

Penggunaan Struktur *FP-Tree* dan Algoritma *FP-Growth* dalam Rekomendasi Promosi Produk pada Situs Belanja Online

Irene Edria Devina / 13515038¹

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

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

¹13515038@itb.ac.id

Abstrak—Rekomendasi promosi produk merupakan suatu bentuk promosi yang baik dan efektif. Rekomendasi promosi produk menggunakan data hasil observasi terhadap keadaan dan keinginan pelanggan untuk membeli suatu produk. Dengan menemukan pola barang yang dibeli oleh pelanggan, maka promosi suatu produk akan lebih tepat sasaran. Salah satu pendekatan yang digunakan adalah pendekatan analisis asosiasi dengan menggunakan Struktur *FP-tree* dan Algoritma *FP-growth*. Dalam makalah ini akan dibahas bagaimana pembangunan dan aplikasi penggunaan struktur *FP-tree*. Struktur *FP-tree* ini kemudian digunakan bersamaan dengan *FP-growth* untuk menentukan data yang paling sering muncul dari suatu data (*frequent itemset*), yang kemudian berguna dalam penentuan rekomendasi promosi suatu produk.

Kata Kunci—Analisis Asosiasi, *FP-Tree*, *FP-Growth*, Rekomendasi Promosi Produk, Situs Belanja Online .

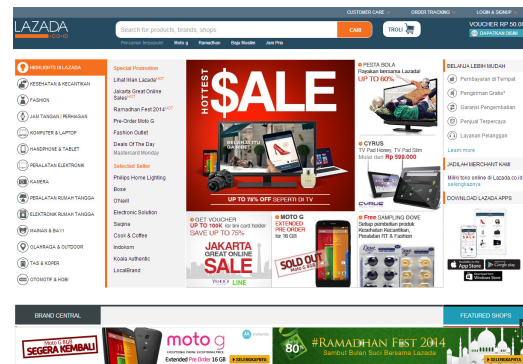
I. PENDAHULUAN

Seiring berkembangnya zaman, kecanggihan teknologi informasi juga berkembang dengan pesat. Teknologi banyak memudahkan kehidupan manusia dan membuat banyak hal menjadi lebih mudah dan lebih cepat diselesaikan. Misalnya, saat ini untuk memesan tiket kereta api, kita hanya perlu memesan melalui situs resmi kereta api, tidak perlu harus datang ke stasiun, dan masih banyak teknologi lain yang memudahkan kehidupan kita.

Salah satu teknologi yang saat ini paling diminati masyarakat adalah situs belanja online. Situs belanja online adalah suatu sarana jual beli yang memanfaatkan sistem elektronik seperti Internet, dimana pembeli dan penjual tidak perlu melakukan tatap muka secara langsung namun transaksi dapat tetap berjalan. Situs online ini dipilih oleh banyak orang karena alasan kemudahan dan efisiensi. Kini sudah ada bermacam-macam situs belanja online, seperti lazada.co.id, elevenia.co.id, mataharimall.com, tokopedia.com, mapemall.com, dan masih banyak lagi.

Barang-barang yang dijual di situs belanja online ini sangat beragam, mulai dari perlengkapan rumah tangga,

perlengkapan kantor, alat elektronik, pakaian, peralatan olahraga, hingga bahan-bahan kebutuhan sehari-hari seperti beras, minyak, susu, dan lain-lain.



Gambar 1.1 Salah satu situs belanja online, lazada.co.id
(Sumber: <http://startupbisnis.com/wp-content/uploads/2014/06/Screenshot-Lazada1.png>)

Situs belanja online ini juga memiliki strategi promosi dalam memasarkan situsnya dan juga membuat pengguna mengunjungi kembali situs tersebut dan melakukan transaksi lagi. Rekomendasi promosi produk merupakan salah satu strategi yang dilakukan situs-situs belanja online. Rekomendasi promosi produk merupakan suatu metode untuk mendapatkan pola tertentu tentang pola belanja pembeli akan suatu produk. Misalnya, pembeli 1 membeli beras dan minyak di situs belanja online X, kemudian pembeli 2 membeli minyak, gula, dan kecap di situs belanja online X, lalu data-data ini akan disimpan dan kemudian diolah dan dianalisis untuk dicari polanya dan akan didapatkan *frequent itemset*, yang berguna untuk rekomendasi promosi produk di kemudian hari. Sehingga misalnya situs belanja online X ingin mempromosikan suatu produk baru, yaitu beras organik, maka dengan mengetahui *frequent itemset* dari beras, promosi beras bisa 'dikaitkan' dengan promosi minyak, atau produk lain yang merupakan *frequent itemset* dari beras. Selain itu, dengan mengetahui *frequent itemset* juga, situs belanja online dapat memberikan rekomendasi produk, misalnya

seorang pembeli 1 membeli beras, maka pada bagian rekomendasi akan dimunculkan *frequent itemset* dari beras.

Oleh karena itu, makalah ini bertujuan untuk membahas penggunaan algoritma *FP-Growth* dan struktur *FP-Tree* dalam menentukan *frequent itemset* suatu produk, yang kemudian dapat digunakan untuk menentukan rekomendasi promosi suatu produk, khususnya pada situs belanja *online*. Makalah ini akan menggunakan pendekatan analisis asosiasi dalam pembahasannya.

II. TEORI MENGENAI ANALISIS ASOSIATIF, ALGORITMA *FP-GROWTH* DAN STRUKTUR *FP-TREE*

2.1 Analisis Asosiatif

Analisis Asosiasi merupakan suatu teknik penggalian data untuk data yang muncul bersamaan dengan frekuensi tertentu. Terdapat dua tahap dalam melakukan analisis asosiatif, yaitu *frequent itemset candidate generation* dan *rule generation*. Tahap *frequent itemset candidate generation*, merupakan tahap untuk mencari *frequent itemset* dengan menggunakan algoritma *FP-Growth* dan struktur *FP-Tree*. Tahap *rule generation* adalah tahap menggeneralisasi *rule* dengan *confidence* yang tinggi dari setiap *frequent itemset*. Berikut adalah hal-hal penting yang berhubungan dengan analisis asosiatif :

2.1.1 Support (Nilai Penunjang)

Support adalah persentase kombinasi item tersebut dalam basis data.

$$\text{Support}(A) = \frac{\text{Jumlah transaksi yang mengandung } A}{\text{Total transaksi}}$$

$$\text{Support}(A \cap B) = \frac{\text{Jumlah transaksi yang mengandung } A \text{ dan } B}{\text{Total transaksi}}$$

2.1.2 Confidence (Nilai Kepastian)

Confidence adalah persentase kuatnya hubungan antar item

$$\text{Confidence} = P(B|A) = \frac{\text{Support}(A \cap B)}{\text{Support}(A)}$$

2.1.3 Lift Ratio

Lift ratio mengukur seberapa pentingnya pola yang telah terbentuk, berdasarkan nilai *confidence* dan *support*.

$$\text{Lift} = \frac{\text{Support}(A \cap B)}{\text{Support}(A)} \times \text{Support}(B)$$

2.2 Algoritma *FP-Growth*

Algoritma *FP-growth* adalah algoritma yang digunakan untuk menentukan himpunan data yang paling sering muncul dalam kumpulan data-data (*frequent itemset*). Algoritma *FP-growth* menggunakan konsep pembangunan pohon prefix, yaitu *FP-tree*, dalam pencarian himpunan data yang paling sering muncul dengan menggunakan prinsip *divide and conquer*.

Algoritma *FP-Growth* dibagi menjadi tiga langkah

utama, yaitu :

2.2.1 Tahap Pembangkitan *Conditional Pattern Base*

Conditional Pattern Base merupakan subdatabase yang berisi *prefix path* (lintasan prefix) dan *suffix pattern* (pola akhiran). Pembangkitan *conditional pattern base* didapatkan melalui *FP-tree* yang telah dibangun sebelumnya.

2.2.2 Tahap Pembangkitan *Conditional FP-Tree*

Pada tahap ini *support count* dari setiap item pada setiap *conditional pattern base* dijumlahkan, kemudian setiap item yang memiliki *support count* lebih besar daripada *minimum support count* akan dibangkitkan dengan *conditional FP-Tree*.

2.2.3 Tahap Pencarian *frequent itemset*

Pada tahap ini, apabila *conditional FP-Tree* merupakan lintasan tunggal, maka didapatkan *frequent itemset* dengan melakukan kombinasi *item* untuk setiap *conditional*. Jika bukan lintasan tunggal, maka dilakukan pembangkitan *FP-Growth* secara rekursif.

Procedure $FP_{growth}^*(T)$

Input: A conditional *FP-tree* T

Output: The complete set of all FI's corresponding to T .

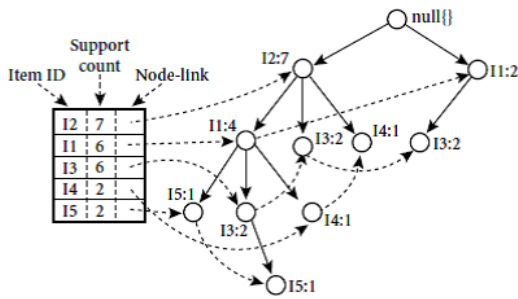
Method:

1. if T only contains a single branch B
2. for each subset Y of the set of items in B
3. output itemset $Y \cup T.base$ with count = smallest count of nodes in Y ;
4. else for each i in $T.header$ do begin
5. output $Y = T.base \cup \{i\}$ with $i.count$;
6. if $T.FP-array$ is defined
7. construct a new header table for Y 's *FP-tree* from $T.FP-array$
8. else construct a new header table from T ;
9. construct Y 's conditional *FP-tree* T_Y and possibly its *FP-array* A_Y ;
10. if $T_Y \neq \emptyset$
11. call $FP_{growth}^*(T_Y)$;
12. end

Gambar 2.1 Algoritma *FP-Growth* (Sumber : <http://www.computer.org/cms/Computer.org/dl/trans/tk/2005/10/figures/k13473.gif>)

2.3 Struktur *FP-Tree*

Struktur *FP-tree* adalah struktur penyimpanan data dengan memetakan setiap data ke dalam lintasan tertentu. *FP-tree* dibentuk oleh sebuah akar yang diberi nama null, cabang-cabang yang terdiri dari item-item tertentu, dan tabel *frequent header*. Setiap simpul pada *FP-tree* memiliki 3 informasi penting, yaitu label item, yang menginformasikan jenis item pada simpul tersebut, *support count*, yang menginformasikan jumlah lintasan yang melalui simpul tersebut, dan pointer penghubung yang menghubungkan simpul yang satu dengan simpul lain yang memiliki label item yang sama.



Gambar 2.2 : Contoh FP-tree untuk permasalahan penggalan frekuensi pola pada basis data (Sumber : <http://shareengineer.blogspot.co.id/2012/09/association-rule-mining.html>)

Suatu pola dikatakan sering muncul (*frequent pattern*), bila *support count* dari pola tersebut tidak kurang dari ambang *minimum support* yang telah didefinisikan

III. METODOLOGI

3.1 Pembangunan FP-Tree

Misalnya, diberikan sebuah tabel data transaksi, dengan *minimum support count* = 2

TID	Transaksi
100	{f, a, c, d, g, i, m, p}
200	{a, b, c, f, l, m, o}
300	{b, f, h, j, o}
400	{b, c, k, s, p}
500	{a, f, c, e, l, p, m, n}

Tabel 3.1.1 Tabel data transaksi mentah

Selanjutnya, kita perlu menyusun kemunculan tiap item berdasarkan frekuensinya

Item	Frekuensi
f	4
c	4
a	3
b	3
m	3
p	3

Tabel 3.1.2 Tabel Frekuensi Item

Selanjutnya, karena diketahui bahwa *minimum support count* adalah 2, dan pada pemindaian pertama didapatkan item yang memiliki frekuensi diatas *minimum support count* 2, yaitu f, c, a, b, m, dan p. Kelima item ini yang berpengaruh dan akan dimasukkan ke *FP-Tree*, sedangkan item sisanya akan dibuang, karena tidak berpengaruh (signifikan).

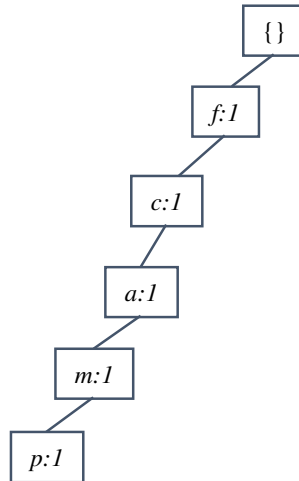
Kemudian, kita perlu mendata kemunculan item yang frequent dalam setiap transaksi, yang diurutkan berdasarkan frekuensi tertinggi.

TID	Transaksi
100	{f, c, a, m, p}

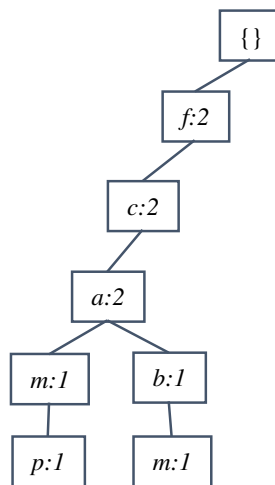
200	{f, c, a, b, m}
300	{f, b}
400	{c, b, p}
500	{f, c, a, m, p}

Tabel 3.1.3 Tabel data transaksi

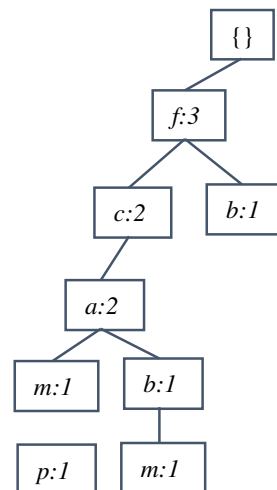
Selanjutnya, kita membentuk *FP-Tree* mulai dari pembacaan TID 100 sampai TID 500



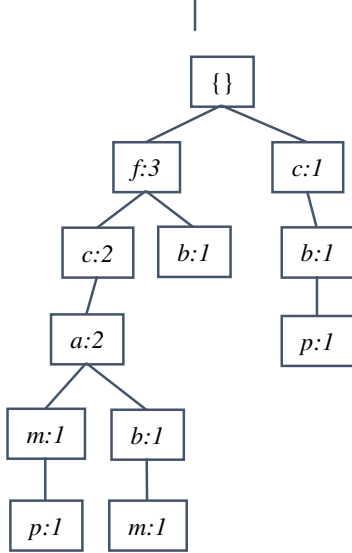
Gambar 3.1.1 Hasil pembentukan FP-Tree setelah pembacaan TID 100



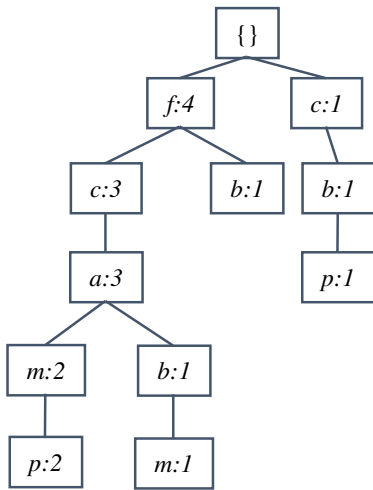
Gambar 3.1.2 Hasil pembentukan FP-Tree setelah pembacaan TID 200



Gambar 3.1.3 Hasil pembentukan FP-Tree setelah pembacaan TID 300



Gambar 3.1.4 Hasil pembentukan FP-Tree setelah pembacaan TID 400



Gambar 3.1.5 Hasil pembentukan FP-Tree setelah pembacaan TID 500

Cara pembangunan FP-Tree dengan merepresentasikan data transaksi yang terdapat pada tabel transaksi 3.1.1 adalah :

3.1.1 Kumpulan data dipindai pertama kali untuk menentukan *support count* dari setiap *item*. *Item* yang tidak *frequent* dibuang, sedangkan *frequent item* dimasukkan dan disusun dengan urutan menurun

3.1.2 Pemindaian kedua, yaitu pembacaan TID 100

{f,c,a,m,p} akan membuat simpul f,c,a,m, dan p, sehingga terbentuk lintasan transaksi $\text{Null} \rightarrow f \rightarrow c \rightarrow a \rightarrow m \rightarrow p$. *Support count* dari setiap simpul bernilai awal 1

3.1.3 TID 200 memiliki *prefix* transaksi yang sama dengan transaksi pertama, yaitu f, maka lintasan transaksi ketiga dapat ditimpakan di f, sambil menambah *support count* dari f, dan selanjutnya membuat lintasan baru sesuai dengan transaksi TID 200. Hal serupa juga sama untuk TID 300.

3.1.4 Setelah pembacaan TID 400 {c,b,p}, terbentuk lintasan kedua yaitu $\text{Null} \rightarrow c \rightarrow b \rightarrow p$. *Support count* masing- masing juga bernilai awal 1. Walaupun c ada pada transaksi pertama, namun karena *prefix* transaksinya tidak sama, maka transaksi kedua ini tidak bisa dimampatkan dalam satu lintasan.

3.1.5 Proses ini dilanjutkan sampai FP-tree berhasil dibangun berdasarkan tabel data transaksi yang diberikan.

3.2 Penerapan Algoritma FP-Growth

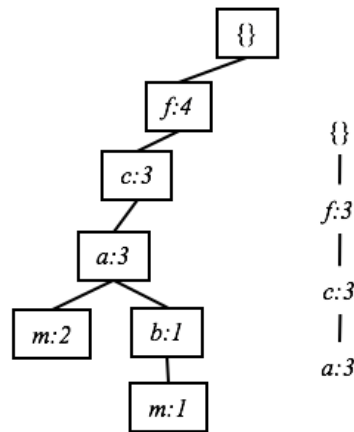
3.2.1 Tahap Pembangkitan Conditional Pattern Base

Item	Conditional Pattern Base
p	fcam:2, cb:1
m	fca:2, fcab:1
b	fca:1, f:1, c:1
a	fc:3
c	f:3
f	{}

Tabel 3.2.1 Conditional Pattern Base

3.2.2 Tahap Pembangkitan Conditional FP-Tree

Misalnya, kita ingin mengetahui frequent itemset dari m.

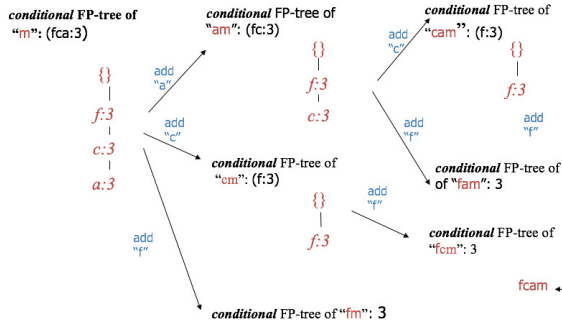


Gambar 3.2.2.1 Conditional FP-Tree item m

IV. STUDI KASUS

Misalnya, sebuah situs belanja online memiliki data yang terdiri dari ID pelanggan, dan item-item barang yang dibeli oleh pelanggan tersebut

3.2.3 Tahap Pencarian frequent itemset



Gambar 3.2.3.1 Tahapan pencarian frequent itemset m

Cara penerapan algoritma FP-Growth berdasarkan contoh persoalan diatas adalah :

1. Langkah pertama adalah membangun sebuah upapohon *FP-tree* dengan hanya menyertakan lintasan yang berakhir di m.
2. Nilai *support count item m* dihitung dan dibandingkan dengan *minimum support*, karena memenuhi, maka {m} termasuk ke *frequent itemset*.
3. Karena *item m frequent*, maka perlu dipecahkan subproblem untuk menemukan *frequent itemset* yang berakhir dengan am, cm, dan fm. Sebelum memecahkan subproblem ini, maka upapohon *FP-Tree* harus diubah terlebih dahulu menjadi *conditional FP-Tree*. *Conditional FP-Tree* mirip seperti *FP-Tree* biasa, namun *conditional FP-Tree* dimaksudkan untuk mencari frequent itemset yang berakhir item tertentu.
4. Hal-hal yang perlu diperhatikan saat membuat *Conditional FP-Tree* :
 - a. Setiap lintasan yang tidak mengandung m, dibuang.
 - b. Kemudian buang semua simpul m. Selanjutnya, cari lintasan *frequent itemset* yang berakhir di am, cm, dan fm.
 - c. Misalnya, {f,c,a} adalah *conditional pattern base m*, maka {f,c,a,m} adalah *frequent itemset* jika dan hanya jika {f,c,a} *frequent*.
5. Dengan menggunakan *Conditional FP-Tree*, buat pohon lintasan prefix untuk menemukan *frequent itemset* yang berakhir am, cm, dan fm. Lihat gambar 3.2.3.1 agar lebih jelas.
6. Didapatkan *frequent itemset* untuk item m adalah {m}, {f,m}, {c,m}, {a,m}, {f,c,m}, {f,a,m}, {c,a,m}, {f,c,a,m}.

TID	Transaksi
100	{gula, garam, minyak, piring, gelas, lampu, beras, pewangi ruangan}
200	{garam, gula, minyak, gula, sensok, beras, keripik tempe}
300	{gula jawa, gula, handuk, pelembut pakaian, keripik tempe}
400	{gula jawa, minyak, sabun pel, sabun cuci piring, pewangi ruangan}
500	{garam, gula, minyak, tahu susu, sendok, pewangi ruangan, beras, cangkir}

Tabel 4.1 Tabel data transaksi mentah

Misalnya situs belanja *online* tersebut ingin mempromosikan beras, maka perlu diketahui *frequent itemset* dari beras. Jika kita mengikuti langkah-langkah pada bagian 3, maka akan diperoleh bahwa *frequent itemset* untuk beras adalah {beras}, {gula,beras}, {garam,beras}, {minyak, beras}, {gula, minyak, beras}, {gula, garam, beras}, {garam, minyak, beras} dan { gula, garam, minyak, beras}.

Sehingga, dapat disimpulkan bahwa jika situs belanja online itu ingin mempromosikan beras, situs belanja online tersebut dapat ‘mengaitkan’ beras dengan garam, minyak, atau gula, karena garam, minyak dan gula merupakan *frequent itemset* untuk beras.

Kesimpulan ini akan sangat berguna bagi situs belanja *online*, karena strategi promosi yang dilakukan terus bergerak sesuai dengan pola belanja pembeli. Strategi promosi yang dilakukan akan berorientasi pada keinginan pasar, sehingga dengan mengikuti keinginan pasar, maka situs belanja *online* tersebut akan terus dikunjungi oleh konsumen, dan keuntungan situs belanja *online* itu juga akan bertambah.

V. KESIMPULAN

Tree merupakan salah satu struktur yang sangat berguna dalam menentukan rekomendasi promosi suatu produk pada situs belanja *online*. Melalui pengembangan struktur tree, yaitu *FP-Tree* dan menggunakan algoritma *FP-Growth* dengan pendekatan analisis asosiasi, situs belanja online dapat menyusun strategi usahanya yang bertujuan untuk memudahkan pelanggan dan menaikkan pendapatan. Karena data yang berasal dari pengguna itu sendiri, maka rekomendasi promosi yang didapatkan akan

terus mengikuti keinginan pembeli dan pasar.

VI. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Tuhan Yang Maha Esa, karena tanpa berkat dan pertolongan-Nya, penulis tidak akan dapat menyelesaikan makalah ini. Penulis juga mengucapkan terima kasih kepada Bapak Rinaldi Munir dan Ibu Harlili selaku dosen mata kuliah IF2120 Matematika Diskrit atas ilmu-ilmu yang telah diberikan dan bimbingannya selama satu semester. Tidak lupa penulis juga berterima kasih kepada keluarga dan teman-teman penulis yang telah memberikan dukungan moral kepada penulis.

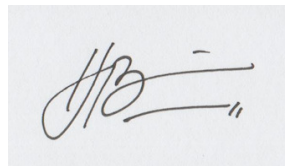
REFERENSI

- [1] Borgelt, Christian. "An Implementation of FP-Growth Algorithm" osdm.ua.ac.be/papers/p1-borgelt.pdf
Waktu akses : 3 Desember 2016, Pukul 22.00
- [2] Han, J., Micheline, K., "Data Mining: Concept and Techniques", *Simon Fraser University, Morgan Kaufman Publisher, 2000.*
Waktu akses : 3 Desember 2016, Pukul 23.00
- [3] Han, J., Yin, Y., "Mining Frequent Pattern Without Candidate Generation", <http://www-faculty.cs.uiuc.edu/~hanj/pdf/sigmod00.pdf>
Waktu akses : 8 Desember 2016, Pukul 19.00
- [4] <http://www.metode-algoritma.com/2013/02/algoritma-fp-growth.html>
Waktu akses : 8 Desember 2016, Pukul 19.00

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, 8 Desember 2016



Irene Edria Devina - 13515038