

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Badan Pusat Statistik (BPS) Kota Bandung menyatakan bahwa jumlah penduduk di kota Bandung pada tahun 2016 mencapai 2,490,622 penduduk [1]. Peningkatan jumlah penduduk di Kota Bandung menyebabkan banyaknya permasalahan yang muncul, salah satu permasalahan yang ada diantaranya pelayanan aspirasi warga atau pengaduan warga terhadap kinerja pemerintahan dan pelayanan terhadap kesejahteraan masyarakat Kota Bandung sendiri. Menyadari hal tersebut, Pemkot Bandung meluncurkan fasilitas Layanan Aspirasi dan Pengaduan *Online* Rakyat (LAPOR!), sebagai wadah menampung keluhan dan aspirasi dari masyarakat. LAPOR! merupakan berupa aplikasi media sosial yang melibatkan fasilitas publik dengan sistem yang bersifat dua arah, untuk menciptakan mekanisme aspirasi dan pengaduan dari masyarakat yang nantinya didisposisikan ke berbagai instansi atau lembaga terkait untuk direspons [2].

Jumlah LAPOR! dalam setahun Pemkot Bandung mendapat 5,571 laporan. Jumlah aduan terbanyak disampaikan melalui pesan singkat SMS yaitu sebanyak 67% dan sisanya 33% disampaikan melalui laman *website* [3]. Hasil dari berbagai aspirasi atau pengaduan masyarakat tersebut tentunya menghasilkan jumlah data yang tidak sedikit, banyaknya data yang masuk menyebabkan lamanya waktu dan mempersulit admin dalam proses pengklasifikasian data. Keterbatasan pengetahuan oleh pihak admin yang berhubungan dengan data pengaduan dan pengetahuan tentang dinas atau lembaga yang bersangkutan menyebabkan kesalahan dalam proses pengklasifikasian data. Oleh karena itu, pihak pengadu harus menunggu

terlebih dahulu supaya data yang diadukan dapat segera didisposisikan.

Berdasarkan salah satu data pengaduan laporan pada laman LAPOR! yang telah memberikan laporannya pada tanggal 11 Desember 2017, kemudian didisposisikan pada tanggal 18 Desember 2017 dan belum adanya respon atau tanggapan dari pihak yang didisposisikan [4]. Hal tersebut tentunya cukup lama dalam proses pengelolaan data yang masuk, karena data laporan tersebut baru didisposisikan selama 7 hari kerja sedangkan berdasarkan Peraturan Walikota Bandung Nomor 1265 Tahun 2015 Tentang Pedoman Pengelolaan Layanan Aspirasi dan Pengaduan *Online* Rakyat di Lingkungan Pemerintah Kota Bandung bahwa data laporan yang masuk harus diproses selama 2 hari kerja oleh pihak pengelola data dan adanya respon selama 5 hari kerja oleh pihak yang didisposisikan. Oleh karena itu, proses pengelolaan data yang cukup lama dalam pengklasifikasiannya diperlukan adanya pengklasifikasian data dengan menggunakan metode *text mining*.

Penelitian yang dilakukan B. Herwijayanti, dkk mengenai metode klasifikasi dalam *text mining* yang dapat digunakan adalah metode *Term Frequency – Inverse Document Frequency* (TF-IDF) dan *Cosine Similarity*. Berdasarkan hasil pengujian dalam mengklasifikasikan artikel berita, metode tersebut mampu mengelompokkan artikel berita dengan tingkat akurasi sebesar 91.25%. Tujuan dari penelitian tersebut adalah untuk mempermudah dalam pengkategorian dan mengurangi adanya *human error* serta mengurangi terjadinya kesalahan pengkategorian [5]. Selain itu juga, penelitian juga dilakukan oleh Verry Prima Anindya Rhosady mengenai pengklasifikasian halaman *web* berdasarkan konten menggunakan metode *tf-idf* dan *cosine similarity*. Menurutny algoritma ini mampu

mengklasifikasikan seluruh halaman *web* dengan mengacu 8 kategori yaitu nasional, internasional, ekonomi, politik, hiburan, olahraga, teknologi, sains dan edukasi [6].

Data dalam penelitian ini bersumber dari Dinas Komunikasi dan Informatika (Diskominfo) Kota Bandung yaitu berupa data teks mengenai seluruh aspirasi atau pengaduan dari masyarakat Kota Bandung. Data yang diperoleh di klasifikasikan kata per kata dalam bentuk bahasa Indonesia menggunakan metode *text mining* sehingga mampu menghasilkan data pengaduan yang sesuai dengan kategori dinas yang dituju sehingga dapat diteruskan kepada dinas atau lembaga terkait. Oleh karena itu, dilakukanlah penelitian yang berjudul **Pengklasifikasian Pengaduan Masyarakat Menggunakan Algoritma *Cosine Similarity***.

## 1.2 Rumusan Masalah

1. Apakah pengimplementasian algoritma *cosine similarity* dapat membantu dalam mengklasifikasikan pengaduan masyarakat sesuai dengan lembaga disposisinya?
2. Apakah pengimplementasian algoritma *cosine similarity* dapat membantu mempercepat waktu proses pengelolaan dalam pengklasifikasian pengaduan masyarakat?

## 1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Sumber data yang digunakan dalam pengklasifikasian pengaduan masyarakat adalah data LAPOR! Kota Bandung berbahasa Indonesia;

2. Data yang digunakan adalah data LAPOR! dari bulan Agustus 2017- November 2017;
3. Data LAPOR! yang digunakan berasal dari *platform* web, android, pesan singkat sms, media sosial seperti twitter dan facebook;
4. Sistem yang dibangun berbasis *website* dengan menggunakan bahasa pemrograman PHP dan MySQL sebagai basis datanya.

#### **1.4 Tujuan Penelitian**

Tujuan yang hendak dicapai dalam penelitian ini adalah membangun sistem pengklasifikasian pengaduan masyarakat menggunakan algoritma *cosine similarity* untuk membantu dalam proses pengklasifikasian data pengaduan masyarakat.

#### **1.5 Manfaat Penelitian**

Manfaat dari pembuatan sistem ini adalah dapat mempermudah pihak pengelola data dalam mengelola seluruh data pengaduan masyarakat sehingga data pengaduan secara otomatis dapat diketahui harus kemana data pengaduan tersebut didisposisikannya.

#### **1.6 State Of The Art**

Penyusunan laporan tugas akhir ini mengambil beberapa referensi penelitian sebelumnya termasuk jurnal-jurnal yang berhubungan dengan penelitian ini.

Tabel 1.1 *State of The Art*

No	Judul	Masalah	Solusi	Hasil
1	Klasifikasi Berita <i>Online</i> dengan menggunakan Pembobotan <i>TF-IDF</i> dan <i>Cosine Similarity</i> Peneliti: Bening Herwijayanti Dian Eka Ratnawati Lailil Muflikhah (2018)	Banyaknya berita yang dimasukkan ke dalam <i>website</i> , sehingga terkadang berita tersebut terposting tidak sesuai dengan kategorinya.	Penggunaan metode pembobotan <i>tf-idf</i> dan <i>Cosine Similarity</i> .	Dari pengujian I, II, III, dan IV dari 154 data uji diketahui bahwa hasil akurasi pengujian tertinggi pada percobaan ke-3 dengan akurasi 100%.
No	Judul	Masalah	Solusi	Hasil
2	Klasifikasi Data Pengaduan Masyarakat Pada Laman Pesduk Cimahi Menggunakan <i>Rocchio</i> Peneliti: Agung Wahana Khusnul Khuluqiyah Tacbir Hendro Pudjiantoro (2016)	Bertambahnya jumlah data pengaduan mengakibatkan banyaknya data yang belum tertanggapi. Terbatasnya pengetahuan operator tentang tugas pokok dan fungsi dari setiap dinas menyebabkan kesalahan dalam pengklasifikasian pesan.	Penggunaan metode <i>text mining</i> serta menggunakan metode <i>Rocchio</i> dikombinasikan dengan fitur pembobotan <i>tf-idf</i> .	Metode <i>Rocchio</i> mampu melakukan klasifikasi pesan pengaduan masyarakat dan dapat mengklasifikasikan pesan secara otomatis tanpa harus dibaca satu persatu terlebih dahulu.

Tabel 1.1 *State of The Art* (Lanjutan)

No	Judul	Masalah	Solusi	Hasil
3	Rancang Bangun Aplikasi Pengklasifikasian Halaman Web Berdasarkan <i>Content</i> Menggunakan <i>Cosine Similarity</i>  Peneliti: Verry Prima Anindya Rhosady (2013)	Tidak adanya sistem pengkategorian, sehingga sering kali informasi yang didapat tidak akurat dan tidak tepat.	Penerapan metode <i>TF-IDF</i> dan <i>Cosine Similarity</i>	Mampu mengklasifikasikan halaman <i>web</i> , namun tingkat akurasi klasifikasi berbanding jauh dengan klasifikasi halaman <i>web</i> yang sudah ditentukan oleh <i>web</i> itu sendiri yaitu sebesar 30%.
No	Judul	Masalah	Solusi	Hasil
4	Implementasi <i>Text Mining</i> Untuk Menentukan Pola Klasifikasi Surat Dengan Metode <i>Tf-Idf</i> Dan <i>Cosine Similarity</i> Pada Pengadilan Agama Cianjur  Peneliti: Silmi Maulida (2016)	Pola klasifikasi surat di Pengadilan Agama Cianjur yang secara manual dapat menyulitkan dan memakan waktu <i>user</i> dalam melakukannya.	Penggunaan <i>text mining</i> dan metode pembobotan <i>TF-IDF</i> serta <i>Cosine Similarity</i> .	Pola klasifikasi dengan kategori surat didapat hasil akurasi mencapai 70% dari 20 data uji dengan 14 data percobaan akurat dan 6 data tidak akurat.

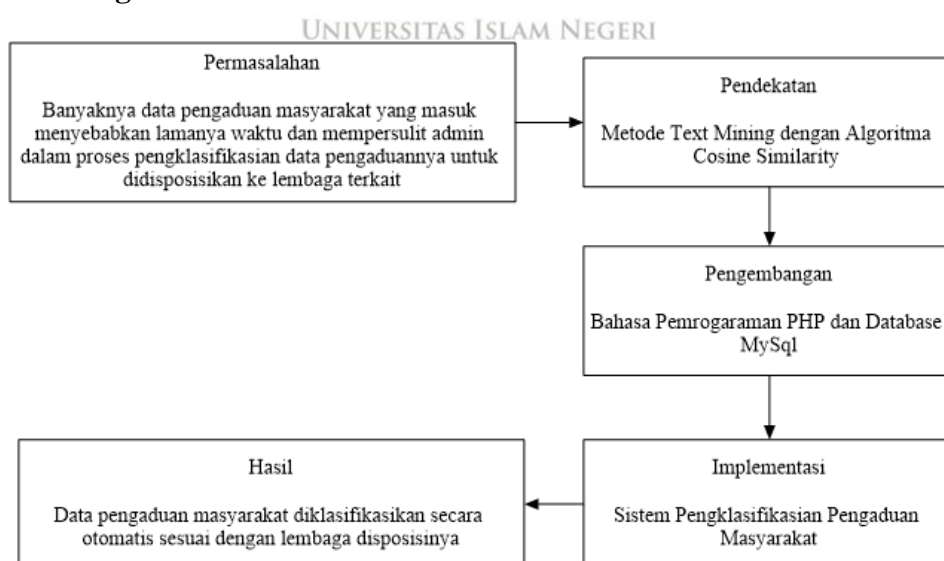
Tabel 1.1 *State of The Art* (Lanjutan)

No	Judul	Masalah	Solusi	Hasil
5	<i>Cosine Similarity to Determine Similarity Measure: Study Case in Online Essay Assessment</i>  Peneliti: Alfirna Rizqi Lahitani Adhistya Erna Permanasari Noor Akhmad Setiawan (2016)	Mekanisme penilaian jawaban esai yang lebih rumit karena mempertimbangkan komponen seperti sisi pengetahuan, sisi pendapat, keterampilan menulis dan sisi perilaku sehingga perlu membacanya satu per satu untuk pengecekan jawaban kemudian memberikan nilai.	Penerapan bobot <i>Term Frekuensi - Inverse Document Frequency</i> metode (TF-IDF) dan <i>Cosine Similarity</i> dengan konsep tingkat pengukuran dari segi kesamaan dalam dokumen.	Dengan menggunakan prinsip kesamaan, pengenalan karakter teks yang relevan dengan dokumen ahli memberikan hasil penilaian yang lebih obyektif dan mempercepat proses koreksi dalam kategori tugas esai berbasis teks.
No	Judul	Masalah	Solusi	Hasil
6	<i>Rating Of Indonesian Sinetron Based On Public Opinion In Twitter Using Cosine Similarity</i>  Peneliti: Vincentius Riandaru Prasetyo (2016)	Banyak sinetron yang menampilkan adegan yang tidak cocok untuk kelompok usia tertentu. Penonton yang merasa kecewa sering mengungkapkan pendapat mereka pada <i>twitter</i> .	Penerapan metode <i>TF-IDF</i> dan <i>Cosine Similarity</i>	Akurasi perhitungan similaritas dengan menggunakan kombinasi metode kesamaan <i>Cosine</i> dan <i>TF-IDF</i> dihasilkan akurasi sebesar 82.48%.

Tabel 1.1 *State of The Art* (Lanjutan)

No	Judul	Masalah	Solusi	Hasil
7	<i>Automated Document Classification for News Article in Bahasa Indonesia based on Term Frequency Inverse Document Frequency (TF-IDF) Approach</i>	Banyaknya data berita <i>online</i> yang semakin meningkat, sehingga sebagian besar data tidak dapat dikelola dengan mudah.	Penerapan metode <i>Term Frekuensi Inverse Document Frequency</i> (TF-IDF)	Algoritma <i>TF-IDF</i> dapat mengkategorikan artikel berita <i>online</i> dalam akurasi yang tinggi yaitu akurasi rata-rata adalah 98.3%.
	Peneliti Ari Aulia Hakim Alva Erwin Kho I Eng Maulahikmah (2014)			

## 1.7 Kerangka Pemikiran



Gambar 1.1 Kerangka Pemikiran



Gambar 1.1 menjelaskan bahwa proses pengklasifikasian data pengaduan masyarakat masih perlu dibaca satu persatu oleh admin pengelola, sebelum data tersebut didisposisikan ke lembaga terkait untuk direspon sehingga menyebabkan lamanya waktu dan mempersulit admin dalam pengklasifikasiannya. Oleh karena itu, dibangun sistem pengklasifikasian pengaduan masyarakat dengan menggunakan salah satu algoritma *text mining* yaitu algoritma *cosine similarity*. Pengembangan aplikasi yang dibangun menggunakan bahasa pemrograman PHP dan menggunakan MySql sebagai penyimpanan datanya. Adanya sistem tersebut maka pengklasifikasian pengaduan masyarakat dapat dilakukan secara otomatis tanpa harus dibaca satu persatu oleh admin. Hasil dari sistem tersebut adalah klasifikasi data pengaduan yang sesuai dengan dinas yang didisposikannya.

## 1.8 Metodologi Penelitian

Metode penelitian ini terdiri dari dua tahap, yaitu tahap pengumpulan data dan tahap pengembangan perangkat lunak.

### 1.8.1 Metode Pengumpulan Data

Metode yang digunakan untuk proses pengumpulan data dalam penelitian ini adalah sebagai berikut:

#### a. Wawancara (Interview)

Pengumpulan data dengan cara melakukan tanya jawab secara langsung dengan narasumber dari pihak instansi terkait.

**b. Pengamatan (Observasi)**

Pengumpulan data dengan cara mengadakan pengamatan atau peninjauan langsung ke objek atau masalah yang sedang diteliti tentang bagaimana data pengaduan masyarakat yang ada dapat ditunjukkan ke dinas terkait.

**c. Studi Literatur**

Pengumpulan data dengan cara mengumpulkan bacaan-bacaan yang berkaitan dengan masalah yang diteliti, misalnya mempelajari situs-situs yang berhubungan dengan objek penelitian seperti situs-situs sekolah dan pengumpulan data yang dilakukan dengan mempelajari buku-buku literatur.

**d. Dokumentasi**

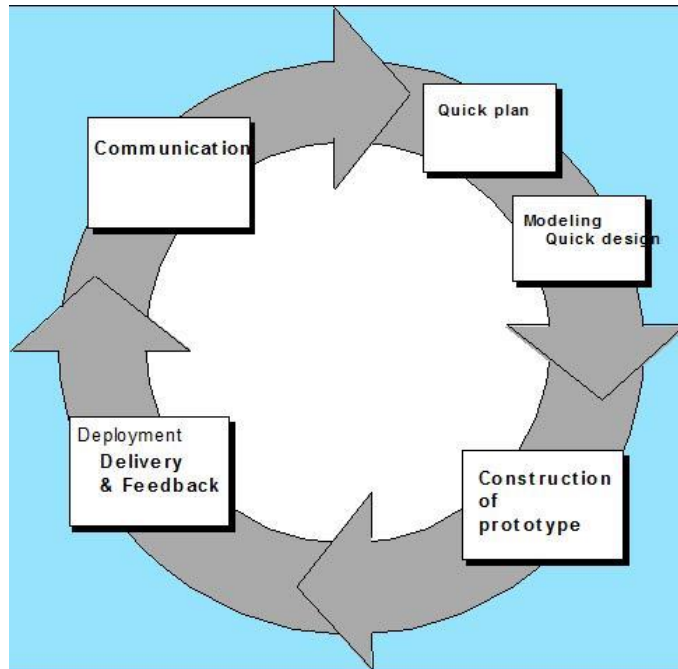
Selain melalui wawancara, pengamatan dan studi literatur peneliti juga melakukan pengumpulan data melalui dokumentasi yang berasal dari data sekunder, dalam hal ini adalah data pengaduan masyarakat Kota Bandung.

### **1.8.2 Metode Pengembangan Perangkat Lunak**

Pembuatan sistem ini menggunakan metode pengembangan perangkat lunak *Prototype*, karena metode ini lebih memudahkan proses dalam membangun perangkat lunak dan pelanggan dapat saling berinteraksi selama proses pembuatan sistem. Keuntungan dari *Prototype* antara lain [7]:

- a. *User* dapat mempertimbangkan sedikit perubahan selama masih bentuk *prototype*;
- b. Memberikan hasil yang lebih akurat dari pada perkiraan sebelumnya, karena fungsi yang diinginkan dan kerumitannya sudah dapat diketahui dengan baik;

- c. *User* merasa puas. Pertama, *user* mengenai komputer dan aplikasi yang akan dbuatkan untuknya. Kedua, *user* terlibat langsung dari awal dan memotivasi semangat untuk mendukung analisis selama proyek berlangsung.



Gambar 1.2 Metode Pengembangan Perangkat Lunak *Prototype* [7]

Berdasarkan Gambar 1.2, proses metode pengembangan *Prototype* dapat dijelaskan sebagai berikut:

- a. *Communication*

*Developer* dan klien bertemu dan menentukan tujuan umum, kebutuhan yang diinginkan dan gambaran bagian-bagian yang akan dibutuhkan.

- b. *Quick Plan*

Perancangan dilakukan cepat dan mewakili semua aspek *software* yang diketahui, dan rancangan ini menjadi dasar pembuatan *prototype*.

c. *Modelling Quick Design*

Berfokus pada representasi aspek *software* yang bisa dilihat *user*. *Modelling Quick Design* cenderung ke pembuatan *prototype*.

d. *Construction of Prototype*

Membangun kerangka atau rancangan *prototype* dari *software* yang akan dibangun.

e. *Deployment Delivery & Feedback*

*Prototype* yang telah dibuat oleh *developer* akan disebarakan kepada *user*, untuk dievaluasi, kemudian *user* akan *memberikan feedback* yang akan digunakan untuk merevisi kebutuhan *software* yang akan dibangun. Pengulangan proses ini terus berlangsung sampai semua kebutuhan terpenuhi.

## 1.9 Sistematika Penulisan

Penulisan skripsi ini dibagi kedalam lima bab, yang disusun berdasarkan sistematika sebagai berikut:

### **Bab 1: Pendahuluan**

Bab ini berisi mengenai latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, kerangka pemikiran, metodologi penelitian dan sistematika penulisan.

### **Bab II: Studi Pustaka**

Dalam bab ini akan dituliskan mengenai teori-teori yang akan digunakan dan relevan dengan penelitian ini serta penjelasan mengenai penelitian-penelitian sebelumnya terkait dengan klasifikasi pengaduan masyarakat.

### **Bab III: Perancangan Sistem**

Dalam bab ini akan dijelaskan mengenai analisis dan perancangan terhadap penelitian yang akan dilakukan. Dimulai dari analisis sistem, analisis kebutuhan fungsional, analisis kebutuhan perangkat lunak maupun perangkat keras, analisis arsitektur sistem, serta perancangan dalam pembuatan sistem tersebut yang meliputi *Data Flow Diagram* (DFD) serta *user interface*.

### **Bab IV: Implementasi Sistem**

Dalam bab ini akan membahas tentang implementasi klasifikasi pengaduan masyarakat yang dilakukan, skenario pengujian dan hasil pengujian tentang penelitian yang dilakukan serta perhitungan akurasi dari metode yang digunakan pada penelitian.

### **Bab V: Penutup**

Dalam bab ini akan dibahas mengenai kesimpulan dan saran yang didapatkan dari serangkaian proses penelitian yang telah dilakukan.

