihnya besar. Kelemahan KNN adalah perlunya menentukan nilai dari parameter K (jumlah dari tetangga terdekat), training berdasarkan jarak tidak jelas mengenai jenis jarak apa yang harus digunakan dan atribut mana yang harus digunakan untuk mendapatkan hasil terbaik, dan biaya komputasi cukup tinggi karena diperlukan perhitungan jarak dari tiap query instance pada keseluruhan *training sample* [7].

Kedekatan didefinisikan dalam jarak metrik, seperti jarak Euclidean. Jarak Euclidean dapat dicari dengan menggunakan persamaan pada langkah-langkah algoritma KNN berikut.

Langkah-langkah untuk menghitung metode K-Nearest Neighbor antara lain:

1. Tentukan parameter *K* (jumlah tetangga paling dekat).
2. Hitung kuadrat jarak *Euclid (query instance)* masing-masing objek terhadap data sampel yang diberikan menggunakan rumus:

$$D(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan:

D : jarak kedekatan

x : data training

y : data testing

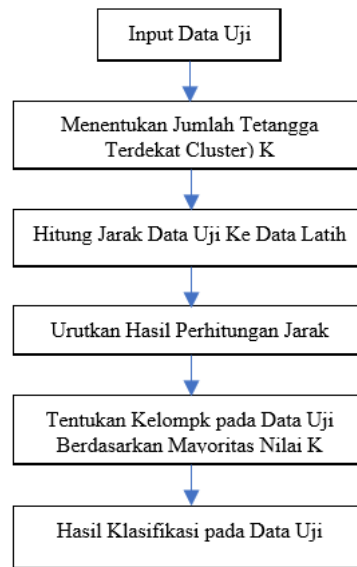
$\Sigma$  : jumlah atribut individu antara 1 s.d. *n*

f : fungsi similitary atribut *i* antara kasus *X* dan kasus *Y*

*i* = Atribut individu antara 1 sampai dengan *n*

3. Urutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *Euclid* terkecil.
4. Kumpulkan kategori *Y* (Klasifikasi *Nearest Neighbor*)
5. Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka dapat diprediksi nilai *query instance* yang telah dihitung. [8]

Berikut adalah proses algoritma *K-Nearest Neighbor (K-NN)*:



**Gambar 1. Proses Metode *K-Nearest Neighbor*  
Sumber : Kurniawan, 2019**

Proses metode K-NN dimulai dari menginput data uji, lalu menentukan tetangga terdekat(k), lalu hitung jarak data uji ke data latih, urutkan hasil perhitungan jarak, setelah itu tentukan kelompok pada data uji berdasarkan mayoritas nilai K, dan setelahnya akan didapat hasil klasifikasi pada data uji.

### III. HASIL DAN PEMBAHASAN

Pada penelitian ini menggunakan *dataset* 100 komentar. Data ini diambil dari Sosial Media PT. PLN Pekanbaru. Berikut beberapa data komentar dari pelanggan PLN :

Tabel 1. Data Komentar Pelanggan PLN

No	Komentar
1	JL. JEND. SUDIRMAN, HARI INI LISTRIK MATI SAMPAI 2X, DAN SAMPAI SAAT SEKARANG BELUM JUGA HIDUP.. INI KENAPA????
2	Assalamualaikum pa,,kok lampu sudah mati lgi. Padahal tadi sudah mati lampu dari jam6 sampai jam 8,,sekarang sudah mati lagi. Apa ada perbaikan??? Wasalam
3	Kabel listrik tegangan tinggi sudah sangat kendor, tolong diperbaiki pak. Lokasi jalan sei mintan
4	Tolong cara cari tau id planggan di meteran prabayar. Saya baru ngontrak, barcode dan id gak ada di meteran...???
5	Pak, ini listriknya hidup jam brp ya? Wilayah jln tanjung uban tanjung rhu. Mati sejak pukul 17.20. Hingga saat ini blm juga hidup.
6	Lapor pak, di daerah kapau sari bukit barisan listriknya sering anjlok/turun naik sdh hampir 1 bln pak, apa lagi sejak pengantian tiang minggu kemaren..... apa solusinya & kapan di diperbaiki ? (terima kasih pak)
7	Yth bpk2, Mohon informasi, dan pertanggungjawabannya .. Di kawasan Rumbai Pesisir mendadak tegangan listrik kami meninggi pd pukul 11 wib. Akibatnya, peralatan elektronik kami meledak dan terbakar : televisi, kipas angin, lampu. TOLONG DIRESPON SEGERA.... Terimakasih

8	Tolong pejbt pln jgn mau di atur oleh oknum tertentu. Krn di tempat kami dusun Iv lading tinggan desa senama nenek sampai sekarng jaringan atau kabel pln belum terpasang. Padahal tiangnya sudah capek berdiri...
9	Mohon bantuannya pln...listrik di rowo bening jl.teratai sukarno hatta padam dari jam 7 malam sampai sekarang....mohon sangat bantuannya..trimakasih
10	Pak dirktur pln yg terhormat,klomter saya kok gk bisa di isi pulsa ya,,daya nya 900
11	diGaruda sakti kilo listrik mati dari pagi sapai sekarak tolong dicik dok pak
12	bagus pelayanannya karena
13	proses pelayanan sangat mumuas kan bagi masyarakat
14	Para Pejabat PLN Pekanbaru, tlg lah kami warga jl Patria sarivl rumbai, Dr td siang mati lampu, org PLN ga ada yg dtg periksa dmn kerusakannya
23	pelayanannya gangguan memuaskan
24	Dapatkan informasi seputar PLN Riau - Kepulauan Riau
25	listrik tanpa padam di Area Pekanbaru
26	Pelayanan yg sangat memuaskan baru lapor setengah jam petugas langsung datang ke rumah
27	Pak PLN tadi baik kalau tanya sopan Dan senyum, itu baru pak PLN yg baik Dan bisa di contoh
28	pak tokennya mahal 50 buat lima hari ( 50 ribu = 32kwh ), kami mahasiswa cuman ngontrak, uang jajan 800 sebulan. saya ngontrak berdua. cobak dihitung sebulan kami harus keluarin berapa untuk listrik ? (50 x 6 =300 ribu) Rp.300,000,00-. itu berat pak buat kami. uang jajan kami gak sebanyak anak-anak bapak yang bisa sampek puluhan juta atau bahkan ratusan juta. kalau udah 20 kwh berisik pak. gak konsen ngerjain tugas. tolong pak di survie siapa pengguna KWH, bukan di pukul rata. KWH kami di ganti, pulsa setelah penggantian kwh tidak di kemablikan. penggantian KWH di pungut biaya Rp.100,000. itu kwh udah di beli pak, dan kami harus bayar lagi ? yang saya tahu penggantian kwh meter tidak di pungut biaya. dan saya mohon pekerja lapangan yang datang untuk melakukan pengecekan atau penggantian perangkat kwh memiliki sopan santun, jangan kayak maling. tolong pak direspon.
29	Dengan bapak terhormat. Nie Listrik sampai kapan matinya. Dari jam 06 sampai sekarang masih mati udh 3-4 jam mati PASAR MANDAU DAN SEKITARNYA Mohon Segera di hidupkan kembali.. aktivitas terganggu. DURI-Riau Terima kasih
30	Saya heran apa maksud nya biaya listrik naik 2 kli lipat... apa klian gak mikirkan orang2 yg hnya pas2an ntuk hidup tiap bulan hnya hbis buat bayar listrik... sya biasa 150 perbulan skarang sampe 300 pmakaian kayak biasa aj gak terlalu berlebihan...
31	Kapankah PLN bisa membuat masyarakat tersenyum, kapankah PLN tidak mematikan jaringan padahal kami bayar secara elektronik loohh... Apa untungnya bagi PLN dengan mematikan lampu secara sepihak tanpa pemberitahuan. Mal ini saja sudah 2 x loh matinya
96	Mohon bsntuanya dan solusinya Meteran Dirumah saya prabayar token eliktrik permasalahanya setiap magrib lampunya mati,aliran listrik terputus,sementara lampu dimeteran hidup,lampu di ruma tetangga menyala. Apa solusinya merek meteran Fed.
97	Mantap pt.pln wilayah propinsi riau kami aptesiaai bantuanya...mantap dan sukses pln
98	Allhamdulillah...ini yg ditunggu-tunggu masyarakat ladan/palmatak..smoga proses sdg operasi lancar dan aman, insha allah barokah...aamiin yra
99	Lampu mati sebagian dikarenakan gangguan cuaca, sudah 6jam PLN tidak merespon, sudah 3org menghubungi PLN masih tidur
100	Kubang sudah padam dr jam 7 pagi boss!! Kebangetan!!

### III.1 Perhitungan KNN

Pada perhitungan metode ini akan dilakukan perhitungan berdasarkan sample data yang didapat, namun pada perhitungan manual ini hanya akan digunakan 4 data 3 sebagai data latih/sample dan 1 sebagai data uji atau hasil proses yang *KNN*. Berikut ini contoh datanya dari komentar pelanggan PLN:

Tabel 2. Data Uji Komentar Pelanggan

No	Username	Komentar	Kelas
1	Antarctic Osram	Pak PLN tadi baik kalau tanya sopan Dan senyum , itu baru pak PLN yg baik hati Dan bisa di contoh	Positif
2	Kiss Kiss Scene	Pln banyak mati lampunya kacau	Negatif
3	Yulia Sarah	Adu ini ibu ibu yang katanya buru2 mau jemput anaknya yang sakit malah kegirangan foto2 dulu, geli tapi ngakakðŸœ€	Netral
4	Norman Khamurun	Bapak-bapak dari pln tadi baik sopan	?

Setelah proses pengambilan data pelabelan pada setiap komentar, selanjutnya dilakukan proses *text preprocessing*. *Text Preprocessing* merupakan tahapan untuk mempersiapkan teks menjadi data yang akan siap diolah lebih lanjut. Suatu teks tidak dapat diproses langsung oleh algoritma pencarian, oleh karena itu dibutuhkan *preprocessing text* untuk mengubah teks menjadi data *numeric*[9].

Tahapan dari *Text preprocessing* sebagai berikut:

#### 1. Case Folding

*Case folding* merupakan untuk mengubah seluruh huruf yang ada pada dokumen menjadi huruf kecil. Berikut merupakan hasil data komentar yang telah di-*Case Folding*.

Tabel 3. Hasil Case Folding

No	Komentar	Hasil Case Folding
1	Pak PLN tadi baik kalau tanya sopan Dan senyum , itu baru pak pln yg baik hati Dan bisa di contoh	pak pln tadi baik kalau tanya sopan dan senyum , itu baru pak pln yg baik hati dan bisa di contoh
2	Pln banyak mati lampunya kacau	pln banyak mati lampunya kacau
3	Adu ini ibu ibu yang katanya buru2 mau jemput anaknya yang sakit malah kegirangan foto2 dulu, geli tapi ngakakðŸœ€	adu ini ibu ibu yang katanya buru2 mau jemput anaknya yang sakit malah kegirangan foto2 dulu, geli tapi ngakakðŸœ€

#### 2. Cleaning

*Cleaning* adalah tahapan membersihkan dokumen dari komponen yang tidak memiliki hubungan dengan informasi yang ada pada dokumen, seperti karakter atau simbol, angka, *emoticon* dan *link URL*. Berikut merupakan hasil data komentar yang telah di *cleaning*.

Tabel 4. Hasil Cleaning

No	Komentar	Hasil Cleaning
1	pak pln tadi baik kalau tanya sopan dan senyum , itu baru pak pln yg baik hati dan bisa di contoh	pak pln tadi baik kalau tanya sopan dan senyum itu baru pak pln yg baik hati dan bisa di contoh
2	pln banyak mati lampunya kacau	pln banyak mati lampunya kacau
3	adu ini ibu ibu yang katanya buru2 mau jemput anaknya yang sakit malah kegirangan foto2 dulu, geli tapi ngakakðŸœ€	adu ini ibu ibu yang katanya buru mau jemput anaknya yang sakit malah kegirangan foto dulu geli tapi ngakak

### 3. Tokenizing

*Tokenizing* yaitu proses untuk mengelompokkan isi teks yang semula berupa kalimat hingga menjadi satuan kata-kata. Berikut merupakan hasil data komentar yang telah di *Tokenizing*.

Tabel 5. Hasil *Tokenizing*

D1	D2	D3
Pak	Pln	adu
Pln	banyak	ini
Tadi	mati	ibu
Baik	lampunya	ibu
Kalua	kacau	yang
Tanya		katanya
Sopan		baru
Dan		mau
Senyum		jemput
Itu		anaknya
Baru		yang
Pak		sakit
Pln		malah
Yang		kegirangan
Baik		foto
Hati		dulu
Dan		geli
Bias		tapi
Di		ngakak
contoh		

### 4. Normalisasi

Proses normalisasi merupakan untuk mengubah kata yang tidak baku atau sesuai ejaan, pada tahap ini dilakukan dengan menggunakan *database* kamus kata Bahasa baku dan tidak baku yang dibuat sendiri berdasarkan dari data komentar yang digunakan.

### 5. Filtering

Tahapan *Filtering* yaitu proses pengambilan kata-kata penting dari hasil *tokenizing*, atau proses pembuangan *stopword*. Pembuangan *stopword* adalah proses pembuangan *term* yang tidak memiliki arti atau tidak relevan. *Term* yang diperoleh dari tahap *tokenisasi* dicek dalam suatu daftar *stopword*, apabila sebuah kata tidak masuk didalam daftar *stopword* maka kata tersebut akan masuk ke proses berikutnya[10].Berikut merupakan hasil data komentar yang telah di *Filtering*.

Tabel 6. Hasil *Filtering*

D1	D2	D3
bapak	pln	ibu
pln	banyak	ibu
baik	mati	katanya
tanya	lampunya	buru
sopan	kacau	mau
senyum	pln	jemput
bapak		anaknya
pln		sakit
baik		kegirangan

hati	foto
contoh	geli
	lucu

**6. Stemming**

*Stemming* merupakan proses mencari kata dasar dari setiap kata hasil proses *filtering* sebelumnya. *Stemming* akan melakukan proses penemuan kata dasar dari sebuah kata, dengan menghilangkan semua imbuhan[11]. Berikut merupakan hasil data komentar yang telah di *Stemming*.

Tabel 7. Hasil *Stemming*

D1	D2	D3
bapak	pln	ibu
pln	banyak	ibu
baik	mati	kata
tanya	lampu	jemput
sopan	kacau	anak
senyum		sakit
bapak		girang
pln		foto
baik		geli
hati		lucu
contoh		

**7. Tahap Pembobotan (TF-IDF)**

Pembobotan adalah proses mengubah kata menjadi bentuk angka atau *vector*. Sedangkan *TF-IDF* merupakan menentukan nilai frekuensi sebuah kata di dalam sebuah dokumen dan pemberian bobot pada setiap kata pada setiap dokumen untuk mencari dan menghitung berapa kali kata itu muncul.

Tabel 8. Hasil Pembobotan Kata

Term (t)	TF			DF	IDF	TF.IDF		
	D1	D2	D3			D1	D2	D3
bapak	2	0	0	1	$\log(3/1)=0,477$	0,954	0	0
pln	2	1	0	2	$\log(3/2)=0,176$	0,352	0,176	0
baik	2	0	0	1	$\log(3/1)=0,477$	0,954	0	0
tanya	1	0	0	1	$\log(3/1)=0,477$	0,477	0	0
sopan	1	0	0	1	$\log(3/1)=0,477$	0,477	0	0
senyum	1	0	0	1	$\log(3/1)=0,477$	0,477	0	0
hati	1	0	0	1	$\log(3/1)=0,477$	0,477	0	0
contoh	1	0	0	1	$\log(3/1)=0,477$	0,477	0	0
banyak	0	1	0	1	$\log(3/1)=0,477$	0	0,477	0
mati	0	1	0	1	$\log(3/1)=0,477$	0	0,477	0
lampu	0	1	0	1	$\log(3/1)=0,477$	0	0,477	0
kacau	0	1	0	1	$\log(3/1)=0,477$	0	0,477	0
ibu	0	0	2	1	$\log(3/1)=0,477$	0	0	0,954
kata	0	0	1	1	$\log(3/1)=0,477$	0	0	0,477



Term (t)	TF				IDF	TF.IDF		
	D1	D2	D3	DF		D1	D2	D3
jemput	0	0	1	1	log(3/1)=0,477	0	0	0,477
anak	0	0	1	1	log(3/1)=0,477	0	0	0,477
sakit	0	0	1	1	log(3/1)=0,477	0	0	0,477
girang	0	0	1	1	log(3/1)=0,477	0	0	0,477
foto	0	0	1	1	log(3/1)=0,477	0	0	0,477
geli	0	0	1	1	log(3/1)=0,477	0	0	0,477
lucu	0	0	1	1	log(3/1)=0,477	0	0	0,477
Total						4,654	2,084	4,954

Kemudian dihitung probabilitas pada setiap *term* dari semua dokumen. jumlah keseluruhan *term* yang digunakan pada perhitungan ini sebanyak 26, 11 *term* kelas positif, 15 *term* kelas negative dan 10 *term* dari kelas netral. Jumlah banyak *term* merupakan tergantung dari hasil *preprocessing* data. Berikut adalah merupakan perhitungan probabilitas pada setiap *term*:

$$P(W_k | Pos/Neg/Net) = \frac{(n_{k,Pos/Neg/Net})+1}{(n_{Pos/Neg/Net})+|kosakata|} \tag{2}$$

Diketahui:

|kosakata| = 26

Term Positif = 11

Term Negatif = 5

Term Netral = 10

Total nilai positif, negatif dan netral = 4.645 + 2.084 + 4.954 = 11.683

i. Probabilitas kata “bapak”

$$P("bapak"|"Pos") = \frac{("bapak"|"Pos") + 1}{("Pos") + |kosakata|} = \frac{0,954 + 1}{4.645 + 11.683} = 0.119$$

$$P("bapak"|"Neg") = \frac{("bapak"|"Neg") + 1}{("Neg") + |kosakata|} = \frac{0 + 1}{2.084 + 11.683} = 0.072$$

$$P("bapak"|"Net") = \frac{("bapak"|"Net") + 1}{("Net") + |kosakata|} = \frac{0 + 1}{4.954 + 11.683} = 0.060$$

ii. Probabilitas kata “pln”

$$P("pln"|"Pos") = \frac{("pln"|"Pos") + 1}{("Pos") + |kosakata|} = \frac{0,352 + 1}{4.645 + 11.683} = 0.082$$

$$P("pln"|"Neg") = \frac{("pln"|"Neg") + 1}{("Neg") + |kosakata|} = \frac{0,176 + 1}{2.084 + 11.683} = 0.085$$

$$P("pln"|"Net") = \frac{("pln"|"Net") + 1}{("Net") + |kosakata|} = \frac{0 + 1}{4.954 + 11.683} = 0.060$$

Setelah menghitung probabilitas setiap *term*, lalu hitung total probabilitas dari semua kata dengan rumus :

$$P(Pos/Neg/Net|d) = \prod_i P(a_i | Pos/Neg/Net) \tag{3}$$

Untuk kalimat pertama atau dokumen pertama :

“bapak pln baik tanya sopan senyum bapak pln baik hati contoh”

$$P(dok1|Pos) = P(bapak | Pos) \times P(pln | Pos) \times P(baik | Pos) \times P(tanya | Pos) \times P(sopan | Pos) \times P(senyum | Pos) \times P(bapak | Pos) \times P(pln | Pos) \times P(baik | Pos) \times P(hati | Pos) \times P(contoh | Pos)$$

$$P(dok1|Pos) = 0.119 * 0.082 * 0.119 * 0.090 * 0.090 * 0.090 * 0.119 * 0.082 * 0.119 * 0.090 * 0.090 = 0.00000000000796211$$

$$P(dok1|Neg) = P(bapak | Neg) \times P(pln | Neg) \times P(baik | Neg) \times P(tanya | Neg) \times P(sopan | Neg) \times P(senyum | Neg) \times P(bapak | Neg) \times P(pln | Neg) \times P(baik | Neg) \times P(hati | Neg) \times P(contoh | Neg)$$

$$P(dok1|Neg) = 0.072 * 0.085 * 0.072 * 0.072 * 0.072 * 0.072 * 0.072 * 0.085 * 0.072 * 0.072 * 0.072 = 0.000000000000375691$$

$$P(dok1|Net) = P(bapak | Net) \times P(pln | Net) \times P(baik | Net) \times P(tanya | Net) \times P(sopan | Net) \times P(senyum | Net) \times P(bapak | Net) \times P(pln | Net) \times P(baik | Net) \times P(hati | Net) \times P(contoh | Net)$$

$$P(dok1|Net) = 0.060 * 0.060 * 0.060 * 0.060 * 0.060 * 0.060 * 0.069 * 0.060 * 0.060 * 0.060 * 0.060 = 0.000000000000362797$$

Perhitungan diteruskan hingga dokumen ke 3. Sehingga hasil data perhitungan probabilitasnya dapat dimasukkan kedalam tabel dibawah ini dan digunakan untuk perhitungan *KNN*.

Tabel 9. Hasil Perhitungan Probabilitas

Dok	Positif	Negatif	Netral	Keterangan
1	0.0000000000 0796211	0.0000000000 00375691	0.0000000000 000362797	Positif
2	0.0000013536	0.0000111418	0.0000007776	Negatif
3	0.0000000000 00713343	0.0000000000 0374391	0.0000000000 492304	Netral
4	0.000012436	0.00000228428	0.0000007776	

Tahapan ini menjelaskan pengklasifikasian data komentar menggunakan data *mining* dengan menggunakan algoritma *K-Nearest Neighbor*. Berikut penjelasan dari tahapan yang dilalui :

1. Menentukan Nilai K

Untuk nilai K akan digunakan K=3. Penentuan K dilakukan oleh penulis untuk menentukan jumlah tetangga yang berdekatan yang akan diambil. Sehingga nilai K akan dilakukan pengujian untuk jumlah tetangga 3, 5, dan 7 sebagai percobaan data hasil yang paling relevan.

2. Menghitung jarak

Jarak yang akan dihitung adalah jarak dokumen yang ada pada tabel 9 diatas. Dari data diatas dibagi menjadi 3 data latih dan 1 data uji, maka perbandingannya adalah :

$$ED1 = \sqrt{\sum (0.00000000000796211 - 0.000012436)^2 + (0.000000000000375691 - 0.00000228428)^2 + (0.000000000000362797 - 0.0000007776)^2} = 0.00001266807$$

Tabel 10. Hasil Perhitungan Jarak

Dokumen	ED
1	0.00001266807
2	0.00001453031
3	0.00001266767

3. Menentukan jarak terhadap K

Pada tahap ini dilakukan mencari nilai terkecil berdasarkan K yang sudah ditentukan. Untuk K yang ditentukan adalah 3, maka cari nilai terkecil sebanyak 3 dari nilai ED :

Tabel 11. Hasil Nilai terkecil dari ED

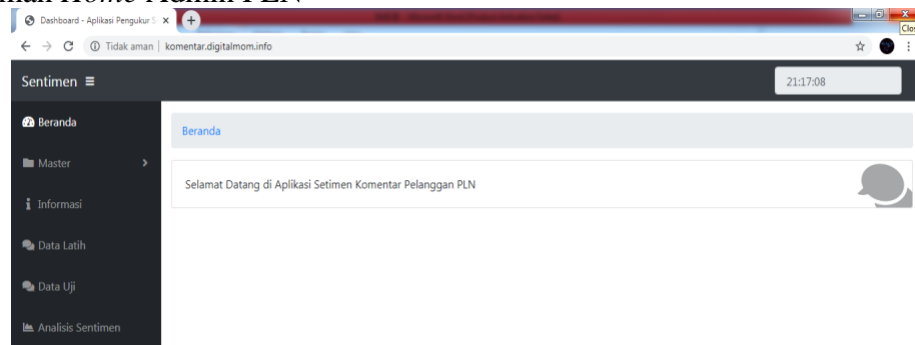
Dokumen	ED	Keterangan
1	0.00001266807	Positif
2	0.00001453031	Negatif
3	0.00001266767	Netral

Dari hasil pencarian 3 nilai terkecil dihasilkan keterangan positif, negatif dan netral.

### III.2 Implementasi Sistem

Setelah melakukan analisa dan perancangan, maka tahap selanjutnya adalah implementasi dengan menggunakan perangkat lunak. Berikut tampilan dari beberapa halaman system yang telah dirancang:

a. Halaman *Home Admin PLN*



Gambar 2. Halaman *Home Admin PLN*

b. Halaman Data Komentar Pelanggan

No	Komentator	Komentar	Kategori	Preprocessing
1	Fitri Ynedi	Kubang sudah padam dari jam 7 pagi bos kebangetan	Negatif	Kubang padam jam pagi
2	Latif Mutiar	Lampu mati sebagian dikarenakan gangguan cuaca sudah 6 jam PLN tidak merespon sudah 3org menghubungi PLN masih tidur	Negatif	lampu mati gangguan cuaca jam orang hubung tidur
3	maman salahudin	Alhamdulillah ini yang ditunggu-tunggu masyarakat ladan palmatask smoga proses sdg operasi lancar dan inshallah barokah amin yra	Positif	tunggu tunggu masyarakat ladan proses operasi lancar aman tuhan
4	rio aji	Mantap pt pln wilayah propinsi riau kami apesiasi bantuannya mantap dan sukses pln	Positif	mantap wilayah mantap sukses
5	Jannah putri	Mohon bantuannya dan solusinya Meteran dirumah saya prabayar token elektrik permasalahannya setiap magrib lampunya mati aliran listrik terputus sementara lampu dimeteran hidup lampu diruma tetangga nyala apa solusinya merek meteran fed	Negatif	mohon solusi meter rumah tiap magrib lampu mati alir listrik putus lampu meter hidup lampu tetangga nyala solusinya merek meter
6	Antarctic Osram	Pak PLN tadi baik kalau tanya sopan dan semoy, itu baru pak PLN yang baik hati dan bisa dicontoh	Positif	
7	Kiss Kiss Scene	PLN banyak mati lampunya kacau	Negatif	
8	Yulla Sarah	Adu ini ibu ibu hanya katanya buru2 mau jemput anaknya yang sakit malah kegrangan foto2 dulu, gell tapi ngakak	Netral	
9	Norman Kamarun	Bapak-bapak dari pln tadi baik sopan	?	

Gambar 3. Data Komentar Pelanggan

Beberapa komentar pelanggan PLN yang masuk ke aplikasi.

c. Halaman Hasil Pengolahan Data

1. Case Folding		
No	Komentar	Hasil Case Folding
1	Pak PLN tadi baik kalau tanya sopan dan senyum, itu baru pak PLN yang baik hati dan bisa dicontoh	pak pln tadi baik kalau tanya sopan dan senyum, itu baru pak pln yang baik hati dan bisa dicontoh
2	PLN banyak mati lampunya kacau	pln banyak mati lampunya kacau
3	Adu ini Ibu hanya katanya buru2 mau jemput anaknya yang sakit malah kegrangan foto2 dulu, geli tapi ngakak	adu ini Ibu hanya katanya buru2 mau jemput anaknya yang sakit malah kegrangan foto2 dulu, geli tapi ngakak

2. Cleaning		
No	Komentar	Hasil Cleaning
1	pak pln tadi baik kalau tanya sopan dan senyum, itu baru pak pln yang baik hati dan bisa dicontoh	pak pln tadi baik kalau tanya sopan dan senyum itu baru pak pln yang baik hati dan bisa dicontoh
2	pln banyak mati lampunya kacau	pln banyak mati lampunya kacau
3	adu ini Ibu hanya katanya buru2 mau jemput anaknya yang sakit malah kegrangan foto2 dulu, geli tapi ngakak	adu ini Ibu hanya katanya buru2 mau jemput anaknya yang sakit malah kegrangan foto2 dulu, geli tapi ngakak

Gambar 4. Hasil Pengolahan Data

Gambar diatas merupakan proses pengolahan data komentar dengan *case folding* dan *cleaning*. Setelah itu dilakukan tahap *tokenizing*, normalisasi, sampai dengan mendapatkan nilai probabilitas dari masing-masing dokumen. Berikut nilai probabilitas dari dokumen latihan dan uji:

Hasil Perhitungan Probabilitas				
Dok	Positif	Negatif	Netral	Keterangan
1	0.00000000001794211	0.00000000000375691	0.000000000000362797	Positif
2	0.0000013536	0.0000111418	0.0000007776	Negatif
3	0.000000000000713343	0.00000000000374391	0.00000000000492304	Netral
4	0.000012436	0.00000228428	0.0000007776	

Gambar 5. Hasil Pengolahan Data Komentar

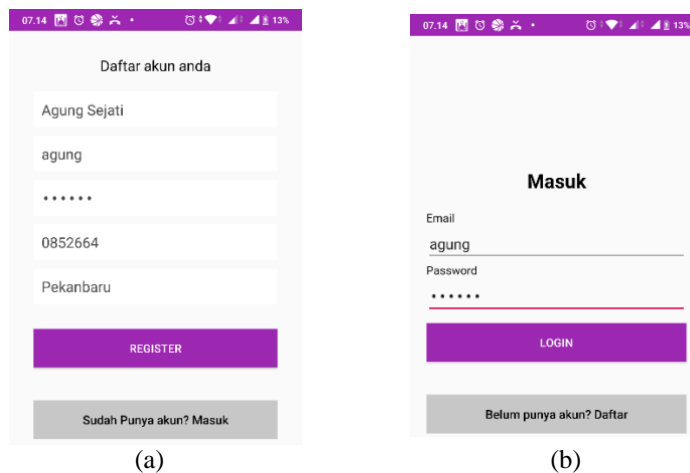
Dari tiga dokumen, dokumen pertama bernilai positif, dokumen kedua bernilai negatif dan dokumen ketiga bernilai netral, lalu dicari nilai terkecilnya berdasarkan K yang sudah ditentukan.

Hasil Nilai Terkecil dari ED		
Dokumen	ED	Keterangan
1	0.00001266807	Positif
2	0.00001453031	Negatif
3	0.00001266767	Netral

Gambar 6. Hasil Nilai ED terkecil

Dari hasil pencariann nilai terkecilnya, sehingga dokumen ke empat hasilnya netral.

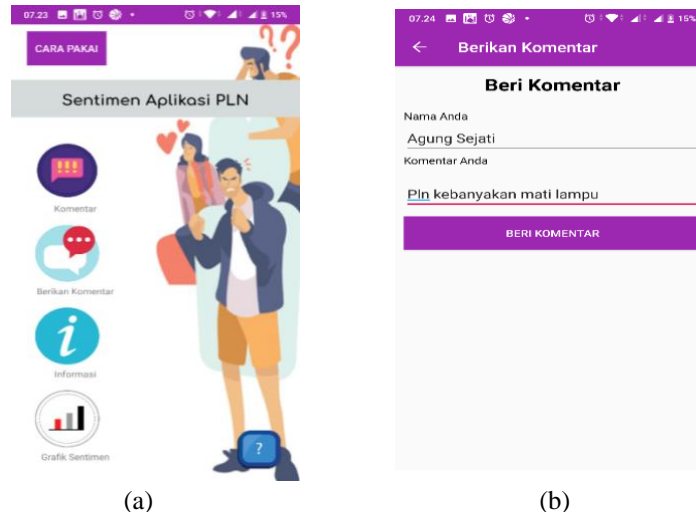
d. Tampilan *Form Register* Pelanggan dan *Login* Pelanggan



Gambar 5. (a) Tampilan *Form Register* Pelanggan (b) Tampilan *Login* Pelanggan

Tampilan gambar (a) diatas merupakan tampilan *form register* pelanggan. Setiap pelanggan yang akan memberikan komentar atau saran terhadap PLN harus melakukan pendaftaran di aplikasi. Tampilan gambar (b) diatas adalah halaman *Login* Pelanggan. Pelanggan yang sudah register kan memiliki akun yang digunakan untuk masuk ke aplikasi.

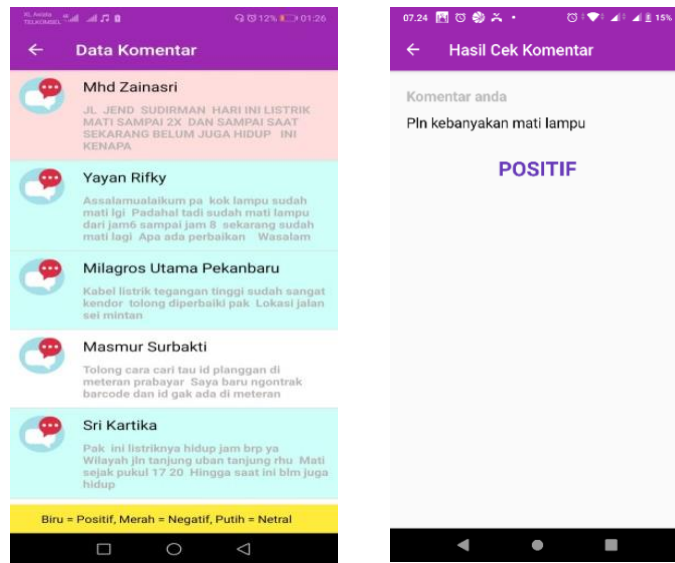
e. Tampilan *Home* Pelanggan dan Tampilan Beri Komentar Pelanggan



Gambar 6. (a) Tampilan *Home* Pelanggan (b) Tampilan Beri Komentar Pelanggan

Tampilan gambar (a) diatas adalah tampilan *home* pelanggan dan tampilan gambar (b) adalah menu pelanggan dalam memberi komentar. Pada tampilan home, pelanggan bisa memasukkan komentar, melihat komentar pelanggan lain, dan melihat beberapa menu lainnya. Pada tampilan komenar, pelanggan boleh memberikan komentar apa saja dengan menginputkan nama lalu komentar dan saran terhadap PLN.

## f. Tampilan Komentar Pelanggan dan Hasil Cek Komentar Pelanggan



(a) (b)  
Gambar 6. (a) Tampilan Data Komentar Pelanggan  
(b) Tampilan Hasil Cek Komentar Pelanggan

Tampilan gambar (a) diatas adalah tampilan data komentar yang diberikan oleh pelanggan dan tampilan gambar (b) adalah hasil cek komentar pelanggan. Komentar yang diberikan pelanggan akan diklasifikasikan dalam 3 kategori, kategori positif, negatif dan netral. Tampilan diatas contoh komentar yang positif.

#### IV. KESIMPULAN

Beberapa kesimpulan yang dapat diambil dari pembangunan aplikasi sentimen komentar pelanggan PLN, diantaranya adalah sebagai berikut:

1. Secara umum sistem yang dibangun telah bisa berjalan dan dapat menghasilkan data yang sesuai dengan proses input komentar dan hasil komentar.
2. Hasil penggunaan metode *KNN* diterapkan pada klasifikasi komentar dapat memberikan hasil positif, negatif dan netral.
3. Grafik klasifikasi komentar dapat ditampilkan pada aplikasi dengan klasifikasi komentar positif, negatif dan netral.

#### REFERENS

- [1] Sani, R. R., Zeniarza, J., & Luthfiarta, A. (2016). *Pengembangan Aplikasi Penentuan Tema Tugas Akhir Berdasarkan Data Abstrak Menggunakan Algoritma K-Nearest Neighbor*. Seminar Nasional Multi Disiplin Ilmu, 103-111.
- [2] Sipayung, Evasaria M. dkk. (2016). *Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier*. Jurnal Sistem Informasi Vol. 8 No.1. ISSN : 2085-1588.
- [3] Hanggara, Surya. Dkk (2017). *Aplikasi Web Untuk Analisis Sentimen Pada Opini Produk Dengan Metode Naive Bayes Classifier*. Seminar Nasional Inovasi dan Aplikasi Teknologi di Industri. ITN Malang.

- [4] Mariana, Novita, Rara Sriartati Redjeki, Jeffri Alfa Razaq.2015. Penerapan Algoritma K-NN (*Nearest Neighbor*) untuk Deteksi Penyakit (Kanker Serviks). *Dinamika Informatika*. Vol 7 No 1. ISSN 2085-3343
- [5] Kurniawan, Denni, dan Ade Saputra. 2019. Penerapan K-Nearest Neighbour Dalam Penerimaan Peserta Didik Dengan Sistem Zonasi. *Jurnal Sistem Informasi Bisnis* VOL. 2 No. 9 ISSN: 2502-2377
- [6] Nanja, Muis., Purwanto. 2015. Metode *K-Nearest Neighbour* Berbasis *Forward Selection* untuk Prediksi Harga Komoditi Lada. *Jurnal Pseudocode* Vol. 2 No. 1 ISSN:2355-5920
- [7] Yustanti, Wiyli. 2012. Algoritma *K-Nearest Neighbour* untuk Memprediksi Harga Jual Tanah. *Jurnal Matematika, Statistika dan Komputasi* Vol.9 No.1 ISSN: 2614-8811
- [8] T, Asahar Johar, Delfi Yanosma, dan Kurnia Anggriani. 2016. Implementasi Metode *K-Nearest Neighbor* (KNN) dan *Simple Additive Weighting* (Saw) dalam Pengambilan Keputusan Seleksi Penerimaan Anggota Paskibraka. *Jurnal Pseudocode* Vol. 3 No. 2, ISSN:2355-5920
- [9] Melita, Ria, Victor Amrizal, Hendra Bayu Suseno, Taslimun Dirjam. Penerapan Metode *Term Frequency Inverse Document Frequency* (Tf-Idf) Dan *Cosine Similarity* Pada Sistem Temu Kembali Informasi Untuk Mengetahui *Syarah Hadits* Berbasis Web (Studi Kasus: *Syarah Umdatil Ahkam*). *Jurnal Teknik Informatika* Vol 11 No. 2, Oktober 2018
- [10] Ma'arif, Abdul Aziz. 2015. Penerapan Algoritma Tf-Idf Untuk Pencarian Karya Ilmiah. *Jurnal Jurusan Teknik Informatika*.
- [11] Maulina, Dina, Dan Rofie Sagara. Klasifikasi Artikel Hoax Menggunakan Support Vector Machine Linear Dengan Pembobotan Term Frequency – Inverse Document Frequency. *Jurnal Mantik Penusa*. Volume 2 No. 1 Juni 2018

#### UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada PLN sebagai sumber data penelitian dan semua pihak yang telah membantu hingga selesainya penelitian ini.