

ALGORITMA KNUTH-MORRIS-PRATT DAN PENGELOMPOKAN KLASIFIKASI DATA UNTUK MENENTUKAN SENTIMEN *TWEET* POSITIF ATAU NEGATIF

Ahmad Farhan Ghifari - 13515602

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia

13515602@std.stei.itb.ac.id

Abstrak—Twitter merupakan salah satu media sosial yang saat ini sedang populer di kalangan masyarakat. Banyak hal yang dapat dieksplorasi dari twitter. Salah satu bentuk eksplorasinya yaitu dari sebuah *tweet* dapat ditentukan apakah *tweet* tersebut termasuk kedalam bernada sentimen positif ataupun negatif. Dengan menggunakan algoritma pencocokan string Knuth-Morris-Pratt dapat dicari kata-kata yang bermakna negatif dan kata-kata yang bermakna positif. Selanjutnya dengan algoritma pengelompokan klasifikasi data dapat ditentukan apakah *tweet* tersebut bersentimen positif ataukah negatif. Melalui penelitian dasar ini selanjutnya dapat ditentukan suatu algoritma yang dapat memeriksa sentimen topik secara keseluruhan.

Kata kunci—Twitter, KMP, Klasifikasi Data.

I. PENDAHULUAN

Komunikasi merupakan salah satu hal dasar dalam melakukan hubungan sosial kepada orang lain. Di zaman sebelum teknologi berkembang setiap orang hanya dapat melakukan komunikasi dengan berhadapan langsung ataupun melalui surat. Namun saat ini perkembangan teknologi semakin pesat. Alat komunikasi mulai ditemukan dari telepon lokal hingga kini telepon pintar (*smartphone*). Semakin berkembangnya teknologi maka cara berkomunikasi pun menjadi bertambah. Salah satunya yaitu dengan menggunakan media sosial.

Sosial media adalah sebuah media untuk bersosialisasi satu sama lain dan dilakukan secara online yang memungkinkan manusia untuk saling berinteraksi tanpa dibatasi ruang dan waktu^[1]. Sosial media dapat dikelompokkan menjadi beberapa bagian yaitu^[1] 1. Social Networks, media sosial yang memfasilitasi sekelompok orang untuk bersosialisasi dan berinteraksi (Facebook, myspace, hi5, Linked in, bebo, dll) ; 2. Discuss, media sosial yang memfasilitasi sekelompok orang untuk melakukan obrolan dan diskusi (google talk, yahoo!, skype, phorum); 3. Share, media sosial yang memfasilitasi kita untuk saling berbagi file, video, music, dll (youtube, slideshare, feedback, flickr, crowdstorm, dll); 4. Publish, (wordpredss, wikipedia, blog, wikia, digg, dll); 5. Social game, media sosial

berupa game yang dapat dilakukan atau dimainkan bersama-sama (koongregate, doof, pogo, cafe.com, dll); 6. MMO (kartrider, warcraft, neopets, conan, dll); 7. Virtual worlds (habbo, imvu, starday, dll); 8. Livecast (y! Live, blog tv, justin tv, listream tv, livecastr, dll); 9. Livestream (socializr, froendsfreed, socialthings!, dll); 10. Micro blog (twitter, plurk, pownce, twirxr, plazes, tweetpeek, dll)

Salah satu media sosial yang dalam beberapa tahun ini sering digunakan oleh warga Indonesia adalah Twitter. Indonesia telah menjadi negara dengan pengguna Twitter terbesar di Asia dan ke tiga di dunia (PeerReach, 2013) dengan banyak pengguna Twitter di tahun 2010 adalah 6,5% dari seluruh pengguna di dunia. Angka ini naik dari 2,41% di tahun 2010 (Sysmosos Inc., 2010). Dari banyaknya jumlah pengguna Twitter di Indonesia ini dapat dihasilkan sebuah informasi yang dapat digunakan untuk menentukan sentimen masyarakat terhadap sebuah topik. Dengan menggunakan hasil sentimen ini selanjutnya dapat dianalisa seberapa tingkat kepuasan masyarakat terhadap suatu topik.

II. DASAR TEORI

Pencarian *string* di dalam teks disebut juga pencocokan string (*string matching* atau *pattern matching*). Persoalan pencarian string dapat dirumuskan menjadi dua yaitu teks dan *pattern*. Teks merupakan bagian kumpulan kalimat yang memiliki panjang n karakter. *Pattern* merupakan string dengan panjang m karakter ($m < n$) yang akan dicari di dalam teks. Dalam melakukan pencocokan string terdapat tiga algoritma yang biasa digunakan yaoti algoritma *brute force*, algoritma Knuth-Morris-Pratt (KMP), dan algoritma Boyer-Moore. Namun dalam makalah ini hanya akan dibahas mengenai algoritma KMP saja..

Dalam melakukan pengelompokan data digunakan metode pengelompokan data secara klasifikasi. Dengan cara ini akan dihasilkan data apakah *tweet* tersebut termasuk ke dalam tweet dengan sentimen negatif atau positif.

A. Algoritma Knuth-Morris-Pratt (KMP)

Algoritma KMP merupakan algoritma pencocokan string yang dirancang oleh Donald Ervin Knuth. Donald Ervin Knuth dilahirkan pada 10 Januari 1938. Beliau adalah seorang ilmuwan komputer dan Professor Emeritus di Stanford University. Dia adalah penulis seminal multi-volume yang bekerja The Art of Computer Programming. Knuth juga disebut sebagai “bapak” dari analisis algoritma. Ia memberikan kontribusi untuk pengembangan analisis ketat dari kompleksitas komputasi algoritma dan teknik matematika formal sistematis. Dalam proses ini ia juga mempopulerkan notasi asimtotik.

Algoritma KMP mencocokkan pattern dengan menelusuri text dari kiri ke kanan. Jika ditemukan ketidakcocokan pada saat mencocokkan, maka pergeseran dilakukan sebanyak jumlah terbesar prefik yang juga merupakan sufik dari pattern. Proses menentukan jumlah pergeseran tersebut dilakukan dengan fungsi pinggiran. Fungsi pinggiran melakukan pencocokan prefik dari pattern dengan pattern itu sendiri. Pada implementasi KMP proses fungsi pinggiran dilakukan sebelum mencocokkan pattern dengan text.

Kompleksitas waktu yang dibutuhkan algoritma KMP untuk menghitung fungsi pinggiran sebanyak $O(m)$ dengan m adalah panjang pattern. Selain itu, kompleksitas waktu yang dibutuhkan untuk proses pencocokan string sebanyak $O(n)$. Sehingga, kompleksitas waktu algoritma KMP adalah $O(m+n)$.

Berikut adalah algoritma KMP dalam bahasa Java:

a. Fungsi Pinggiran

```
public static int[] computeFail(String
pattern){
    int m = pattern.length();
    int j = 0;
    int i = 1;
    int fail[] = new int[m];
    fail[0] = 0;
    while(i < m){
    if(pattern.charAt(j)==pattern.charAt(i))
    {
        fail[i] = j + 1;
        i++; j++;
    }
    else if(j > 0) j = fail[j -1];
    else{
        fail[i] = 0;
        i++;
    }
    }
    return fail;
}
```

b. Algoritma KMP

```
public static int kmpMatch( String
pattern,String text)
```

```
{
    int n = text.length();
    int m = pattern.length();
    int fail[] =computeFail(pattern);
    int i = 0, j = 0;
    while(i < n){
        if(pattern.charAt(j) ==
        text.charAt(i)){
            if(j == m -1){
                return i - m + 1;
            }
            i++; j++;
        }
        else if(j > 0){
            j = fail[j -1];
        }
        else { i++; }
    }
    return -1;
}
```

B. Pengelompokan Data Klasifikasi

Pengelompokan data merupakan cara untuk memisahkan data agar dapat ditentukan data tersebut masuk pada kelompok tertentu. Pengelompokan ini dapat dilakukan dengan memisah data menjadi beberapa bagian sesuai dengan kebutuhan.

Data adalah suatu istilah majemuk yang berarti fakta atau bagian dari fakta yang mengandung arti yang dihubungkan dengan kenyataan, simbol-simbol, gambar-gambar, angka-angka huruf-huruf atau symbol-symbol yang menunjukkan suatu ide, objek, kondisi, atau situasi dan lain lain. Kumpulan dari data yang telah diolah dapat menjadi sebuah informasi.

Klasifikasi adalah teknik pengumpulan data yang digunakan untuk memprediksi kelompok keanggotaan untuk data instans. Sedangkan menurut Tan Klasifikasi adalah sebuah proses untuk menemukan model yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang kelasnya tidak diketahui

Klasifikasi data menurut jenisnya dapat dibedakan menjadi sebagai berikut.

a. Klasifikasi Menurut Jenis Data

1. Data Hitung (Enumeration / Counting Data)

Data hitung adalah hasil perhitungan atau jumlah tertentu. yang termasuk data hitung adalah persentase dari suatu jumlah tertentu. Mencatat jumlah mahasiswa dalam suatu kelas atau

persentasi mahasiswa/mahasiswi dalam kelas itu menghasilkan suatu data hitung.

2. Data Ukur (Measurement Data)

Data ukur adalah data yang menunjukkan ukuran mengenai nilai sesuatu. Angka tertentu atau huruf tertentu yang diberikan oleh seorang dosen kepada mahasiswa setelah memeriksa hasil tentamennya di sebut data ukur. Angka yang ditunjukkan alat barometer atau thermometer adalah hasil proses pengukuran.

b. Klasifikasi Menurut Sifat Data

1. Data Kuantitatif (Quantitative Data)

Data kuantitatif adalah data mengenai penggolongan dalam hubungannya dengan penjumlahan. Sebagai contoh adalah jika SMA dibagi menjadi dua maka terdapat dua golongan yaitu yang jumlah siswanya lebih dari 100 orang dan SMA dengan jumlah siswanya kurang dari 100 orang. Ini merupakan penggolongan kuantitatif.

2. Data Kualitatif

Data kualitatif adalah data mengenai penggolongan dalam hubungannya dengan kualitas atau sifat tertentu. Penggolongan fakultas-fakultas pada universitas negeri menjadi exacta dan fakultas exacta merupakan pemisahan menurut sifatnya. Penggolongan mahasiswa pada fakultas yang menggunakan sistem kredit ke dalam penilaian studi dengan grade A,B,C,D didasarkan pada sifat-sifat kualitatifnya.

c. Klasifikasi Menurut Sumber Data

1. Data Internal

Data internal adalah data yang asli, artinya data sebagai hasil observasi yang dilakukan sendiri, bukan data hasil karya orang lain.

2. Data Eksternal

Data eksternal adalah data hasil observasi orang lain, seseorang boleh saja menggunakan data untuk suatu keperluan, meskipun data tersebut merupakan hasil kerja orang lain. Data eksternal ini di bagi menjadi 2:

1.) Data Eksternal Primer

Data primary eksternal adalah data dalam bentuk ucapan lisan atau tulisan dari pemiliknya sendiri, yakni orang yang melakukan observasi sendiri.

2.) Data Eksternal Sekunder

Data eksternal sekunder adalah data yang diperoleh bukan dari orang lain yang melakukan observasi melainkan melalui seseorang atau sejumlah orang lain.

d. Klasifikasi Data Menurut Cara Memperolehnya

1. Data Primer

Data primer adalah secara langsung diambil dari objek / obyek penelitian oleh peneliti perorangan maupun organisasi. Contoh : Meneliti langsung pengunjung restoran x untuk meneliti preferensi konsumen bioskop.

2. Data Sekunder

Data sekunder adalah data yang didapat tidak secara langsung dari objek penelitian. Peneliti mendapatkan data yang sudah jadi yang dikumpulkan oleh pihak lain dengan berbagai cara atau metode baik secara komersial maupun non komersial. Contohnya adalah pada peneliti yang menggunakan data statistik hasil riset dari surat kabar atau majalah.

e. Klasifikasi Data Menurut Waktu Pengumpulannya

1. Data Cross Section

Data cross-section adalah data yang menunjukkan titik waktu tertentu. Contohnya laporan keuangan per 31 desember 2006, data pelanggan PT. angin ribut bulan mei 2004, dan lain sebagainya.

2. Data Berkala

Data berkala adalah data yang datanya menggambarkan sesuatu dari waktu ke waktu atau periode secara historis. Contoh data time series adalah data perkembangan nilai tukar dollar amerika terhadap euro eropa dari tahun 2004 sampai 2006, jumlah pengikut jamaah nurdin m. top dan doktor azahari dari bulan ke bulan, dll.

f. Klasifikasi Data Berdasarkan Sifat Nilai Data

1. Data Diskrit

Data diskrit adalah data yang nilainya adalah bilangan asli. Contohnya adalah berat badan ibu-ibu PKK Warasatih, nilai rupiah dari waktu ke waktu, dan lain-sebagainya.

2. Data Kontinyu

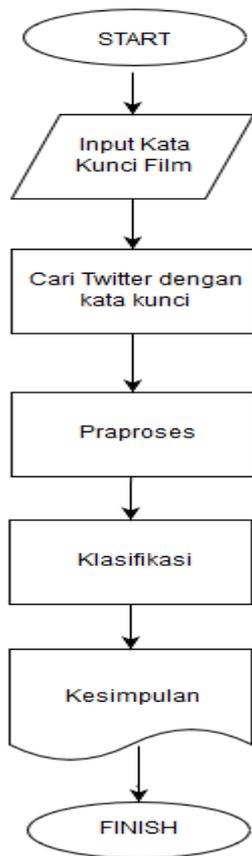
Data kontinyu adalah data yang nilainya ada pada suatu interval tertentu atau berada pada nilai yang satu ke nilai yang lainnya. Contohnya penggunaan kata sekitar, kurang lebih, kira-kira, dan sebagainya. Dinas pertanian daerah mengimpor bahan baku pabrik pupuk kurang lebih 850 ton.

Untuk makalah ini akan diterapkan model klasifikasi menurut jenis data. Data akan dikelompokkan sesuai dengan jumlah perhitungannya.

III. ANALISIS DATA SENTIMEN

Data Twitter diambil dengan menggunakan Twitter API yang disediakan oleh Twitter. Pencarian twitter dengan kata

kunci tertentu dilakukan dengan menggunakan algoritma KMP. Selanjutnya dari data ini dianalisis dengan menggunakan kamus data yang sudah ada. Proses pengekstraksian data dari *fetching* hingga menjadi sebuah kesimpulan dapat dilihat pada diagram alir berikut.



Gambar 1. Diagram Alir

Proses analisis dimulai dengan memasukkan kata pencarian kedalam sistem. Kata pencarian yang dimasukkan adalah bebas. Pengguna dapat menggunakan kata pencarian apapun sesuai dengan yang dibutuhkan.

Selanjutnya tweet dapat diperoleh dengan menggunakan Twiiter API. Proses ini menggunakan website sebagai pencariannya. Algoritma yang digunakan untuk mencocokkan string teks dengan kata kunci adalah algoritma KMP.

Praproses merupakan salah satu langkah penting dalam melakukan klasifikasi data untuk menentukan sentimen publik. Karakter yang didapat dari twitter dapat bermacam – macam sehingga diperlukan langkah – langkah praproses sebagai berikut :

1. *Tokenizer.*

Tokenizer merupakan proses untuk memisahkan kata atau word. Proses ini dilakukan karena adanya emoticon yang sering digunakan oleh pengguna Twitter dalam menulis tweetnya. Tahapannya dimulai dengan memisahkan bagian tweet yang dipisahkan dengan karakter spasi. Selanjutnya, bagian ayng hanya

memiliki satu karakter non alfabet dan angka tidak akan digunakan atau dengan kata lain dibuang. Bagian yang termasuk dalam daftar emoticon akan diubah. Proses terakhirnya yaitu bagian yang memiliki karakter selain alfabet, angka dan garis bawah akan dipecah sesuai posisi karakter tersebut.

2. *Normalisasi Fitur.*

Beberapa komponen khas yang biasanya ada di tweet yaitu username, url, “RT” (tanda retweet), dan juga hashtag(#). Komponen diatas tidak akan memiliki pengaruh apapun sehingga akan dibuang. Komponen username diidentifikasi dengan kemunculan karakter ‘@’, sedangkan komponen url dikenali melalui ekspresi reguler. Terakhir untuk kasus hashtag maka akah diubah menjadi string hashtag.

3. *Casefolding.*

Proses ini dilakukan untuk mengubah semua huruf menjadi lowercase atau huruf kecil.

4. *CleanNumber.*

Menurut Aqsath & Ayu (2011) dalam jurnal Analisis Sentimen dan Ekstraksi Topik Penentu Sentimen pada Opini terhadap Tokoh Publik yang ditulis oleh Ismail Sunni, kemunculan angka di depan dan diakhir kata memiliki pengaruh ayng tidak signifikan dari nilai sentimen suatu tweet.

5. *ConvertNumber.*

Masih berdasarkan Aqsath & Ayu dalam jurnal Ismail Sunni, kemunculan di tengah-tengah kata perlu diubah menjadi karakter serupa. Konversi angka akan dilakukan berdasarkan Tabel 1 Konversi Angka sebagai berikut.

Angka	Konversi	Angka	Konversi
0	o	5	s
00	u	6	g
1	i	7	t
2	Sama seperti karakter sebelum angka 2	8	b
3	e	9	g
4	a		

Tabel 1. Pengkonversian Angka

6. *ConvertEmoticon.*

Emoticon merupakan gambar yang dibentuk dari kumpulan karakter untuk mengungkapkan perasaan secara tekstual. Hal ini dapat membantu untuk

menentukan sentimen terhadap tweet. Setiap emotikon akan dikonversi menjadi string yang sesuai dengan tabel 2 Konversi Emotikon ini.

Emotikon	Konversi
>:] :-) :) :-o) :] :3 :c) :> =] 8) =) :} :^)	-e-senang-
>:D :-D :D 8-D 8D x-D xD XD XD =-D =D =-3 =3	-e-ketawa-
>:[:-(:(: :-c :c :-< :< :-[:[:{ >.>.<.<.>.<	-e-horor-
>:P :-P :P X-P x-p xp XP :-p :p =p :-P :P :-b :b	-e-melet-
>:o >:O :-O :O °o° °O° :O o_o o.O 8-0	-e-shock-
>:\ >:/ :-/ :-.\ :/\ =/= \:S	-e-bingung-
;\ :-	-e-muka-datar-

Tabel 2. Pengkonversian Emotikon

7. *ConvertWord.*

Beberapa kata dalam twitter dapat berupa kata yang tidak baku. Hal ini seperti kata – kata gaul yang digunakan oleh anak zaman masa kini. Oleh karena itu, untuk kata yang tidak baku akan diubah ke kata terdekat dengan menggunakan kamus yang dibuat dengan melihat pola kemunculan kata-kata tidak baku tersebut.

8. *ConvertNegation.*

Di dalam sebuah kalimat dapat terdapat kata yang bersifat negasi yang dalam hal ini akan merubah nilai sentimen suatu tweet. Oleh karena itu kata-kata ini dihilangkan dan diganti oleh penanda saja sehingga tidak perlu diproses lebih lanjut. Ketika banyak kata negasi adalah ganjil, maka sentimen tweet tersebut akan diubah. Beberapa kata yang bersifat negasi ini adalah “bkn”, “bukan”, “tidak”, “enggak”, “g”, “ga”, “jangan”, “nggak”, “tak”, “tdk”, dan “gak”.

9. *StopWordRemoval*

Di dalam proses ini akan menghilangkan kata-kata yang sering muncul namun tidak memiliki pengaruh apapun dalam ekstraksi sentimen suatu tweet. Kata ini termasuk ke dalam kata gaul yaitu seperti “gw”, “cc”, “otw”, atau “ente”; lalu kata hubung seperti “di”, “dengan”, atau “ke”, penunjuk waktu dan kata tanya.

Tahap selanjutnya yaitu tahap klasifikasi. Tahap klasifikasi yaitu tahap untuk menghitung banyak kata yang tergolong ke dalam negatif dan positif. Pengklasifikasian tersebut berdasarkan pada kamus yang telah dibuat. Sampel dari kata kata negatif dan positif tersebut adalah sebagai berikut.

Positif	Negatif
Keren	Jelek
Bagus	Cupu
Menarik	Aneh
Indah	buruk

Tabel 3. Sampel Kamus Kata Positif dan Negatif

Setiap tweet selanjutnya di cek berapa banyak tweet tersebut mengandung kata positif atau negatif. Setelah itu hasil yang terbanyak antara jumlah kata positif atau negatif dalam tweet tersebut akan menentukan tweet tersebut termasuk sentimen positif atau negatif. Algoritma untuk melakukan pengecekan berapa banyak jumlah kata positif dan negatif dari suatu tweet adalah sebagai berikut.

```
private int posLength = 0;
private int negLength = 0;
public void searchData(String text){
    kmp kmpp = new kmp();
    int temp;
    for(int i=0;i<posLength;i++){
        temp = kmpp.kmpMatch(text, positif[i]);
        if (temp!=-1)
            pos++;
    }
    for(int i=0;i<negLength;i++){
        temp = kmpp.kmpMatch(text, negatif[i]);
        if (temp!=-1)
            neg++;
    }
}
```

Tahap akhir dari sistem ini yaitu penarikan kesimpulan. Setelah melihat jumlah kata positif maupun negatif nya selanjutnya adalah membandingkan mana yang lebih banyak. Jika jumlah kata positif lebih banyak maka tweet tersebut termasuk tweet positif, jika jumlah kata negatif lebih banyak maka tweet tersebut termasuk tweet negatif. Apabila jumlah keduanya sama maka dapat ditarik kesimpulan bahwa tweet tersebut netral.

IV. IMPLEMENTASI DAN ANALISIS

Implementasi dari algoritma tersebut adalah sebagai berikut.

1. Implementasi ke-1



Gambar 3. Data tes pertama

Selanjutnya teks twitter tersebut diklasifikasikan menggunakan sistem pengujian sehingga hasilnya adalah sebagai berikut.

	Positif	Negatif	Kesimpulan
Jumlah	0	2	Negatif

Tabel 4. Hasil analisis tes pertama

Dari tabel tersebut dapat dilihat bahwa twitter tersebut mengandung 2 kata negatif yaitu “jelek” dan “males” dan tweet tersebut tidak mengandung kata positif sehingga kesimpulannya adalah negatif.

2. Implementasi ke-2



Gambar 4. Data tes kedua

Selanjutnya teks twitter tersebut diklasifikasikan menggunakan sistem pengujian sehingga hasilnya adalah sebagai berikut.

	Positif	Negatif	Kesimpulan
Jumlah	1	0	Positif

Tabel 5. Hasil analisis tes kedua

Dari tabel tersebut dapat dilihat bahwa twitter tersebut mengandung 1 kata positif yaitu “keren” dan tweet tersebut tidak mengandung kata negatif. Sehingga kesimpulan yang dapat ditarik dari tabel tersebut adalah tweet tersebut merupakan tweet yang positif.

3. Implementasi ke-3



Gambar 5. Data tes ketiga

Selanjutnya teks twitter tersebut diklasifikasikan menggunakan sistem pengujian sehingga hasilnya adalah sebagai berikut

	Positif	Negatif	Kesimpulan
Jumlah	0	0	Netral

Tabel 5. Hasil analisis tes ketiga

Dari tabel tersebut dapat dilihat bahwa tidak ditemukan kata yang tergolong kedalam kata positif maupun negatif.

Sehingga kesimpulan yang dapat ditarik adalah tweet tersebut merupakan tweet yang netral.

V. KESIMPULAN

Kesimpulan yang dapat ditarik dari hasil implementasi dan analisa tersebut adalah dengan menggunakan algoritma pencocokan string Knuth-Morris-Pratt dan pengelompokan data klasifikasi dapat ditentukan suatu tweet termasuk kedalam tweet yang positif atau negatif. Kesimpulan tersebut diambil dari perbandingan jumlah kata positif atau negatif dalam tweet tersebut.

Namun dengan menggunakan algoritma ini hanya sebatas menyimpulkan sentimen dari jumlah kata positif atau negatifnya saja. Hal tersebut merupakan faktor internal. Faktor-faktor eksplisit masih belum dipertimbangkan seperti bagaimana apabila tweet tersebut mengandung kata negatif namun secara eksplisit sebenarnya tweet tersebut memiliki sentimen positif.

VI. SARAN

Saran untuk pengembangan selanjutnya yaitu :

1. Optimalkan faktor eksplisit
2. Tambahkan klasifikasi dengan cara pengelompokan data menggunakan algoritma statistika lain.

VII. UCAPAN TERIMAKASIH

Alhamdulillah, penulis mengucapkan rasa syukur kepada Allah SWT yang telah memberikan kesempatan untuk selalu berkarya dan bermanfaat bagi masyarakat. Limpahan serta curahan rahmat Nya begitu besar sehingga penulis masih bisa terus belajar, memperbaiki kesalahan dan terus mengembangkan diri lagi menjadi pribadi yang lebih bermanfaat untuk masyarakat. Penulis juga mengucapkan terimakasih kepada orang tua khususnya untuk Ibu yang tidak pernah berhenti berdoa dan memberikan support untuk penulis. Tidak lupa penulis juga mengucapkan terimakasih kepada dosen pembimbing yaitu Bapak Rinaldi Munir dan Ibu Nur Ulfa M yang telah memberikan bimbingan dan membagikan ilmunya kepada penulis. Ucapan terimakasih juga penulis sampaikan kepada pihak lain yang tidak bisa penulis sebutkan satu – persatu. Semoga apa yang telah kita lakukan hari ini dapat memberikan manfaat untuk Indonesia.

REFERENSI

- [1] Munir, Rinaldi. 2008. Diktat Kuliah IF2091 Matematika Diskrit. Bandung: Penerbit Informatika.
- [2] <http://www.unpas.ac.id/apa-itu-sosial-media/>. Diakses pada 5 Mei 2016, pukul 23.00 WIB.
- [3] <https://1pack.wordpress.com/2008/09/06/data-mining-klasifikasi-part-1/>. Diakses pada 08 Mei 2016, pukul 10.14 WIB.
- [4] <http://rofiah.blogspot.co.id/2013/12/makalah-klasifikasi-data.html>. Diakses pada 08 Mei 2016, pukul 10.16 WIB.
- [5] Sunni, Ismail dan Widyantoro, Dwi Hendratmo. 2012. Analisis Sentimen dan Ekstraksi Topik Penentu Sentimen pada Opini Terhadap

Tokoh Publik. Jurnal Sarjana Institut Teknologi Bandung Bidang Teknik Elektro dan Informatika. Volume 1, Number 2, halaman 200-206.

- [6] Darmawan, Najib. 2015. Algoritma Pencocokan String. Paper Strategi Algoritma.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 6 Mei 2016

A handwritten signature in black ink, appearing to read 'Ahmad Farhan Ghifari', written over a light grey rectangular background.

Ahmad Farhan Ghifari
13515602