

# Penerapan String Matching untuk Sensor Kata Bersifat Negatif pada Platform Forum KASKUS

Hilmi Naufal Yafie / 13517035

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

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

13517035@std.stei.itb.ac.id

**Abstrak**—Forum diskusi merupakan wadah untuk orang-orang yang memiliki minat yang sama dapat mendiskusikan suatu topik atau permasalahan sesuai dengan minat yang mereka miliki. Di jaman modern ini, sudah banyak bermunculan Forum diskusi online yang menghubungkan orang-orang dengan minat yang sama untuk saling berdiskusi tentang hal yang mereka senangi. Namun, tidak semua diskusi yang terjadi berjalan secara lancar. Terkadang, terdapat beberapa diskusi yang memanaskan hingga keluar kata-kata negatif dari para pendiskusi. Atau terjadi pula orang-orang yang memanfaatkan kebebasan berpendapat pada forum diskusi online untuk hal-hal yang negatif. Oleh karena itu, diperlukan adanya sensor pada setiap postingan konten yang ada pada forum diskusi online guna menjaga para pengguna melihat hal-hal yang dirasa negatif.

**Kata Kunci**—Forum, Diskusi, Negatif;

## I. PENDAHULUAN

Forum merupakan suatu wadah yang memfasilitasi berbagai orang yang memiliki suatu kesamaan minat dan tujuan untuk bersama-sama membahas suatu permasalahan atau topic secara bebas yang sesuai dengan minat atau tujuan dari orang-orang tersebut. Keberadaan forum merupakan salah satu tanda dari kebebasan berdemokrasi, dimana setiap perkataan yang disampaikan dalam suatu forum tidak dapat dibatasi oleh siapapun, namun tetap harus dapat dipertanggungjawabkan oleh yang mengatakannya. Di era saat ini, forum sudah bukan lagi sekadar orang-orang yang berada di dalam satu tempat berkumpul dan mendiskusikan suatu permasalahan atau topik, namun orang-orang yang berasal dari tempat yang jauh dan berbeda-beda namun memiliki minat yang sama dapat bertemu di forum diskusi online. Di Indonesia sendiri, forum diskusi online sudah banyak merambah dan bermunculan karena beragamnya minat yang dimiliki orang-orang Indonesia dan besarnya keinginan untuk berdiskusi dengan orang-orang yang memiliki kesamaan minat membuat forum diskusi online yang ada di Indonesia masih tetap ada hingga saat ini. Meskipun begitu, karena berbagai macam sifat dan kepribadian orang-orang benar-benar dibebaskan untuk ikut dalam suatu forum diskusi online, maka tidak selamanya forum yang ada membahas suatu topik secara baik-baik. Bisa saja terjadi perselisihan pendapat yang menyebabkan seseorang menuliskan kata-kata kasar, ataupun forum online digunakan untuk hal-hal yang negatif, seperti contohnya menyebarkan situs-situs pembajakan yang ada di internet. Oleh karena itu,

penyedia layanan forum diskusi online juga harus menyediakan sistem untuk dapat menyensor kata/kalimat negatif yang dapat merugikan orang lain maupun penyedia layanan itu sendiri serta guna menciptakan lingkungan forum diskusi online yang sehat dan bisa diterima oleh semua orang. Salah satu contoh forum diskusi online terbesar di Indonesia saat ini, KASKUS, telah menyediakan fitur ini, meskipun masih belum bisa dianggap sempurna.

## II. DASAR TEORI

### A. Forum Diskusi Online KASKUS

Platform forum diskusi online KASKUS didirikan pada tanggal 6 November 1999 oleh tiga orang mahasiswa asal Indonesia yang sedang melanjutkan studinya di Seattle, Amerika Serikat, mereka adalah Andrew Darwis, Ronald dan Budi. Pada awalnya, KASKUS dibuat untuk mempermudah orang-orang yang bertempat tinggal jauh dari Indonesia, terutama para mahasiswa Indonesia yang berkuliah di luar negeri, agar dapat mengetahui berita-berita tentang Indonesia saat itu. Hingga pada tahun 2008, para pembuat KASKUS memutuskan untuk mengelola KASKUS secara profesional dengan mengembangkan infrastruktur dan pengelolaan yang menunjang KASKUS hingga saat ini dapat menjadi salah satu forum diskusi online terbesar di Indonesia dengan jumlah pengguna lebih dari 10 juta pengguna dengan berbagai macam jenis pembahasan yang ada di dalamnya.

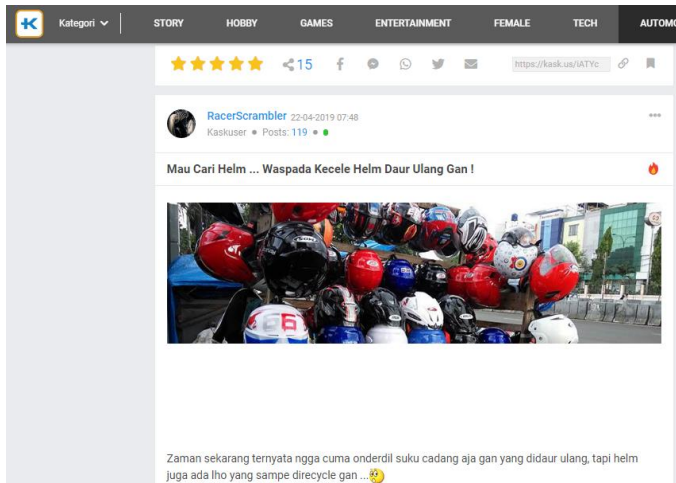


Gambar 2.1.1 Forum Diskusi KASKUS

Sumber: google.com

Dalam KASKUS, biasanya pengguna akan dibebaskan membuat sebuah Thread (isu atau topik yang ingin diangkat)

dengan tema apapun yang nantinya akan didiskusikan oleh para pengguna KASKUS yang berminat dengan Thread tersebut. Diskusi yang terjadi dapat berjalan dengan sangat baik seperti saling menimpali argument-argumen yang dibuat oleh para pengguna, namun tak jarang banyak pengguna yang pada akhirnya terbawa emosi ketika berada ditengah-tengah diskusi yang terjadi hingga keluarlah kata-kata negatif yang justru bisa memancing pengguna lain untuk saling serang. Selain itu, terdapat pula para pengguna yang menyalahgunakan kebebasan yang diberikan dengan membuat Thread yang berisi konten-konten yang bersifat negatif, seperti pembullying, penyebaran hoaks, ataupun penyebaran situs-situs illegal. Namun hal itu selalu berusaha dicegah oleh KASKUS dengan melakukan penyensoran pada konten-konten atau diskusi yang menggunakan kalimat-kalimat negatif.



Gambar 2.1.2 Contoh Forum Diskusi KASKUS  
Sumber: google.com

### B. String Matching

Algoritma *string matching* merupakan suatu algoritma yang bertujuan untuk melakukan pengecekan suatu kata atau sub-kalimat apakah terdapat pada kumpulan kalimat atau sebuah text yang dimiliki. Kata atau kalimat yang ingin dicek biasanya merupakan sebuah input, baik itu input langsung dari user atau input dari file. Ketentuan dari string matching adalah sebagai berikut:

- Panjang kata atau sub-kalimat (*pattern*) yang ingin dicocokkan adalah  $n$  huruf
- Panjang kumpulan kalimat atau teks yang dimiliki adalah  $m$  huruf
- $n \ll m$  (besar dari *pattern* dengan panjang  $n$  huruf cukup kecil jika dibandingkan dengan besar kalimat atau teks dengan panjang  $m$  huruf)

Algoritma string matching memiliki beberapa macam jenis, seperti algoritma Brute Force, algoritma Knutt-Morris-Pratt, algoritma Boyer-Moore, ataupun Regex. Masing-masing dari jenis algoritma *string matching* tersebut memiliki keunggulannya tersendiri.

### C. Algoritma Knutt-Morris-Pratt (KMP)

Algoritma Knutt-Morris-Pratt, atau sering disingkat dengan sebutan KMP, merupakan salah satu dari jenis algoritma *string matching*. Algoritma KMP merupakan algoritma string matching yang melakukan pengecekan secara terurut dari kiri ke kanan. Jika dibandingkan dengan algoritma Brute Force, algoritma KMP akan terlihat sama dalam kompleksitas waktunya, namun sebenarnya algoritma KMP memiliki kompleksitas waktu yang lebih efisien dibanding Brute Force jika mengacu pada perhitungan komparasi karakter yang dilakukan.

Misal terdapat sebuah *pattern* dan kalimat sebagai berikut:  
Kalimat: “Di Indonesia, kantor KASKUS pertama berlokasi di daerah Mangga Besar, yang dibantu dengan 2 orang tenaga profesional”  
*Pattern*: “KASKUS”

Maka pencarian tersebut akan menghasilkan true dengan kata KASKUS yang ditemukan dimulai dari karakter ke-22.

Meskipun begitu, tidak selamanya algoritma KMP akan menghasilkan kompleksitas waktu yang lebih efisien. Karena pengecekan pada algoritma KMP selalu dimulai dari awal kata, maka untuk kalimat atau teks yang banyak memiliki awal kata yang mirip namun bukan kata yang dicari akan menghasilkan pencocokan karakter string yang kurang efisien.

### D. Sensor

Sensor merupakan suatu pengawasan dan pemeriksaan pada surat-surat atau sesuatu yang akan disiarkan ke publik melalui media-media yang tersedia, seperti berita, buku, internet (kbbi.web.id). Sensor diberikan untuk mencegah hal-hal negatif yang dirasa terdapat pada sesuatu yang disiarkan melalui media yang ada sampai kepada masyarakat. Dalam hal ini, misal jika suatu konten yang terdapat pada internet mengandung unsur negatif, akses terhadap konten tersebut akan sangat sulit untuk dibatasi oleh siapa saja yang dapat mengaksesnya, oleh karena itu sensor diberikan terhadap konten tersebut sehingga siapapun yang mengaksesnya tidak dapat melihat unsur negatif yang ada di dalam konten yang dapat mempengaruhi aksesnya.

## III. CARA KERJA PENYENSORAN THREAD KASKUS

### A. Pembuatan Thread

Ketika membuat Thread, pengguna diharuskan untuk memiliki akun KASKUS, sehingga setiap Thread yang ditampilkan di KASKUS merupakan buatan seorang pengguna KASKUS yang bisa dipertanggungjawabkan isinya.

Langkah dalam membuat Thread di KASKUS adalah sebagai berikut:

Gambar 3.1.1 Form Pengisian Thread KASKUS  
Sumber: kaskus.co.id

- Klik “Buat Post / Tulis Thread” untuk membuat thread baru.
- Klik “Pilih Forum” untuk menentukan forum yang paling sesuai dengan topic thread.
- Cantumkan judul thread pada Title Box. (Max 85 Karakter)
- Pilih ikon yang melambangkan isi thread di Ikon Box (Optional)
- Isi thread ditulis di Body Content
- Untuk mengepost konten yang telah dibuat klik “Post Thread”

Thread yang telah dibuat nantinya akan ditampilkan di dalam forum yang sesuai dengan pilihan forum yang diisi pada form pembuatan Thread.

Sebuah thread juga bisa ditambahkan gambar untuk mencegah adanya hoaks di thread yang dibuat, dengan cara:

Gambar 3.1.2 Upload Gambar

Sumber: kaskus.co.id

- Klik “Buat Post / Upload Gambar” untuk Upload Gambar baru
- Tarik Gambar atau Klik untuk memilih gambar
- Cantumkan judul thread pada Title Box
- Masukkan Deskripsi gambar
- Masukkan Tags yang sesuai dengan gambar
- Klik “Upload Gambar”

#### B. Penyensoran pada Isi Thread

Ketika pengguna mempost suatu thread, maka thread yang dibuat tersebut akan langsung ditampilkan kepada pengguna lainnya. Setiap thread yang dibuat oleh para pengguna akan selalu masuk pengetesan string matching dengan database kata-kata atau kalimat negatif yang dimiliki oleh KASKUS, sehingga setiap post thread yang dibaca pengguna lainnya tidak akan ditemui kata-kata negative tersebut. Secara umumnya, cara kerja dari pengecekan thread yang dilakukan oleh KASKUS dapat dijelaskan dengan algoritma KMP pada pseudo code sebagai berikut:

```

procedure KMPSearch(string pat, string
txt, string repl)
{
    int M = pat.length;
    int N = txt.length;

    // create lps[] that will hold the

```

```

longest prefix suffix
// values for pattern
int lps[M];

// Preprocess kata (calculate lps[]
array)

computeLPSArray(pat, M, lps);

int i = 0; // index for txt[]
int j = 0; // index for pat[]
while (i < N) {
    if (pat[j] == txt[i]) {
        j++;
        i++;
    }

    if (j == M) {
        j = lps[j - 1];
        //replace text
        for (k=j;k>0;k++) {
            txt[i-k] = repl[j-k];
        }
    }

    // mismatch after j matches
    else if (i < N && pat[j] !=
txt[i]) {
        // Do not match lps[0..lps[j-
1]] characters,
        // they will match anyway
        if (j != 0)
            j = lps[j - 1];
        else
            i = i + 1;
    }
}

// Fills lps[] for given patttern
pat[0..M-1]
void computeLPSArray(string pat, int M,
int[] lps)
{
    // length of the previous longest
prefix suffix
    int len = 0;

    lps[0] = 0; // lps[0] is always 0

    // the loop calculates lps[i] for i =
1 to M-1
    int i = 1;
    while (i < M) {
        if (pat[i] == pat[len]) {
            len++;
            lps[i] = len;
            i++;

```

```

        }
        else // (pat[i] != pat[len])
        {
            // This is tricky. Consider
the example.
            // AAACAAAA and i = 7. The
idea is similar
            // to search step.
            if (len != 0) {
                len = lps[len - 1];

                // Also, note that we do
not increment
                // i here
            }
            else // if (len == 0)
            {
                lps[i] = 0;
                i++;
            }
        }
    }

// Driver program to test above function
int main()
{
    Konten = readFile(Konten);
    kataNegatif =
readFile(DatabaseKataNegatif);
    char replace = "***Sensor***";

    char txt[] = Konten.read();
    char pat[][];
    for (line in kataNegatif) {

        pat[line][] =
kataNegatif.readLine();
        KMPSearch(pat, txt, replace);
    }
    return 0;
}

```

Pada pseudo code tersebut, bisa dilihat ketika pada konten yang dibuat jika terdapat kata yang bersifat negatif maka akan direplace dengan sensor, sehingga kata-kata negatif yang ditampilkan di thread akan diganti dengan **\*\*Sensor\*\***. Dengan begitu, pengguna lain tidak akan dapat membaca kata-kata negatif pada setiap thread yang ditampilkan.

Selain itu, pengguna yang terlalu banyak menggunakan kata-kata negatif pada setiap konten/thread yang dibuatnya akan mendapat teguran karena seringnya menggunakan kata-kata negatif. Jika sudah diperingati namun tetap saja menggunakan kata-kata negatif pada setiap konten/thread yang dibuat, maka akun dari pengguna tersebut akan di banned dari KASKUS. Dengan begitu, tidak sembarangan orang akan membuat konten dengan isi yang negatif di dalamnya.

```
Inilah daftar kata yang akan disensor dan autocorrect jika agan tulis di kaskus. coba aja sendiri

Forbidden / sensor
[apapun].co.cc -> [apapun]*Forbidden*
ceriwis.us -> *Forbidden*
knowyourmeme -> ***
bangsat -> ** SENSOR **
dinomarket -> ***
kikil-> ***forbidden***
anakayam -> ***forbidden***
duniasex -> ***forbidden***
glodokshop -> *banned*/QUOTE]

autocorrect
[QUOTE]meganil-> megawati
kaskus.us -> kaskus.co.id
ueuo -> yahoo
paypal -> paypal
kawin -> kimpoi
cracker -> kracker
ziddu.com -> zid*du.com
rapidshare.com -> rapid*share.com
4shared.com -> 4*shared.com
selapa -> kaskus
[URL=#]www.tokopedia.com [/URL]-> www.tokopedia.com[/QUOTE]
```

Gambar 3.2.1 Forum Diskusi KASKUS  
Sumber: kaskus.co.id

### C. Pengawasan Pengguna

Kaskus memberikan pengawasan terhadap para penggunanya dengan adanya peraturan-peraturan. Meskipun Para pengguna diberi kebebasan untuk membuat thread apapun yang dimau, tapi para pengguna juga harus bertanggung jawab untuk setiap thread yang dibuatnya.

Sebagai pengguna, kita dapat membantu KASKUS dalam melakukan pengawasan pada konten-konten yang ada di KASKUS dengan mereport thread/konten yang dirasa lolos dari pengecekan konten/thread KASKUS. Selain itu, dapat juga dilakukan report terhadap akun yang dirasa meresahkan dalam setiap pembuatan thread/kontennya. Termasuk thread yang dirasa melanggar hukum. Dengan begitu, forum yang berada di KASKUS akan menjadi lebih terjaga dengan sistem yang ada serta bantuan dari para pengguna KASKUS.

## IV. ANALISIS

Konten atau thread yang ada di KASKUS dapat menyensor kata-kata negatif bagi para pengguna dengan menggunakan string matching. Dengan mencocokkan setiap kata yang ada di database kata kasar yang dimiliki KASKUS, kata-kata kasar yang ditemukan melalui string matching lalu di replace dengan kata **\*\*Sensor\*\*** guna memberi tahu kepada pengguna KASKUS yang membaca thread tersebut bahwa kata-kata yang digunakan digunakan oleh pembuat thread pada threadnya merupakan kata-kata negatif.

## V. KESIMPULAN

Para pengguna dapat membantu KASKUS dalam memperbaiki sistem sensor dengan string matching dengan cara memberikan saran ke KASKUS daftar kata-kata negatif yang ditemukan pengguna dalam thread yang ada di KASKUS dan masih tidak tersensor.

## VI. REFERENSI

- [1] <http://informatika.stei.itb.ac.id/~rinaldi.munir/Stmik/2018-2019>
- [2] <https://www.kanalinfo.web.id/pengertian-forum>
- [3] <https://bantuan.KASKUS.co.id/hc/id/articles/214603738-Sejarah-KASKUS>
- [4] <https://piptools.net/algoritma-kmp-knuth-morris-pratt/>

## VII. 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, 26 April 2019

ttd

Hilmi Naufal Yafie/13517035