

**SISTEM KLASIFIKASI KELUHAN PELANGGAN PT PLN  
SEMARANG MENGGUNAKAN ALGORITMA NAÏVE BAYES  
*CLASSIFIER***



**SKRIPSI**

**Disusun Sebagai Salah Satu Syarat  
untuk Memperoleh Gelar Sarjana Komputer  
pada Departemen Ilmu Komputer/ Informatika**

**Disusun oleh :**

**FARID MAHMUD ROHMATULLAH**

**24010314120063**

**DEPARTEMEN ILMU KOMPUTER/INFORMATIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO**

**2018**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini :

Nama : Farid Mahmud Rohmatullah

NIM : 24010314120063

Judul : Sistem Klasifikasi Keluhan Pelanggan PT PLN Semarang Menggunakan  
Algoritma Naïve Bayes *Classifier*

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang sepengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Semarang, 10 Agustus 2018



Farid Mahmud Rohmatullah

24010314120063

## HALAMAN PENGESAHAN

Judul : Sistem Klasifikasi Keluhan Pelanggan PT PLN Semarang Menggunakan  
Algoritma Naïve Bayes *Classifier*

Nama : Farid Mahmud Rohmatullah

NIM : 24010314120063

Telah diujikan pada sidang skripsi tanggal 24 Juli 2018 dan dinyatakan lulus pada tanggal  
24 Juli 2018.

Semarang, 10 Agustus 2018

Mengetahui,

a.n. Ketua Departemen Ilmu Komputer/Informatika

Sekretaris,

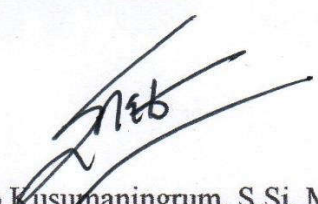


Dr. Eng. Adi Wibowo, S.Si., M.Kom

NIP. 198203092006041002

Panitia Penguji Skripsi

Ketua,

  
Dr. Retno Kusumaningrum, S.Si, M.Kom

NIP. 198104202005012001

## HALAMAN PENGESAHAN

Judul : Sistem Klasifikasi Keluhan Pelanggan PT PLN Semarang Menggunakan  
Algoritma Naïve Bayes Classifier

Nama : Farid Mahmud Rohmatullah

NIM : 24010314120012

Telah diujikan pada sidang skripsi tanggal 24 Juli 2018.

Semarang, 10 Agustus 2018

Pembimbing



Sutikno, M.Cs

NIP. 197905242009121003

## ABSTRAK

PT PLN (Persero) adalah sebuah BUMN yang mengurus semua aspek kelistrikan yang ada di Indonesia. Sebagai perusahaan penyedia layanan listrik terbesar di Indonesia, PT PLN tentunya menghadapi banyak tantangan salah satunya yaitu banyaknya jumlah pengaduan dan keluhan dari pelanggan seiring dengan semakin banyaknya pengguna listrik di Indonesia. Agar dapat merespon keluhan pelanggan secara cepat, petugas perlu melakukan klasifikasi keluhan untuk mengetahui jenis keluhan yang masuk agar dapat segera ditindak lanjuti. Pengklasifikasian keluhan pelanggan yang masih dilakukan secara manual tentunya akan mempengaruhi lama proses respon keluhan pelanggan. Penelitian ini bertujuan untuk menjawab permasalahan tersebut dengan membangun sistem yang dapat mengelompokkan keluhan-keluhan pelanggan berdasarkan jenisnya. Sistem ini menggunakan algoritma Naïve Bayes *Classifier* untuk mengelompokkan keluhan-keluhan pelanggan tersebut. Naive Bayes *Classifier* merupakan algoritma pengklasifikasian dengan metode probabilitas dan statistik. Hasil dari beberapa eksperimen pada penelitian ini menunjukkan bahwa penggunaan pembobotan frekuensi kemunculan kata/TF *Raw* pada algoritma Naïve Bayes *Classifier* menghasilkan akurasi sebesar 97.163%, sedangkan penggunaan pembobotan TF-IDF pada algoritma Naïve Bayes *Classifier* menghasilkan akurasi sebesar 95.763%. Hasil tersebut didapat dengan menggunakan data sebanyak 3000 data keluhan pelanggan dari *call center* 123 PT PLN Semarang.

Kata Kunci : Klasifikasi, Keluhan Pelanggan, Naïve Bayes *Classifier*

## ABSTRACT

PT PLN (Persero) is a SOEs/BUMN that takes care of all aspects of electricity in Indonesia. As the largest electricity service provider in Indonesia, PT PLN certainly faces many challenges, one of which is the number of complaints from customers along with the increasing number of electricity users in Indonesia. In order to respond quickly the customer complaints, the officer needs to classify the complaints to find out the types of complaints to be followed up immediately. The classification of customer complaints that are still done manually will certainly affect the duration of customer complaints response. This research aims to answer the problem by building a system that can classify customer complaints by the type. This system uses the Naive Bayes algorithm to classify the customer complaints. Naive Bayes Classifier is a classification algorithm with probability and statistical methods. The results of the several experiment in this research indicate that the use of word occurrence/TF Raw weighting in Naive Bayes Classifier algorithm produces accuracy of 97.163%, while the use of TF-IDF weighting in Naive Bayes Classifier algorithm produces accuracy of 95.763%. The results obtained by using data as much as 3000 customer complaints data from call center 123 PT PLN Semarang.

Keywords : Classification, Customer Complaints, Naïve Bayes *Classifier*

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Sistem Klasifikasi Keluhan Pelanggan PT PLN Semarang Menggunakan Algoritma Naïve Bayes *Classifier*”.

Skripsi ini dibuat dengan tujuan sebagai salah satu syarat untuk memperoleh gelar sarjana komputer pada Departemen Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.

Dalam pelaksanaan skripsi serta penyusunan dokumen skripsi ini, penulis menyadari banyak pihak yang membantu sehingga akhirnya dokumen ini dapat diselesaikan. Oleh karena itu, melalui kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada :

1. Ibu Dr. Retno Kusumaningrum, S.Si, M.Kom selaku Ketua Departemen Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.
2. Bapak Helmie Arif Wibawa, S.Si, M.Cs, selaku Koordinator Skripsi Departemen Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.
3. Bapak Sutikno, S.T, M.Cs selaku dosen pembimbing skripsi yang telah membantu dalam membimbing dan mengarahkan penulis hingga selesainya skripsi ini.
4. Semua pihak yang telah membantu kelancaran dalam penyusunan skripsi, yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa dokumen skripsi ini masih jauh dari sempurna. Oleh karena itu, saran dan kritik yang membangun sangat penulis harapkan. Akhir kata, semoga skripsi ini dapat bermanfaat bagi semua pihak.

Semarang, 10 Agustus 2018

Penulis

## DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	ii
HALAMAN PENGESAHAN .....	iii
HALAMAN PENGESAHAN .....	iv
ABSTRAK .....	v
ABSTRACT .....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI .....	viii
DAFTAR GAMBAR.....	x
DAFTAR TABEL .....	xii
DAFTAR LAMPIRAN .....	xiv
BAB I PENDAHULUAN .....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Tujuan dan Manfaat .....	3
1.4. Ruang Lingkup .....	3
1.5. Sistematika Penulisan .....	4
BAB II TINJAUAN PUSTAKA .....	5
2.1. <i>Text Mining</i> .....	5
2.2. <i>Preprocessing</i> pada <i>Text Mining</i> .....	6
2.2.1 <i>Case Folding</i> .....	6
2.2.2 <i>Tokenizing</i> .....	7
2.2.3 <i>Filtering</i> .....	7
2.2.4 <i>Stemming</i> .....	8
2.3. Pembobotan Kata.....	10
2.4. <i>K-Fold Cross Validation</i> .....	11
2.5. Klasifikasi <i>Data Mining</i> .....	11
2.6. Algoritma Naïve Bayes.....	12
2.7. Evaluasi.....	13
2.8. Ekstraksi Informasi.....	15
2.9. <i>Flowchart</i> .....	15
2.10. Model Proses <i>Waterfall</i> .....	16



BAB III METODOLOGI PENELITIAN .....	19
3.1. Pengumpulan Data .....	20
3.2. <i>Preprocessing</i> Data .....	20
3.2.1 <i>Case Folding</i> .....	21
3.2.2 <i>Tokenizing</i> .....	22
3.2.3 <i>Filtering</i> .....	23
3.2.4 <i>Stemming</i> .....	25
3.3. <i>K-Fold Cross Validation</i> .....	31
3.4. Pembobotan kata .....	32
3.5. <i>Training</i> Data .....	34
3.6. <i>Testing</i> Data .....	36
3.7. Evaluasi .....	38
3.8. Ekstraksi Tanggal dan Tempat .....	42
3.9. Analisis dan Perancangan Sistem .....	42
3.9.1. Analisis Sistem .....	42
3.9.2. Perancangan Sistem .....	43
BAB IV HASIL DAN IMPLEMENTASI .....	55
4.1. Hasil Pembangunan Sistem .....	55
4.1.1. Lingkungan Implementasi .....	55
4.1.2. Implementasi Data .....	56
4.1.3. Implementasi Proses .....	58
4.1.4. Implementasi Antarmuka .....	62
4.2. Pengujian Sistem .....	66
4.2.1. Pengujian Fungsional .....	66
4.2.2. Pengujian Kinerja .....	67
BAB V PENUTUP .....	71
5.1. Kesimpulan .....	71
5.2. Saran .....	71
DAFTAR PUSTAKA .....	73

## DAFTAR GAMBAR

Gambar 2.1 <i>Flow Diagram Text Categorization</i> (Kumar & Karthika, 2014).....	6
Gambar 2.2 Model <i>Waterfall</i> (Sommerville, 2011) .....	17
Gambar 3.1 Gambaran Umum Penelitian .....	19
Gambar 3.2 <i>Flowchart Preprocessing</i> .....	21
Gambar 3.3 <i>Flowchart Proses Tokenizing</i> .....	22
Gambar 3.4 <i>Flowchart Proses Filtering</i> .....	24
Gambar 3.5 <i>Flowchart Proses Stemming</i> .....	26
Gambar 3.6 Sub proses <i>stemming</i> Sastrawi.....	27
Gambar 3.7 Sub proses <i>stemming plural</i> .....	27
Gambar 3.8 Sub proses <i>stemming singular</i> .....	29
Gambar 3.9 Sub proses loopPengembalianAkhir.....	30
Gambar 3.10 <i>Flowchart</i> proses pelatihan.....	34
Gambar 3.11 <i>Flowchart</i> proses pengujian.....	37
Gambar 3.12 <i>Entity Relational Diagram</i> .....	44
Gambar 3.13 <i>Data Context Diagram</i> (DCD) .....	45
Gambar 3.14 <i>Data Flow Diagram</i> (DFD) level 1 .....	46
Gambar 3.15 DFD level 2 Proses Mengelola Data Keluhan.....	47
Gambar 3.16 DFD Level 2 Proses Klasifikasi Data.....	48
Gambar 3.17 Antarmuka Halaman <i>Dashboard</i> .....	51
Gambar 3.18 Antarmuka Halaman Data <i>Training</i> .....	52
Gambar 3.19 Antarmuka Halaman <i>Add Data</i> .....	53
Gambar 3.20 Antarmuka Halaman <i>Edit Data</i> .....	53
Gambar 3.21 Antarmuka Halaman Klasifikasi.....	54
Gambar 3.22 Antarmuka Halaman Hasil Klasifikasi .....	54
Gambar 4.1 Implementasi Tabel Data Keluhan .....	56
Gambar 4.2 Implementasi Tabel Kelas .....	56
Gambar 4.3 Implementasi Tabel <i>Vocabulary</i> .....	57
Gambar 4.4 Implementasi Tabel <i>Conditional Probability</i> .....	57
Gambar 4.5 Implementasi Tabel <i>K-Fold</i> .....	57
Gambar 4.6 Implementasi Antarmuka Halaman <i>Dashboard</i> .....	62
Gambar 4.7 Implementasi Antarmuka Halaman Data <i>Training</i> .....	63

Gambar 4.8 Implementasi Antarmuka <i>Add Data</i> .....	64
Gambar 4.9 Implementasi Antarmuka <i>Edit Data</i> .....	64
Gambar 4.10 Implementasi Antarmuka Halaman Klasifikasi.....	65
Gambar 4.11 Implementasi Antarmuka Hasil Klasifikasi.....	66
Gambar 4.12 Grafik Perbandingan Jumlah Keluhan per Bulan .....	69
Gambar 4.13 Grafik Perbandingan Jumlah Keluhan per Tempat .....	70

## DAFTAR TABEL

Tabel 2.1 Contoh proses <i>Tokenizing</i> .....	7
Tabel 2.2 Contoh proses <i>filtering</i> .....	8
Tabel 2.3 Tabel Variasi <i>Prefix</i> .....	9
Tabel 2.4 <i>Confusion Matrix</i> .....	14
Tabel 2.5 Simbol <i>Flowchart</i> .....	16
Tabel 3.1 Data Keluhan Sebelum Proses <i>Case Folding</i> .....	21
Tabel 3.2 Data Keluhan Setelah Proses <i>Case Folding</i> .....	22
Tabel 3.3 Data Keluhan Sebelum Proses <i>Tokenizing</i> .....	23
Tabel 3.4 Data Keluhan Setelah Proses <i>Tokenizing</i> .....	23
Tabel 3.5 Data Array Kata Sebelum Proses <i>filtering</i> .....	25
Tabel 3.6 Data Array Kata Setelah Proses <i>Filtering</i> .....	25
Tabel 3.7 Data Array Kata Hasil Proses <i>Filtering</i> .....	31
Tabel 3.8 Data Array Kata Hasil Proses <i>Stemming</i> .....	31
Tabel 3.9 Data hasil <i>preprocessing</i> .....	32
Tabel 3.10 Hasil pembobotan TF <i>raw</i> .....	33
Tabel 3.11 Hasil pembobotan TF-IDF .....	33
Tabel 3.12 Contoh <i>Vocabulary</i> .....	35
Tabel 3.13 Ekstraksi kata berdasarkan <i>vocabulary</i> .....	36
Tabel 3.14 Hasil perhitungan <i>posterior</i> tiap kelas.....	37
Tabel 3.15 Contoh Data Prediksi dan Data Asli/Aktual.....	38
Tabel 3.16 Contoh <i>Confusion Matrix</i> .....	41
Tabel 3.17 Ekstraksi Informasi Tanggal dan Tempat.....	42
Tabel 3.18 Kebutuhan Fungsional Sistem.....	43
Tabel 3.19 Kebutuhan Non-Fungsional Sistem.....	43
Tabel 3.20 Tabel Data Keluhan.....	49
Tabel 3.21 Tabel Kelas.....	49
Tabel 3.22 Tabel <i>Vocabulary</i> .....	50
Tabel 3.23 Tabel <i>Conditional Probability</i> .....	50
Tabel 3.24 Tabel <i>K-Fold</i> .....	50
Tabel 4.1 Skenario Pengujian <i>Black Box</i> .....	67

Tabel 4.2 Hasil Pengujian Kinerja Menggunakan <i>K-Fold Cross Validation</i> dan Pembobotan Frekuensi Kata .....	68
Tabel 4.3 Hasil Pengujian Kinerja menggunakan <i>K-Fold Cross Validation</i> dan Pembobotan TF-IDF .....	68

## DAFTAR LAMPIRAN

Lampiran 1. Perhitungan Naïve Bayes .....	77
Lampiran 2. Hasil Pengujian <i>Black Box</i> .....	87
Lampiran 3. <i>Sample</i> 135 Data Keluhan Pelanggan PT PLN Semarang .....	89

# **BAB I**

## **PENDAHULUAN**

Bab ini membahas latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan dalam pembuatan skripsi mengenai sistem klasifikasi keluhan pelanggan PT PLN Semarang menggunakan algoritma *Naïve Bayes Classifier*.

### **1.1. Latar Belakang**

Perusahaan Listrik Negara (PLN) atau nama resminya adalah PT PLN (Persero), adalah sebuah BUMN yang mengurus semua aspek kelistrikan yang ada di Indonesia. PT PLN bertujuan untuk menyelenggarakan usaha penyediaan tenaga listrik bagi kepentingan umum dalam jumlah dan mutu yang memadai serta memupuk keuntungan dan melaksanakan penugasan Pemerintah di bidang ketenagalistrikan dalam rangka menunjang pembangunan dengan menerapkan prinsip-prinsip Perseroan Terbatas (PLN, 2018).

Sebagai perusahaan penyedia layanan listrik terbesar di Indonesia, PT PLN tentunya mempunyai banyak tantangan salah satunya yaitu jumlah pengaduan dan keluhan pelanggan yang semakin bertambah seiring dengan meningkatnya jumlah pengguna listrik di Indonesia. Pengaduan dan keluhan pelanggan pada PT PLN Semarang dapat melalui beberapa media seperti telepon, *facebook*, *twitter*, serta melalui *email*. Agar dapat merespon keluhan pelanggan dengan cepat, petugas perlu mengetahui jenis dari keluhan yang masuk supaya dapat ditindak lanjuti. Data keluhan pelanggan melalui telepon/*call center* 123 merupakan keluhan pelanggan yang paling banyak dengan rata-rata keluhan mencapai 1000 keluhan per bulannya untuk wilayah Semarang. Oleh karena itu, dibutuhkan sebuah sistem klasifikasi yang dapat memisahkan keluhan-keluhan yang masuk berdasarkan kelas atau jenisnya. Terdapat sembilan jenis keluhan pelanggan pada PT PLN yaitu Tusbung(Putus Sambungan), Instalasi Listrik, Mutu dan Keandalan, PDPB(Perubahan Daya Pemasangan Baru), Prabayar, Cater(Pencatatan Meteran), Lain-Lain, Rekening, serta Pemakaian PTL Tidak Sah.

Dari data keluhan pelanggan yang telah masuk juga dapat dimanfaatkan untuk mendapatkan informasi mengenai tanggal serta tempat kejadian. Dimana dari hasil klasifikasi serta hasil ekstraksi data tanggal dan tempat keluhan, dapat diketahui keluhan dengan kelas mana yang paling sering muncul di tanggal dan tempat tertentu.

Keluhan pelanggan yang masuk melalui telepon kemudian akan dicatat oleh *customer service* menjadi keluhan dalam bentuk *text digital* yang tidak terstruktur. Salah satu cara yang dapat digunakan untuk mengolah data dalam bentuk *text* yaitu dengan analisis *text mining*. *Text mining* bertujuan untuk memudahkan pengguna mengekstrak informasi yang membangun dari sumber tekstual. Klasifikasi teks adalah proses pengkategorian dokumen ke dalam kelas yang telah ditentukan berdasarkan isinya (Rajeswari, et al., 2017). Klasifikasi teks sendiri memiliki beberapa cara pendekatan yang berbeda antara lain, *Naïve Bayes*, *Support Vector Machine (SVM)*, dan *Artificial Neural Network (ANN)*, atau *Decision Tree (DT)* (Saptono, et al., 2016).

Penelitian mengenai klasifikasi keluhan pelanggan pernah dilakukan oleh Eni Irfiani (2014) dan menghasilkan akurasi hingga 80%. Algoritma *Naïve Bayes* untuk klasifikasi teks telah dilakukan oleh beberapa peneliti sebelumnya. Penelitian mengenai *Naïve Bayes* pernah dilakukan oleh Rajeswari, et al (2017), pada penelitian tersebut menunjukkan bahwa *Naïve Bayes* lebih unggul dengan akurasi 66.67% dibandingkan KNN dengan akurasi 38.89% (Rajeswari, et al., 2017). Ting, et al (2011) melakukan perbandingan antara *Naïve Bayes*, SVM, NN, dan DT, pada penelitian tersebut didapatkan akurasi tertinggi yaitu *Naïve Bayes* dengan akurasi mencapai 97%. Pada penelitian tersebut juga diketahui bahwa *Naïve Bayes* membutuhkan waktu lebih sedikit dalam membangun model dari pada SVM dan DT (Ting, et al., 2011).

Pada penelitian dengan teks bahasa Indonesia menggunakan algoritma *Naïve Bayes* juga pernah dilakukan oleh Saptono, et al (2016) dan menghasilkan akurasi sebesar 87%. Darujati & Gumelar (2012) melakukan penelitian dengan menggunakan data latih lebih dari 150 dokumen bahasa Indonesia dengan algoritma *Naïve Bayes* menghasilkan akurasi hingga 90%. Pada penelitian tersebut diketahui bahwa semakin banyak data latih maka akurasi pada algoritma *Naïve Bayes* akan semakin baik (Darujati & Gumelar, 2012). Dari beberapa penelitian diatas diketahui bahwa algoritma *Naïve Bayes* banyak digunakan karena terbukti efektif untuk klasifikasi teks, sederhana, cepat, serta mempunyai akurasi yang cukup baik. Pada penelitian kali ini



menggunakan algoritma Naïve Bayes yang merupakan algoritma dengan pendekatan probabilistik untuk melakukan klasifikasi.

Berdasarkan uraian diatas, maka akan dibangun sistem klasifikasi keluhan pelanggan pada PT PLN Semarang dengan menggunakan algoritma Naïve Bayes. Sistem ini diharapkan mampu memudahkan PT PLN Semarang dalam menentukan jenis keluhan yang masuk.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, dapat dirumuskan suatu permasalahan yaitu bagaimana membuat sistem yang dapat mengklasifikasikan keluhan pelanggan yang masuk ke dalam beberapa kelas yang telah ditentukan dengan menggunakan algoritma Naïve Bayes *Classifier* serta mendapatkan informasi mengenai tanggal dan tempat kejadian dari keluhan pelanggan tersebut.

## **1.3. Tujuan dan Manfaat**

Tujuan yang ingin dicapai pada penelitian ini yaitu untuk menghasilkan sebuah sistem yang dapat melakukan klasifikasi terhadap keluhan pelanggan dengan menggunakan algoritma Naïve Bayes *Classifier* serta dapat mengetahui informasi tanggal dan tempat kejadian dari keluhan pelanggan.

Adapun manfaat dari penelitian ini adalah sistem yang dikembangkan dapat memudahkan pihak PLN dalam menentukan jenis keluhan-keluhan yang masuk sehingga keluhan dapat segera ditanggapi berdasarkan jenis keluhan masing-masing. Pihak PLN juga mendapatkan informasi mengenai tanggal dan tempat kejadian dari keluhan pelanggan tersebut.

## **1.4. Ruang Lingkup**

Ruang lingkup sistem klasifikasi keluhan pelanggan pada PT PLN menggunakan algoritma Naïve Bayes *Classifier* adalah:

1. Sistem ini dibuat dengan menggunakan metode Naïve Bayes *Classifier* dengan pembobotan frekuensi kemunculan kata dan TF-IDF untuk melakukan klasifikasi terhadap keluhan pelanggan.
2. Pengembangan sistem menggunakan model proses *Waterfall* sampai dengan tahap *Integration and system testing*.

3. Data yang digunakan yaitu data keluhan pelanggan pada bulan Oktober, November, dan Desember 2017 sebanyak 3000 data keluhan dari *call center* 123 PT PLN Semarang.
4. Ekstraksi informasi tanggal dan tempat belum dapat mengatasi jika terjadi kesalahan pengetikan atau *typo* dari pengguna.

### 1.5. Sistematika Penulisan

Sistematika penulisan yang digunakan dalam skripsi ini terbagi dalam beberapa pokok bahasan, yaitu :

#### BAB I PENDAHULUAN

Bab ini membahas latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan dari sistem klasifikasi keluhan pelanggan PT PLN Semarang menggunakan algoritma *Naïve Bayes Classifier*.

#### BAB II LANDASAN TEORI

Bab ini menyajikan hasil studi pustaka mengenai teori yang berhubungan dengan pelaksanaan dan penyusunan skripsi seperti *text mining*, *preprocessing* pada *text mining*, pembobotan kata yang digunakan, *K-Fold cross validation*, serta algoritma *Naïve Bayes*.

#### BAB III METODOLOGI PENELITIAN

Bab ini menyajikan tahapan penyelesaian skripsi yang diawali dengan pengumpulan data, pengolahan data/*preprocessing* data, *training* dan *testing* data, dan terakhir yaitu evaluasi terhadap sistem dan kinerja algoritma yang digunakan.

#### BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan tentang implementasi dari hasil analisis dan perancangan sistem serta menjelaskan beberapa eksperimen yang dilakukan dan hasil tiap eksperimen.

#### BAB V PENUTUP

Bab ini membahas mengenai kesimpulan dari uraian yang telah dijabarkan pada bab-bab sebelumnya dan saran untuk pengembangan penelitian lebih lanjut.