BAB 2
LANDASAN TEORI

2.1 Kemacetan

2.1.1 Pengertian Kemacetan

"Kemacetan adalah situasi atau keadaan tersendatnya atau bahkan terhentinya lalu lintas yang disebabkan oleh banyaknya jumlah kendaraan melebihi kapasitas jalan. Kemacetan banyak terjadi di kota-kota besar, terutamanya yang tidak mempunyai transportasi publik yang baik atau memadai ataupun juga tidak seimbangnya kebutuhan jalan dengan kepadatan penduduk" [8].

Berdasarkan penjelasan tersebut, penulis berpendapat bahwa kemacetan merupakan suatu masalah lalu lintas yang timbul akibat tidak seimbangnya jumlah populasi penduduk dan jumlah kendaraan dengan sarana atau prasarana lalu lintas. Di samping itu, berdasarkan fakta-fakta yang terlihat pada keadaaan lalu lintas di kota-kota besar, kemacetan juga merupakan faktor utama timbulnya masalah-masalah lain yang berkaitan erat dengan penggunaan lalu lintas dan juga merupakan masalah yang masih sulit untuk diselesaikan.

2.1.2 Penyebab Kemacetan

Salah satu faktor yang menyebabkan kemacetan adalah karena kurangnya informasi mengenai kemacetan itu sendiri. Pada kenyataannya, pemerintah kota hanya menyediakan sarana dan pra-sarana sebagai pendukung lalu lintas. Akan tetapi, faktor dalam pengaturan sistemnya sendiri terabaikan. Sehingga, faktor-faktor yang terabaikan tersebut menyebabkan kemacaten. Disamping itu, informasi mengenai kemacetan yang telah terjadi pun kurang tersampaikan dengan baik. Oleh karena itu, perlu adanya suatu sistem yang baik agar penyampaian informasi mengenai kemacetan dapat tersampaikan dengan baik pula.

2.2 Text Mining

Text mining adalah cara pengekstraksian informasi yang tidak diketahui sebelumnya dari suatu sumber data yang berbeda ke dalam bentuk baru dimana bentuk baru dan sumber data tersebut mempunyai keterkaitan [3]. *Text Mining* merupakan salah satu bentuk explorasi dan analisis data teks yang bertujuan untuk mendapatkan pengetahuan baru baik itu melalui cara otomatis maupun semi otomatis [13].

Berdasarkan kutipan-kutipan diatas, penulis berpendapat bahwa *Text mining* merupakan proses penggalian informasi-informasi baru dan berguna dari data lama menjadi data baru dimana kedua data tersebut saling berhubungan.

2.3 Ekstraksi Informasi

2.3.1 Pengertian Ekstraksi Informasi

Ekstraksi Informasi adalah pengambilan fakta dan informasi terstruktur dari isi koleksi teks yang besar. Pengertian fakta disini adalah beragam entitas yang diperhitungkan. Secara singkat ekstraksi informasi adalah sebuah proses mendapatkan fakta-fakta terstruktur dari data yang tersedia [10].

Berdasarkan penjelasan dari kutipan di atas, penulis berpendapat bahwa ekstraksi informasi adalah proses penyaringan parameter-parameter informasi fakta dari data yang telah tersedia.

2.3.2 Manfaat Ekstraksi Informasi

Ekstraksi Informasi, mempunyai banyak sekali manfaat tergantung hasil ekstraksi apa yang ingin kita capai. “Pada tahun 1959, Calvin Mooers membuat istilah pengambilan informasi (*information reetrival*) dan membuat kategori informasi. Tujuan awalnya adalah untuk memungkinkan pencarian (searching) secara cepat pada jumlah data set yang banyak. Selanjutnya, Maron dan Kuhn di tahun 1961 membahas lebih jauh tentang relevansi antardokumen dalam permasalahan pengurutan (*sorting*) dan pemberian peringkat dalam pengelolaan dokumen.” [10].

Meskipun pengambilan informasi dan ekstraksi informasi adalah dua metode yang berbeda, akan tetapi perannya hampir sama, perebedaannya adalah "ekstraksi informasi memberikan pengetahuan dengan tingkat (level) yang lebih dalam dibanding pengambilan informasi yang dilakukan secara tradisional yang menganalisa fakta-fakta dari berbagai dokumen." [10].

2.4 Algoritma KNN

Algoritma KNN adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi data pembelajaran [4].

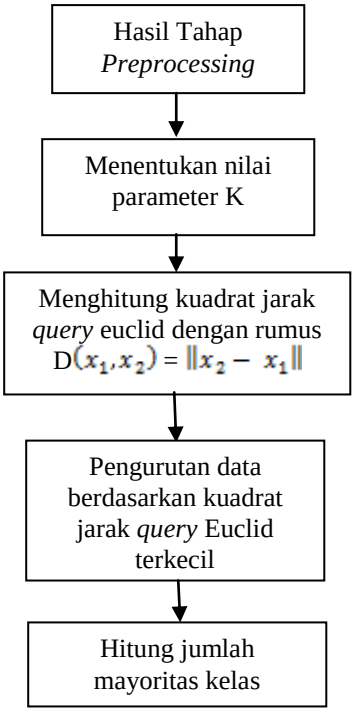
Adapun langkah-langkah dari algoritma KNN, yaitu:

1 Menentukan parameter K (jumlah tetangga paling dekat);.
2 Menghitung kuadrat jarak *query* euclid masing-masing objek terhadap data *sample* yang diberikan;
3 Urutkan objek-objek tersebut berdasarkan urutan jarak euclid terkecil;
4 Hitung jumlah mayoritas kelas berdasarkan K (jumlah terdekat).

Dalam penelitian ini, proses perhitungan kuadrat jarak *query* menggunakan Euclidean Distance. Euclidean Distance memiliki rumus:

$$D(x_1, x_2) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^{p}(x_{2j} - x_{1j})^2}$$

Dimana x_1 adalah nilai dari penyebab dan kondisi dari data uji dan x_2 adalah nilai dari masing-masing penyebab dan kondisi dari data uji.



Gambar 2-1 Alur kerja algoritma KNN

Berikut adalah *pseudocode* dari algoritma KNN:

```
Procedure Knn(input: k:integer, querySet:larikPoint,
trainingSet:larikPoint)

{I.S: Menentukan kelas pada kueri terbaru berdasarkan
mayoritas kelas}
{F.S: Mendapatkan kelas pada kueri terbaru berdasarkan
mayoritas kelas}

Kamus:
   d: double
   map: SortedMap
   query: pointer

Algoritma:
   foreach (query in querySet) do
       foreach (training in trainingSet) do
           map ← SortedMap(k);
               if (query != 0) do
```

```
                    d ← EuclideanDistance(query, training);
                    map(d, training);
                    else query ← 0;
                       d ← EuclideanDistance(query, training);
                       map(d, training);
                endif
        endfor
        query↑label ← majorityVote(map);
endfor
```

2.5 Sumber Data

2.5.1 Twitter

Twitter merupakan sebuah situs web sosial media yang dimiliki dan dioperasikan oleh Twitter Inc. Twitter merupakan media sosial mikroblog sehingga memungkinkan penggunanya untuk mengirim dan membaca pesan yang disebut tweet [9]. Tweet adalah teks tulisan berjumlah 140 karakter yang ditampilkan pada halaman profil pengguna. Tweet bisa dilihat secara luar, namun pengirim dapat membatasi pengiriman pesan ke daftar teman-teman mereka saja.

Selain itu, Twitter mempunyai pilihan untuk memproteksi akun twitter para penggunanya, agar tweet mereka tidak dapat dilihat oleh sembarang orang. Biasanya pengguna akun yang membuka hak akses mereka adalah akun-akun publik seperti akun @BandungTraffic yang melakukan tweet mengenai informasi-informasi seputar kemacetan di kota Bandung.

"Pengguna dapat melihat tweet penulis lain yang dikenal dengan sebutan follower. Semua pengguna dapat mengirim dan menerima tweet melalui situs Twitter, aplikasi eksternal yang kompatibel (telepon seluler), atau dengan pesan singkat (SMS) yang tersedia di negara-negara tertentu" [9]. Selain itu, Twitter memiliki beberapa fitur andalan yang membuat Twitter menjadi salah satu media sosial yang digemari oleh para penggiat internet. Berdasarkan hasil survei www.zocialrank.com, pengguna Twitter di Indonesia sebanyak 1.028.332 akun.

Adapun objek-objek seperti, tweet, re-tweet, at atau mention (@), hashtag (#), waktu tweet dan tweetsearch.

2.5.2 Lewat Mana

Lewat Mana merupakan salah satu *website* yang menyediakan informasi kemacetan. Lewat mana dibangun dengan tujuan untuk menghadirkan informasi terkini kepada penguna jalan dengan harapan dapat membantu merencanakan perjalanan dengan lebih efektif dan efisien [12].

Dalam penyampaian informasi kemacetan, *website* Lewat Mana mempunyai kemiripan dengan sumber data yang telah dijelaskan sebelumnya, yaitu media sosial Twitter. Kemiripan tersebut adalah pada bentuk penyampaiannya yang berupa *timeline*.

2.6 Web Service dan CodeIgniter

Web Service adalah sebuah software aplikasi yang dapat teridentifikasi oleh URI dan memiliki interface yang didefiniskan, dideskripsikan, dan dimengerti oleh XML dan juga mendukung interaksi langsung dengan software aplikasi yang lain dengan menggunakan message berbasis XML melalui protokol internet [15].

Web Service merupakan jembatan koneksi antara server dan client agar keduanya dapat saling berkomunikasi dengan baik, selain itu dengan adanya *web service* dapat memungkinkan banyak *client* dengan *platform* yang berbeda-beda. Salah satu framework yang biasa digunakan untuk pembangunan *web service* adalah CodeIgniter.

CodeIgniter adalah aplikasi open source yang berupa framework dengan model MVC (Model, View, Controller) untuk membangun website dinamis dengan menggunakan PHP. CodeIgniter memudahkan developer untuk membuat aplikasi web dengan cepat dan mudah dibandingkan dengan membuatnya dari awal. [14]

2.7 Android

Android merupakan sebuah sistem operasi yang dimiliki oleh perusahaan Google, Inc. Pada saat perilisan perdana Android, 5 November 2007, Android bersama Open Handset Alliance menyatakan mendukung pengembangan standar terbuka pada perangkat seluler. Di lain pihak, Google merilis kode–kode Android di bawah lisensi Apache, sebuah lisensi perangkat lunak dan standar terbuka perangkat seluler [11].

Saat ini, Android menjadi salah satu *smartphone* yang paling banyak digemari, karena dengan segala fitur yang dimiliki oleh Google seperti Google Maps, Google Drive dan Google Search, hal tersebut sudah dapat terintegrasi dengan *smartphone* Android.