

**ANALISA LOG AKSES DI E-LEARNING MENGGUNAKAN ASSOCIATION  
RULE MINING UNTUK MENGENALI POLA BELAJAR SISWA  
(Studi Kasus di SMUN 1 Pare Kediri)**

Mohammad Khayat Subkhan<sup>1</sup>

<sup>1</sup>Staf Pengajar di Fakultas Teknik, Universitas Nusantara PGRI Kediri  
[aanmks@rocketmail.com](mailto:aanmks@rocketmail.com)

Abstraksi, *Data log aktifitas siswa di aplikasi e-learning menjadi sangat penting sekali manakala seorang guru ingin mempelajari pola perilaku siswanya ketika berinteraksi dengan aplikasi tersebut. Salah satu metode yang dapat digunakan untuk melakukan ekstraksi informasi mengenai pola perilaku menggunakan data log seperti itu yaitu menggunakan metode apriori. Apriori merupakan salah satu algoritma dalam model association rules mining yang saat ini sangat populer sekali dalam dunia machine learning. Association rule mining digunakan untuk menemukan asosiasi antar aktifitas di e-learning. Asosiasi ini berupa rules yang menyatakan hubungan sejumlah aktifitas. Kekuatan hubungan ini diukur menggunakan ukuran minimum support dan confidence. Dari penelitian ini dengan minimum support 0.2 dan confidence 1 dihasilkan sebuah aktifitas dengan asosiasi yang sangat kuat dengan banyak aktifitas lain yaitu course view. Dengan keakuratan mencapai 90%. Hal ini berarti aktifitas tersebut menjadi sangatlah penting untuk menjadi perhatian para guru untuk menyiapkan bahan ajar sebaik – baiknya dalam aplikasi e-learning.*

*Keyword : e-learning, machine learning, apriori, association rule, log aktifitas*

## I. Pendahuluan

### A. LatarBelakang

Tuntutan akan persaingan global telah menjadikan perkembangan zaman, ruang dan waktu menjadi kendala utama dalam upaya peningkatan mutu pembelajaran. Pertumbuhan jumlah peserta didik yang semakin banyak, terkadang tidak diimbangi oleh jumlah ruang dan waktu yang cukup dalam proses pembelajarannya. Sehingga hal ini berpotensi mengurangi kualitas interaksi antara pendidik dengan peserta didiknya. Yang akhirnya berujung pada kurang optimalnya mutu proses pembelajaran.

Dengan kehadiran aplikasi open source pembelajaran e-learning, seperti moodle yang telah digunakan oleh SMUN 1 Pare Kediri, maka menjadi mungkin sekali sebuah proses pembelajaran itu diselenggarakan secara mandiri dan kreatif. Segala bentuk aktifitas manual dikelas, bisa digantikan oleh fasilitas pada aplikasi ini, yang bisa diakses secara online, seperti modul bacaan, modul tugas, modul chat, modul forum, modul pilihan, modul quiz, dan sebagainya.

Moodle mencatat setiap aktifitas belajar siswa didalam databasenyanya, sehingga dengan hal ini kita bisa mempelajari bagaimana pola perilaku siswa ketika

berinteraksi dengan aplikasi dalam rangkain proses belajarnya, apa saja yang mereka kerjakan, kapan melakukannya, berapa kali dan sebagainya.

Untuk itu dalam penelitian ini, machine learning akan difokuskan untuk menemukan pola – pola perilaku belajar siswa dengan melihat akses log moodle menggunakan pendekatan model association rules mining. Metode ini akan menghasilkan pengetahuan berbasis rule yang dapat digunakan sebagai pedoman untuk menganalisa lebih jauh tentang pola perilaku belajar siswa ketika mereka berinteraksi dengan aplikasi e-learning khususnya di SMUN 1 Pare – Kediri.

## B. Association Rules

Association rule mining digunakan untuk menemukan asosiasi atau hubungan yang menarik dalam sekumpulan data yang besar (Large Itemset). Dimana data – data yang secara frequent berasosiasi bisa digunakan untuk membantu dalam pengambilan keputusan.

Representasi matematik pada model association rule bisa dijelaskan sebagai berikut :

Jika  $J = \{i_1, i_2, \dots, i_m\}$  disebut sebagai sekumpulan item Sedangkan  $D =$  set database transaksi dimana tiap transaksi  $T$  adalah sekumpulan item, maka  $T \subseteq J$ . Jika  $A, B =$  set item dari  $T$ , maka  $A$  dikatakan item transaksinya  $T$  jika dan hanya jika  $A \subseteq T$ .

Dalam association rule sebuah bentuk implikasi sederhana di tulis dalam bentuk

$A \Rightarrow B$ , dimana  $A \subset J$ ,  $B \subset J$  dan  $A \cap B = \emptyset$ .

Rule  $A \Rightarrow B$  didalam set databasetransaksi  $D$  memiliki support  $S$ , dimana  $S$  adalah prosentasetransaksi di dalam  $D$  yang mengandung  $A \cup B$ .

Maka support  $(A \Rightarrow B) = P(A \cup B)$

Rule  $A \Rightarrow B$  memiliki confidence  $C$  dalam set transaksi  $D$  jika  $C$  adalah prosentasedaritransaksididalam  $D$  yang mengandung  $A$  dan mengandung  $B$

Maka confidence  $(A \Rightarrow B) = P(B|A) = [\text{support count}(A \cup B) / \text{support count}(A)]$ .

Salah satu algoritma yang diusulkan dalam pendekatan asosiasi ini adalah algoritma apriori.

Berikut pseudocode algoritma apriori

```

1)  $L_1 = \{\text{large 1-itemsets}\};$ 
2) for ( $k = 2; L_{k-1} \neq \emptyset; k++$ ) do begin
3)    $C_k = \text{apriori\_gen}(L_{k-1});$  // New candidates
4)   forall transactions  $t \in D$  do begin
5)      $C_t = \text{subset}(C_k, t);$  // Candidates contained in
6)     forall candidates  $c \in C_t$  do
7)        $c.\text{count}++;$ 
8)   end
9)    $L_k = \{c \in C_k \mid c.\text{count} \geq \text{minsup}\}$ 
10) end
11) Answer =  $\bigcup_k L_k$ 

```

The apriori-gen function takes as argument  $L_{k-1}$ , the set of all large  $(k-1)$ -itemsets. It returns a superset of the set of all large  $k$ -itemsets. The function works as follows. First, in the join step, we join  $L_{k-1}$  with  $L_{k-1}$ .

```

insert into  $C_k$ 
select  $q.\text{item}_1, p.\text{item}_2, \dots, p.\text{item}_{k-1}, q.\text{item}_{k-1}$ 
from  $L_{k-1} p, L_{k-1} q$ 
where  $p.\text{item}_1 = q.\text{item}_1, \dots, p.\text{item}_{k-2} = q.\text{item}_{k-2},$ 
 $p.\text{item}_{k-1} < q.\text{item}_{k-1}$ 

```

Next, in the prune step, we delete all itemsets  $c \in C_k$  such that some  $(k-1)$ -subset of  $c$  is not in  $L_{k-1}$ .

2 operasi utama pada algoritma Apriori adalah :

### 1) Operasi join

Untuk mendapatkan  $L_k$  yaitu sebuah set kandidat  $k$ -itemset, dilakukan dengan cara menggabungkan atau join

$L_{k-1}$  dengan dirinya sendiri. set kandidat ini di simbolkan dengan  $C_k$ . misal  $I_1$  dan  $I_2$  adalah itemset  $L_{k-1}$ , maka  $I_1$  dan  $I_2$  bisa dijoin jika item k-2 nya seperti rumusan berikut :  
 $(L_1[1]=L_2[1]).( L_1[2]=L_2[2]) \dots\dots(L_1[k-2]=L_2[k-2]). (L_1[k-1]<L_2[k-1])$

## 2) Operasi prune

$C_k$  adalah superset dari  $L_k$ . Proses scan databaseakan menghitung tiap kandidat di dalam  $C_k$ , dan hasilnya digunakan untuk menentukan  $L_k$  yaitu itemset yang memiliki support tidak kurang dari minimum support yang telah ditentukan di  $C_k$ . Proses scan dan perhitungan ini dapat digantikan dengan operasi prune pada algoritma apriori. (k-1)-itemset yang tidak frequent tidak bisa menjadi subset dari sebuah frequent k-itemset, begitupula jika ada (k-1)-subset dari kandidat k-itemset tidak menjadi anggota dalam  $L_{k-1}$ , maka kandidat tersebut dianggap tidak frequent dan bisa di hapus dari  $C_k$ .

## C. Moodle

Moodle adalah sebuah aplikasi berbasis web yang berfungsi sebagai learning management system (LMS). Melalui aplikasi ini aktifitas belajar mengajar konvensional bisa dilakukan secara digital misalkan untuk mengerjakan latihan – latihan, diskusi, pre dan post quiz, mereview materi pelajaran, mendownload materi pelajaran sampai mengerjakan final exam. Setiap aktifitas siswa pada aplikasi ini akan dicatat dalam log-nya. Dan hal ini menjadi sebuah basis pengetahuan yang bernilai untuk dianalisa.

Data berasal dari aktifitas siswa yang terekam dalam log aplikasi e-learning, sebelum siap diolah dan di analisa, dilakukan proses cleaning untuk membuang record – record yang tidak berguna

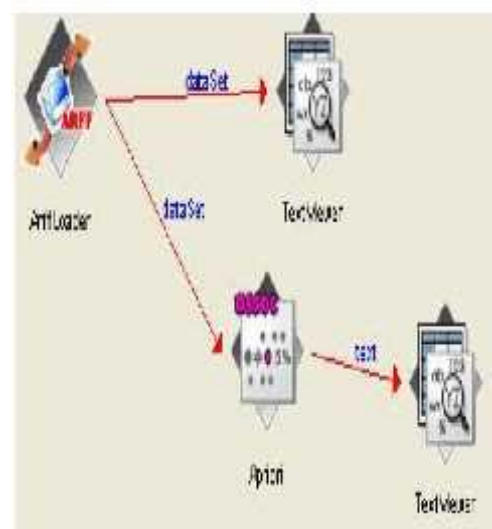
## D. WEKA

Teknik association rule diterapkan dengan menggunakan algoritma Apriori. Dimana analisa datanya menggunakan bantuan aplikasi WEKA (Waikato Environment for Knowledge Analysis), yaitu sebuah tool free license yang didalamnya banyak terdapat implementasi algoritma machine learning yang digunakan untuk menganalisa dataset dari permasalahan data mining.

Duaparameter threshold digunakanpadaalgoritmaaprioriadalah minimum support dan confidence. Padajurnalini minimum support yang digunakanadalah 0.2 dengannilai confidence sebesar 0.9.

Format file untuk proses ini yang digunakanoleh WEKA adalaharff format, data flow di WEKA bisadigambarkansebagai berikut :

Gambar data flow WEKA



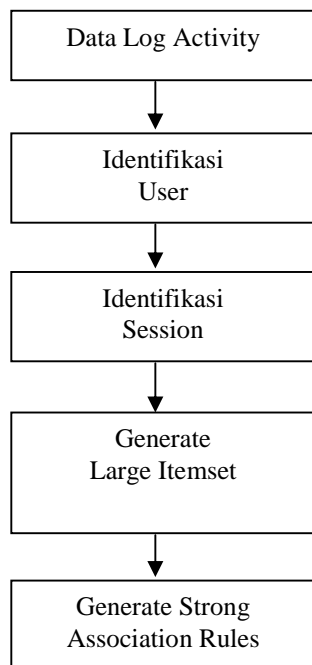
Dari penelitianinihasilkanbanyak rule, tapitidaksemuapengetahuan yang

adapada rule tersebut berguna atau menarik bagi guru untuk mengembangkan strategi pembelajaran.

## II. TEKNIK PENELITIAN

Dalam penelitian ini, tentunya ada langkah – langkah teknis yang harus dilakukan. Langkah – langkah tersebut digambarkan melalui blok diagram dibawah ini :

Diagram Teknik penelitian



## III. Dataset

Dataset yang digunakan dalam penelitian ini bersumber dari log activity database moodle SMUN 1 Pare Kediri, berikut ini adalah beberapa hal yang berhubungan dengan dataset. Dibawah ini struktur datasetnya :

No	Atribut	Tipe Data
1	Id	Bigint
2	Time_Access	Bigint
3	Userid	Bigint
4	IP_Addres	Varchar

5	Module	Varchar
6	Action	Varchar

Dataset diambil dari log activity 50 orang siswa kelas X SMUN1 Pare Kediri pada tahun ajaran 2011/2012 dengan rata – rata aktifitas 1000 aktifitas, yang terkelompokkan dalam rata – rata 200 session. Jadi total sekitar 50.000 activity log.

Berikut ini adalah contoh dataset log activity siswa dengan userid : 11

Id	Time	IP	Module	Action
1	2010-07-17 09:07:49	192.168.0.109	User	Login
2	2010-07-17 09:07:55	192.168.0.109	Course	View
3	2010-07-17 09:07:56	192.168.0.109	User	Logout
4	2010-07-21 07:11:18	192.168.0.239	User	Login
5	2010-07-21 07:11:33	192.168.0.239	Course	View
6	2010-07-21 07:11:42	192.168.0.239	Resource	View
7	2010-07-21 07:20:56	192.168.0.239	Quiz	View
8	2010-07-21 07:25:16	192.168.0.239	User	Logout
9	2010-07-22 10:01:16	192.168.0.111	User	Login
10	2010-07-22 10:50:25	192.168.0.111	Course	View
11	2010-07-22 10:55:22	192.168.0.111	Resource	View
12	2010-07-23 11:07:16	192.168.0.105	User	Login
13	2010-07-23 07:08:17	192.168.0.105	Course	View
14	2010-07-23 07:07:16	192.168.0.105	Resource	View
15	2010-07-23 11:08:16	192.168.0.107	User	Login
16	2010-07-23 11:10:10	192.168.0.107	Course	View
17	2010-07-23 11:30:55	192.168.0.107	Resource	View
18	2010-07-23 11:45:18	192.168.0.107	Quiz	View
19	2010-07-23 12:05:10	192.168.0.107	Message	Write
20	2010-07-23 13:25:11	192.168.0.107	User	Logout

#### IV. Pre-processing

Proses preprocessing dilakukan untuk mengidentifikasi user dan mendapatkan session aktif user. User diidentifikasi melalui field user id sedangkan session ditandai oleh string login dan logout. Dimana session diartikan sebagai rangkaian aktifitas yang dilakukan oleh user dalam satu kesempatan yang diawali oleh aktifitas login dan diakhiri oleh aktifitas logout. Sebuah session diasumsikan sebagai sebuah transaksi dalam pemodelan association rule mining. Artinya setiap siswa nantinya akan mempunyai sekumpulan transaksi.

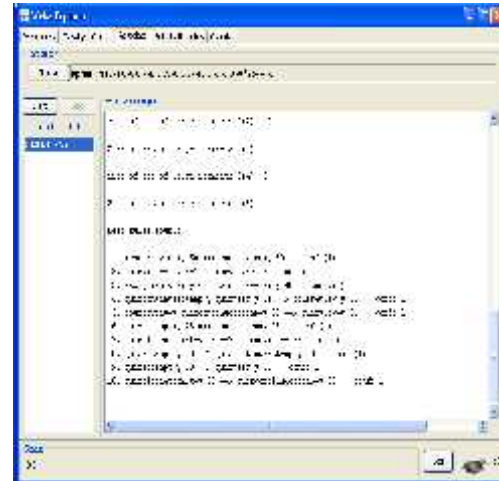
SID	Aktifitas
1	{{Course_view}}
2	{{Course_view},{Resource_view},{Quiz_view}}
3	{{Course_view},{Resource_view}}
4	{{Course_view},{Resource_view}}
5	{{Course_view},{Resource_view},{Quiz_view},{Message_write}}

Dari sekumpulan transaksi untuk setiap siswa ini selanjutnya akan diterapkan algoritma apriori untuk mencari large item set.

#### Generate Strong Association Rules

Tahap ini merupakan tahap inti dari penerapan model association rule mining. Proses dilakukan dengan bantuan WEKA untuk menemukan rule – rule terbaik. Yang selanjutnya bisa dianalisa untuk mengenal pola perilaku tiap siswa di e-learning. Generate strong association rule dilakukan dengan minimum support 0.2 dan confidence 0.9.

Tampak proses di aplikasi WEKA sebagai berikut :



#### V. HASIL PENELITIAN

Dari rangkaian penelitian ini, dihasilkan banyak rules tapi tidak semuanya bisa digunakan untuk mengidentifikasi pola perilaku aku. Dengan minimum support 0.2 dan confidence 0.9 dihasilkan kelompok rule – rule terbaik yang bisa digunakan sebagai berikut :

No	Item X	Item Y	Confidence
1	ResourceView	CourseView	1
2	QuizView	CourseView	1
3	AssignmentView	CourseView	1
4	QuizAttempt	CourseView	1
5	QuizAttempt	QuizView	1

Rules diatas merupakan rata – rata mayoritas dari strong rules yang dihasilkan dari data log aktifitas setiap siswa. Dan disitu terlihat bahwa aktifitas courseview menjadi aktifitas yang paling banyak dan dominan dipilih oleh siswa. Course View

merupakan aktifitas untuk melihat materi se  
buah pelajaran yang  
dipersiapkan oleh seorang guru di aplikasi  
e-learning.

Darisini bisa  
diketahui betapa pentingnya seorang guru  
menyiapkan materi – materi berkualitas,  
sehingga akan mendorong siswa lebih aktif  
dalam e-learning.

Rules”, Proceeding of the 20<sup>th</sup> VLDB  
Conference pp 487 –499

#### DAFTAR PUSTAKA

Kularbphetong, Kunyanuth; Waraporn,  
Phanu; Tongsiri, Cholticha, 2012.  
“*Analysis of student motivation  
Behaviour on e-learning based on  
Association rule mining*”, World  
Academy Of Science, Engineering and  
Technology 66.

Dimitrijevic, Maja; Bosnjak, Zita, 2010.  
“*Discovering Interesting Association  
Rules in The Web Log Usage Data*“,  
Interdisciplinary Journal of Information,  
Knowledge and Management, Volume 5.

Agrawal, R; Mannila, H; Srikant, R;  
Toivonon, H; Verkamo, A, 1996. “Fast  
Discovery of Association Rules”,  
Advanced in Knowledge Discovery and  
Data Mining.

Gunawan, 2012, “*Association Rules  
Mining (Bagian 2)*” ,Bahan Ajar  
Kuliah Knowledge Discovery in Database  
dan Data Mining, Kuliah #8.

MohdHelmy Ab Wahab;  
AzizulAzharRamli; NeureizeArbaiy,  
ZurinahSuradi, 2005. “*Log Information  
Mining Using Association Rules  
Technique: A Case Study of Utusan  
Education Portal*”, Proceeding of the  
International Conference on Robotics,  
Vision, Information and Signal  
Processing, ROVISP.

Agrawal, R; Srikant, R, 1994, “Fast  
Algorithms For Mining Association