

PENCARIAN INFORMASI BUAH DAN SAYURAN LOKAL MENGUNAKAN METODE *BOYER MOORE*

Afin Findi Ariyanto, Daryanto

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

Afin12385@gmail.com, daryanto@unmuhjember.ac.id

ABSTRAK

Ensiklopedia sudah di kenal oleh kalangan pelajar sebagai media untuk mendapatkan informasi tentang topik tertentu yang di inginkan. Ensiklopedia buah dan sayuran lokal Jember digunakan oleh instansi pendidikan, dan masyarakat umum sebagai sumber pembelajaran. Oleh karena itu, penelitian ini bertujuan untuk membangun aplikasi Ensiklopedia buah dan sayuran lokal Jember berbasis web secara akurat. Algoritma *Boyer Moore* adalah algoritma pencarian *String* yang paling efektif saat ini. Penelitian ini membahas tentang bagaimana Algoritma *Boyer Moore* melakukan pencarian terhadap ensiklopedia buah dan sayur lokal daerah Jember. Penelitian akan membahas hasil implementasi algoritma *Boyer Moore* pada data buah dan sayuran lokal jember, tingkat akurasi tertinggi yang dihasilkan dari pengujian terhadap algoritma *Boyer Moore* dan waktu tercepat yang dihasilkan dari pencarian menggunakan algoritma *Boyer Moore*. Data yang digunakan pada penelitian ini sebanyak 125 data, dimana jumlah data buah sebanyak 74 dan data sayur sebanyak 51. Dari hasil tujuh kali pengujian yang dilakukan pada penelitian ini, dimana pengujian dilakukan pada waktu pencarian, tingkat kecocokan, kategori, jenis buah dan manfaat terhadap sesuatu. Kata kunci yang digunakan pada pencarian ini adalah “sayur, “buah”, “apel”, “pisang”, “pear”, “anggur” dan “mencegah kanker”. Hasil pengujian yang diperoleh adalah Algoritma *Boyer Moore* memperoleh tingkat akurasi tertinggi yaitu 100% sebanyak tiga kali dari tujuh kali percobaan dengan waktu pencarian tercepat yaitu 0,18 detik dan paling lama 3,37 detik.

Kata kunci : Klasifikasi, Ensiklopedia, buah, sayur, Jember, *Boyer Moore*

BAB I PENDAHULUAN

1.1. Latar belakang

Ensiklopedia sudah di kenal oleh kalangan pelajar sebagai media untuk mendapatkan informasi tentang topik tertentu yang di inginkan. Kebanyakan produk Ensiklopedia di pasaran dalam bentuk buku, majalah, atlas dan kartu. Produk ensiklopedia fisik tersebut telah digunakan dalam proses pembelajaran siswa-siswi di sekolah dan masyarakat umum. Akan tetapi produk ensiklopedia fisik memiliki keterbatasan dalam kemudahan akses dan kecepatan penyebaran informasi serta bersifat statis (Suharso, 2017). Algoritma *Boyer Moore* adalah algoritma pencarian *String* yang paling efektif saat ini. Algoritma yang ditemukan oleh Bob *Boyer* dan J. Strother *Moore* ini telah menjadi standar untuk berbagai literatur pencarian *String*. Algoritma *Boyer Moore* akan menyimpan informasi pergeseran untuk melakukan pencarian *String*.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah di paparkan, masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana hasil implementasi algoritma *Boyer Moore* pada data buah dan sayuran lokal jember?
2. Berapa tingkat akurasi tertinggi yang dihasilkan dari pengujian terhadap algoritma *Boyer Moore*?
3. Berapa waktu tercepat yang dihasilkan dari pencarian menggunakan algoritma *Boyer Moore*?

1.3. Tujuan

Adapun tujuan penelitian sebagai berikut :

1. Mengimplementasikan metode *Boyer Moore* terhadap data buah dan sayur sehingga menjadi ensiklopedia buah dan sayuran lokal Jember berbasis web sebagai sumber pembelajaran.
2. Mengetahui tingkat akurasi tertinggi dari metode *Boyer Moore*.
3. Mengetahui tingkat kecepatan metode *Boyer Moore*.

1.4. Manfaat

Adapun manfaat penelitian sebagai berikut :

1. Kalangan pelajar dapat menggunakan produk aplikasi ensiklopedia tentang buah dan sayuran lokal Jember sebagai sumber pembelajaran konten lokal secara mudah dan cepat.
2. Membantu pelajar dalam mendapatkan informasi tentang topik buah dan sayuran lokal jember dalam beragam jenis seperti teks dan gambar.

1.5. Batasan Masalah

Batasan-batasan dalam penelitian ini sebagai berikut.

1. Metode pencarian informasi berdasarkan kata kunci yang akan di lakukan *String Matching* dengan metode *Boyer Moore*.
2. Dataset penelitian menggunakan data buah dan sayuran lokal Jember yang di dapat dari penelitian ANIK ANDRIANI (1310211030) dan KUTSIATUL HIDAYAH (1310211006).
3. Jumlah data yang digunakan dalam penelitian ini sebanyak 125 data.
4. Pada penelitian ini dilakukan pengujian waktu yang dihasilkan, hasil pencarian yang benar, data seharusnya yang benar, hasil pencarian yang salah, data yang tidak terseleksi dengan benar, tingkat kemiripan, kebenaran data yang ditampilkan dan tingkat akurasi.

BAB II KAJIAN PUSTAKA

2.1. Data Buah dan Sayur

Sawitri (2017) menyatakan bahwa Buah dan Sayuran lokal Jember merupakan merupakan bahan pangan utama dalam kehidupan sehari-hari. Jenis buah dan sayuran ini memiliki kurang lebih 107 macam jenis yang keanekaragamannya sangat bervariasi. Keanekaragaman warna pada buah bukanlah sekedar pembeda jenis antar buah yang satu dengan yang lainnya. Data Buah dan Sayur ini sudah banyak di sajikan berupa Buku Bacaan, Majalah, Atlas dan lain-lain.

2.2. Web Ensiklopedia

Mauludin (2012) menyatakan bahwa Web Ensiklopedia yang berbasis Teknologi Informasi (TI) merupakan tempat untuk mencari informasi sesuai kebutuhan sehari-hari, sehingga pengunjung web tersebut dapat mengendalikan situs tersebut sesuai fitur-fitur yang di sediakan. Beberapa situs website yang menyediakan layanan yang mengandung Ensiklopedia sebagai berikut: <http://www.wikipedia.org> wikipedia.org merupakan situs yang mengandung layanan Ensiklopedia,. Ensiklopedia bebas berbahasa Indonesia dan di bangun oleh para sukarelawan. serta gratis. Salah satu keunggulannya adalah di sediakan fasilitas pencarian, seperti situs search engine.

2.3. String Matching

String adalah susunan dari karakter-karakter (Angka, alphabet, atau karakter yang lain) dan biasanya direpresentasikan sebagai struktur data array (Syaroni dan Munir, 2004:1). String dapat berupa kata, frase, atau kalimat. Sedangkan *String Matching* diartikan sebagai sebuah permasalahan untuk menemukan pola susunan karakter *String* didalam *String* lain atau bagian dari isi teks (Syaroni dan Munir, 2004:1). *String Matching* dalam bahasa Indonesia dikenal dengan istilah pencocokan *String* (Munir, 2007:1).

Pencarian *String* yang juga bisa disebut pencocokan *String* (*String Matching*) merupakan algoritma untuk melakukan pencarian semua kemunculan *String* pendek $[0...n-1]$ yang disebut pattern di *String* yang lebih panjang $[0...m-1]$ yang disebut teks (Charras, 1997: 11).

2.4. Boyer Moore

Algoritma *Boyer Moore* adalah salah satu algoritma pencarian *String*, dipublikasikan oleh Robert S.Boyer dan J. Strother Moore pada tahun 1977. Algoritma ini tidak seperti algoritma pencarian *String* yang ditemukan sebelumnya, algoritma *Boyer Moore* mulai mencocokkan karakter dari sebelah kanan pattern. Ide dibalik algoritma ini adalah bahwa dengan memulai pencocokkan karakter dari kanan, maka akan lebih banyak informasi yang didapat (Chiquita, 2012).

Algoritma *Boyer Moore* menggunakan dua buah tabel untuk mengolah informasi saat terjadi kegagalan pencocokkan pattern. Tabel pertama disebut dengan istilah bad character shift atau occurrence heuristic. Tabel kedua disebut dengan istilah good suffix shift atau match heuristic (Guidio, 2014).

Algoritma *Boyer Moore* telah dibuktikan sebagai salah satu algoritma yang paling efisien dalam aplikasi pencarian *String* dengan menggunakan natural language (bukan binary language). Algoritma ini telah sering diimplementasikan untuk fungsi "Search" dan "Substitute" pada text editor. Padasarnya cara kerja algoritma ini mirip dengan algoritma Knuth-MorrisPrat (KMP) dimana kedua algoritma ini akan melakukan lompatan pengecekan dalam proses pencarian *String*. Namun berbeda dengan algoritma KMP, algoritma *Boyer Moore* ini melakukan perbandingan pattern mulai dari kanan ke kiri. (Evlyn Dwi Tambun, 2010) Idea of the *Boyer Moore* Algorithm The algorithm uses knowledge of the search text to improve search speed significantly. It uses a step of preprocessing to create an occurrence-

function and shifting function which are used to perform bad character heuristics and good suffix heuristics respectively. (Prabhakar Gupta, et al. 2010).

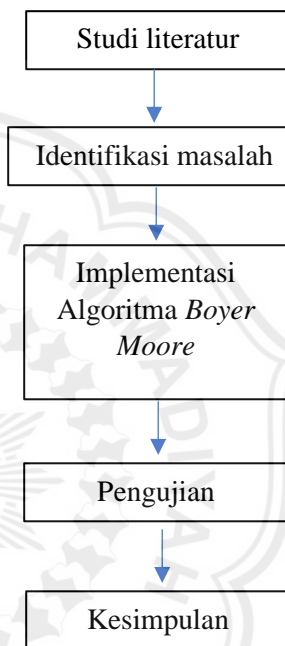
2.5. Confusion Matrix

Menurut Han dan Kamber (2011:365) Confusion matrix adalah alat yang berguna untuk menganalisis seberapa baik Classifier mengenali tuple dari kelas yang berbeda. TP dan TN memberikan informasi ketika Classifier benar, sedangkan FP dan FN memberitahu ketika Classifier salah.

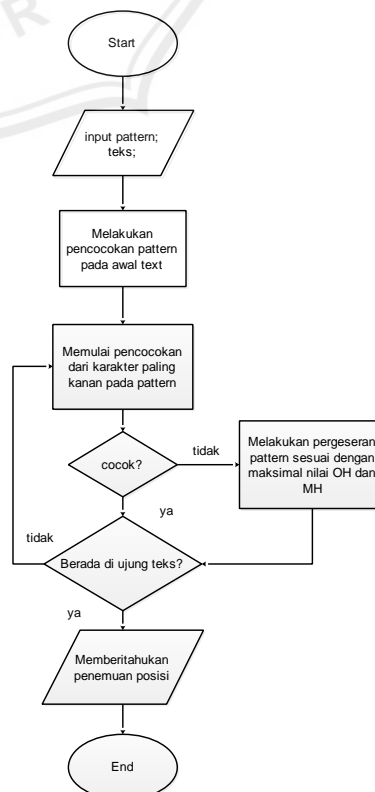
| | | Klasifikasi | |
|--------|---|-------------|----|
| | | Y | N |
| Aktual | Y | TP | FN |
| | N | FP | TN |

BAB III METODOLOGI PENELITIAN

3.1. Tahapan Penelitian



3.2. Implementasi Algoritma Boyer Moore



3.3. Pengujian

Pengujian pada tahap ini meliputi pengujian waktu pencarian atau time remining serta pencarian berdasarkan jenis, klasifikasi, kandungan dan manfaat.

Berikut poin-poin yang akan diujikan pada pencarian *Boyer Moore*:

- Kata/kalimat pencarian** adalah kata atau kalimat yang menjadi *keyword* untuk pencarian. Pada pengujian ini ada 4 *keyword* yang akan diujikan yaitu jenis, klasifikasi, kandungan dan manfaat.
- Jumlah data** adalah jumlah data keseluruhan
- Hasil pencarian yang benar** adalah hasil yang muncul dan terkoreksi benar.
- Data seharusnya yang benar** adalah kondisi sebenarnya pada data.
- Data yang tidak terseleksi dengan benar** adalah data yang tak boleh muncul pada saat pencarian karena tidak dimaksud dalam seleksi atau pencarian.
- Waktu** adalah *time remining* atau waktu yang dibutuhkan metode dalam melakukan pencarian.
- Keterangan berisi keputusan hasil pencarian yaitu tepat dan tidak tepat.**
- Akurasi** adalah nilai akurasi dari metode *Boyer Moore* pada sebuah kasus pencarian. Persamaan yang digunakan adalah $\frac{(TP+TN)}{total\ data}$
- Tingkat kecocokan** diperoleh dari hasil pencarian yang ditemukan dibagi jumlah kata dalam data tersebut, atau dapat dituliskan menjadi $tingkat\ kecocokan = \frac{pattern}{total\ kata\ dalam\ data} \times 100\%$. Pattern adalah kata kunci yang dimasukkan, dan total kata adalah kata yang berada pada tiap data.

BAB IV IMPLEMENTASI DAN HASIL

4.1. Gambaran Data

- Kategori**
Pada data penelitian ini data terdiri dari buah dan sayur dengan total data 125. Pada dataset kategori buah berjumlah 74 data dan kategori sayur berjumlah 51 data.
- Nama**
Data buah dan sayur ini juga terklasifikasi berdasarkan nama. Berikut Tabel nama buah dan sayur.

4.2. Pengujian

- Kategori Buah**
Dari total 125 data pada pencarian dengan kata kunci “buah” menghasilkan 99 data. Dari 99 data tersebut 73 data terkategori buah dan itu bernilai benar (TP) dan 1 data dari kategori buah tidak tampil. 26 data dari 99 data adalah sayur, dimana ini adalah data yang dianggap salah atau tidak dimaksud dalam pencarian yang ditampilkan oleh hasil pencarian. Dari 125 total data terdapat 25 data yang tidak tampil dan bukan berkategori buah, hal ini berarti benar atau true negative

(TN). Dari akumulasi tersebut dapat diketahui tingkat akurasi pada pencarian kategori buah yaitu dengan menggunakan persamaan $\frac{(TP+TN)}{total\ data} \times 100\%$ yaitu $\frac{(73+25)}{125} \times 100\% = 78,4\%$.

- Kategori Sayur**
Dari total 125 data pada pencarian dengan kata kunci “sayur” menghasilkan 26 data. Dari 26 data tersebut 25 data terkategori sayur dan itu bernilai benar (TP) dan 1 data dari kategori sayur tidak tampil. 1 data dari 26 data adalah buah, dimana ini adalah data yang dianggap salah atau tidak dimaksud dalam pencarian yang ditampilkan oleh hasil pencarian. Dari 125 total data terdapat 73 data yang tidak tampil dan bukan berkategori buah, hal ini berarti benar atau true negative (TN). Dari akumulasi tersebut dapat diketahui tingkat akurasi pada pencarian kategori buah yaitu dengan menggunakan persamaan $\frac{(TP+TN)}{total\ data} \times 100\%$ yaitu $\frac{(25+73)}{125} \times 100\% = 78,4\%$.
- Jenis Pisang**
Dari total 125 data pada pencarian dengan kata kunci “pisang” menghasilkan 9 data yang ditampilkan. Dari 9 data tersebut 8 data terkategori jenis pisang dan itu bernilai benar (TP). 1 data berkategori buah sukun, dimana ini adalah data yang dianggap salah atau tidak dimaksud dalam pencarian yang ditampilkan oleh hasil pencarian. Dari 125 total data terdapat 116 data yang tidak tampil dan bukan berkategori buah pisang, hal ini berarti benar atau true negative (TN). Dari akumulasi tersebut dapat diketahui tingkat akurasi pada pencarian kategori buah yaitu dengan menggunakan persamaan $\frac{(TP+TN)}{total\ data} \times 100\%$ yaitu $\frac{(8+116)}{125} \times 100\% = 99,2\%$.
- Jenis Pear**
Dari total 125 data pada pencarian dengan kata kunci “pear” menghasilkan 6 data yang ditampilkan. Dari 6 data tersebut keseluruhan data terkategori jenis pear dan itu bernilai benar (TP). Dari 125 total data terdapat 119 data yang tidak tampil dan bukan berkategori buah pisang, hal ini berarti benar atau true negative (TN). Dari akumulasi tersebut dapat diketahui tingkat akurasi pada pencarian kategori buah yaitu dengan menggunakan persamaan $\frac{(TP+TN)}{total\ data} \times 100\%$ yaitu $\frac{(6+119)}{125} \times 100\% = 100\%$.
- Jenis Apel**
Dari total 125 data pada pencarian dengan kata kunci “apel” menghasilkan 7 data yang ditampilkan. Dari 7 data tersebut keseluruhan data terkategori jenis apel dan itu bernilai benar (TP). Dari 125 total data terdapat 118 data yang tidak tampil dan bukan berkategori buah pisang, hal ini berarti benar atau true negative (TN). Dari akumulasi tersebut dapat diketahui tingkat

akurasi pada pencarian kategori buah yaitu dengan menggunakan persamaan $\frac{(TP+TN)}{total\ data} \times 100\%$ yaitu $\frac{(7+118)}{125} \times 100 = 100\%$.

6. Jenis Anggur

Dari total 125 data pada pencarian dengan kata kunci “anggur” menghasilkan 4 data yang ditampilkan. Dari 4 data tersebut keseluruhan data terkategori jenis anggur dan itu bernilai benar (TP). Dari 125 total data terdapat 121 data yang tidak tampil dan bukan berkategori buah pisang, hal ini berarti benar atau true negative (TN). Dari akumulasi tersebut dapat diketahui tingkat akurasi pada pencarian kategori buah yaitu dengan menggunakan persamaan $\frac{(TP+TN)}{total\ data} \times 100\%$ yaitu $\frac{(4+121)}{125} \times 100 = 100\%$.

7. Pengujian pada manfaat

Dari total 125 data pada pencarian dengan kata kunci “mencegah kanker” menghasilkan 54 data yang ditampilkan. Dari 54 data tersebut 33 data bermanfaat mencegah kanker dan itu bernilai benar (TP). Dari 125 total data terdapat 71 data yang tidak tampil dan bukan berkategori buah pisang, hal ini berarti benar atau true negative (TN). Dari akumulasi tersebut dapat diketahui tingkat akurasi pada pencarian kategori buah yaitu dengan menggunakan persamaan $\frac{(TP+TN)}{total\ data} \times 100\%$ yaitu $\frac{(33+71)}{125} \times 100\% = 82,3\%$.

4.3. Hasil Analisis

Berikut ringkasan hasil pengujian yang dilakukan dalam penelitian ini.

Tabel Ringkasan hasil pengujian

| No | Kata/kalimat pencarian | Jumlah data | Hasil pencarian yang benar | Data sehas yang benar | Hasil Pencarian yang salah | Data yang tidak terselensi dengan benar | Waktu | akurasi (%) |
|----|------------------------|-------------|----------------------------|-----------------------|----------------------------|---|------------|-------------|
| 1 | buah | 125 | 73 | 74 | 26 | 25 | 3.37 detik | 78.4 |
| 2 | sayur | 125 | 25 | 26 | 1 | 73 | 0.93 detik | 78.4 |
| 3 | Pisang | 125 | 8 | 8 | 1 | 116 | 0.52 detik | 99.2 |
| 4 | pear | 125 | 6 | 6 | 0 | 119 | 0.24 detik | 100 |
| 5 | apel | 125 | 7 | 7 | 0 | 118 | 0.24 detik | 100 |
| 6 | anggur | 125 | 4 | 4 | 0 | 121 | 0.18 detik | 100 |
| 7 | mencegah | 125 | 33 | 33 | 21 | 71 | 1.87 detik | 83.2 |

| | | | | | | | | |
|--------|--|--|--|--|--|--|--|--|
| kanker | | | | | | | | |
|--------|--|--|--|--|--|--|--|--|

Dari beberapa pengujian yang dilakukan di atas, peneliti memperoleh beberapa hal dan juga temuan diantaranya:

1. Algoritma *Boyer Moore* mampu melakukan pencarian yang diinginkan pada dataset buah dan sayur dengan tepat.
2. Algoritma *Boyer Moore* memiliki waktu pencarian yang relatif cepat dengan waktu pencarian terlama yaitu 3,37 detik dan tercepat 0,18 detik pada dataset berjumlah 125 data dan 9 atribut.
3. Algoritma *Boyer Moore* memiliki tingkat akurasi tinggi dengan akurasi tertinggi yaitu 100% dan terendah 78,4%.

Hasil pencarian yang tidak sesuai dan ditampilkan oleh aplikasi tidak sepenuhnya salah karena pada deskripsi juga mengandung kata yang dicari. Contohnya dalam pencarian “buah”, hasil pencarian juga menampilkan petai cina yang tergolong sayur tetapi dalam deskripsi petai cina terdapat kalimat “*Buah polong (legumen) berbentuk pipih dan tipis berukuran panjang 14-22 cm lebar 1,5-2 cm dengan sekat diantara biji*” sehingga *Boyer Moore* juga memunculkannya dalam hasil pencarian.

BAB V PENUTUP

5.1. Kesimpulan

Dari 3 kali pengujian implementasi pencarian menggunakan metode *Boyer Moore* terhadap 125 data buah dan sayur dengan pengujian melibatkan kategori, jenis dan manfaat, peneliti memperoleh beberapa hasil yaitu:

1. Algoritma *Boyer Moore* mampu melakukan pencarian data buah dan sayur.
2. Algoritma *Boyer Moore* memiliki tingkat akurasi cukup tinggi. Dari tujuh kali uji coba, tingkat akurasi terendah yaitu 78,4% dan tertinggi 100%.
3. Waktu pencarian yang dilakukan oleh algoritma *Boyer Moore* juga relatif cepat. Dari tujuh kali uji coba tercatat waktu tercepat yaitu 0,18 detik dan paling lama 3,37 detik.

5.2. Saran

Beberapa aspek yang dapat dikembangkan dalam penelitian ini yaitu:

1. Pengembangan dapat menambahkan metode pencarian lain untuk mengakomodasi kesalahan sistem saat menampilkan data.
2. Pengembang dapat menambahkan data lebih banyak agar *time remaining* dari algoritma *Boyer Moore* dapat diketahui keakuratannya.

Daftar Pustaka

- Boyer, Robert. 2015. “Algoritma *Boyer Moore*”. https://id.wikipedia.org/wiki/Algoritme_Boyer-Moore
- Gntur H., Achmad., Octavianto, Hardian., Daryanto. 2018. Aplikasi Deteksi Kemiripan Judul Berbahasa Indonesia Menggunakan *Boyer Moore*. Universitas Muhammadiyah Jember.
- Hadjarat, Srinaning. 2015. “Penerapan *String Matching* Pada Aplikasi Berbasis

Web Di Jurusan Teknik Elektro Fakultas Teknik Universitas Negeri Gorontalo”, Gorontalo.[http://www.academia.edu/27858705/PENERAPAN_STRING_MATCHING_PADA APLIKASI EARSIP BERBASIS WEB DI JURUSAN TEKNIK ELEKTRO FAKULTAS TEKNIK UNIVERSITAS NEGERI GORONTALO](http://www.academia.edu/27858705/PENERAPAN_STRING_MATCHING_PADA_APLIKASI_EARSIP_BERBASIS_WEB_DI_JURUSAN_TEKNIK_ELEKTRO_FAKULTAS_TEKNIK_UNIVERSITAS_NEGERI_GORONTALO)

Jhoni, 2016. “Pengertian *String Matching*”.
<http://repository.usu.ac.id/bitstream/handle/123456789/59859/Chapter%20II.pdf?sequence=4&isAllowed=y>

Leonaerde, Guidio. 2014. Penerapan algoritma *Boyer Moore* Pada Aplikasi Pengajuan Judul Skripsi Berbasis Web.

Mu’Awanah, Asifatuul. 2018. Implementasi Algoritma *Boyer Moore* Pada Pencarian Data Di Sistem Informasi Manajemen Laboratorium Fakultas

Sains Dan Teknologi.Skripsi.Jurusan Teknik Informatika Fakultas Sains Dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Rahmanita, E., 2014, Pencarian *String* Menggunakan Algoritma *Boyer Moore*. Pada Dokumen, Madura, Jurusan Teknik Informatika:Universitas Trunojoyo.

Sawitri, Komarayanti. 2017. “Ensiklopedia Buah-buahan Lokal Jember Berbasis Potensi Alam Jember”. p-ISSN 2527-7111; e-ISSN 2528-1615
<http://jurnal.unmuhjember.ac.id/index.php/BIO/MA/article/download/591/470>

Shofan, Aldyaka, 2016. *Pengembangan Algoritma Boyer Moore*.

[http://www.academia.edu/19848314/Pengembangan Algoritma Boyer Moore](http://www.academia.edu/19848314/Pengembangan_Algoritma_Boyer_Moore)

