

Laporan Penelitian

by Slamet K

Submission date: 24-Aug-2021 10:14PM (UTC-0500)

Submission ID: 1635578672

File name: 3.pdf (815.3K)

Word count: 1261

Character count: 7471

1
**SISTEM DETEKSI DINI UNTUK MENINGKATKAN PERFORMANCE
KELULUSAN MAHASISWA DENGAN ID3
(STUDI KASUS : TEKNIK INFORMATIKA UNITOMO)**

Slamet Kacung¹⁾, Budi Santoso²⁾

^{1,2)}Fakultas Teknik, Universitas Dr. Soetomo Surabaya
slamet@unitomo.ac.id¹⁾, Santosobudi.id@gmail.com²⁾

Abstract

The performance of academic programs higher education is measured by number of graduates are produced by each study program as reflected in the standard III accreditation form in point 3.1.1 and 3.1.4. The study program is required to have a good performance is marked by the increasing number of graduates in proportion to the number of students received so that ratio lecturers with students can be maintained. The more students who are accepted in college if they are not comparable with the number of graduates in each year will have an impact on the quality of learning. The result of this graduation becomes the evaluation material of the study program which will be the input of the study program and the Academic Advisors (DPAM) in order to provide treatment to the problem students so that they can improve the performance of the graduates. DPAM has a very important role to the progress of the learning process of students Guide, but with the amount of guidance that is increasingly causing students to be misdirected and in the end the student performance becomes bad, for that need an early detection system to improve the performance of graduates based on the results of the recommendation from the decision tree classifier. this method can generate a decision tree and give recommendations to students problems with accurate.

Keywords : Performance, DPAM, Decision tree classifier

1. PENDAHULUAN

Program Studi dituntut untuk memiliki performa yang baik ini ditandai dengan meningkatnya jumlah lulusan sebanding dengan jumlah mahasiswa yang diterima sehingga rasio dosen dengan mahasiswa dapat terjaga. Semakin banyak mahasiswa yang diterima diperguruan tinggi bila tidak sebanding dengan jumlah lulusan dalam tiap tahun maka akan berdampak pada kualitas lulusan setiap program studi, dan pada akhirnya akan menambah jumlah nisbah pada forlap dikti. Hasil kelulusan ini menjadi bahan evaluasi program studi yang nanti akan menjadi masukan program studi dan Dosen Pembimbing Akademik (DPAM) agar dapat memberikan *treatment* kepada mahasiswa yang bermasalah sehingga dapat meningkatkan *performance* lulusan. Untuk mengatasi permasalahan tersebut perlu dibangun aplikasi deteksi dini di lingkungan program studi guna mengurangi mahasiswa *Drop Out*.

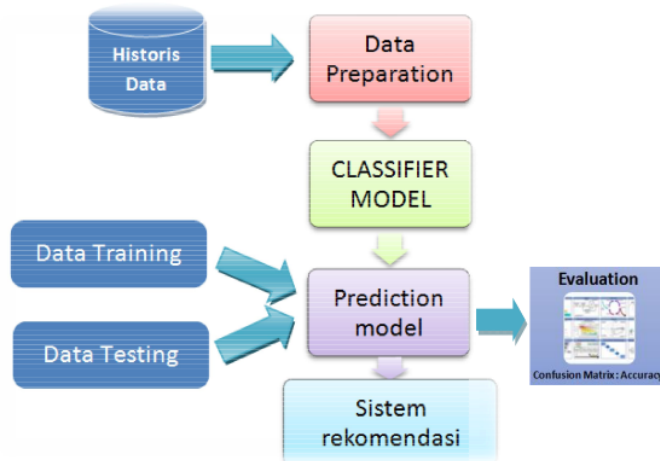
Sistem informasi adalah salah satu sumber daya yang dapat digunakan untuk meningkatkan performa lulusan program studi, Sistem informasi dapat digunakan untuk mendapatkan data, mengolah data menjadi pengetahuan yang awalnya hanya untuk memenuhi kebutuhan informasi operasional sehari-hari sehingga dapat dimanfaatkan untuk pengambilan keputusan yang strategis.

Induksi Decision Tree (ID3) adalah bagian dari metode klasifikasi dan merupakan jenis metode yang dapat memetakan atau memisahkan dua atau lebih class yang berbeda. Dengan metode ini maka sistem dapat menggenerate sebuah pohon keputusan yang dapat menghasilkan rekomendasi terhadap mahasiswa yang bermasalah dengan akurat.

Data Mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *knowledge discovery in database* (KDD) secara keseluruhan (Jiawei Han, Micheline Kamber, 2006).

2. METODE PENELITIAN

Dalam paper ini diajukan metode untuk membangun sistem rekomendasi berdasarkan *Data Preparation*, Membangun model klasifikasi menggunakan Induksi *Decision Tree*, kemudian hasil dari Id3 dilakukan proses training dan testing untuk menghasilkan sistem rekomendasi terhadap mahasiswa yang bermasalah, dan *evaluation* untuk mengukur kinerja berdasarkan *confusion matrix*. Gambar 1. Menunjukkan langkah-langkah secara garis besar metode tersebut.



Gambar 1. Blok Diagram Tahapan Penelitian

Data preparation

Datasheet yang digunakan dalam penelitian ini diambil dari data historis sistem informasi mahasiswa (SIMA) Universitas Dr. Soetomo prodi. Teknik Informatika. Beberapa Feature dari datasheet tersebut formatnya harus disesuaikan dengan kebutuhan metode Id3 dari numeric menjadi nominal seperti ips, ipk, dan totsks seperti pada tabel 1 s.d 3.

Tabel 1. Konfersi attribute

Nama attribute	Range Nilai	Nilai
IPS	3.50 – 4.00	SB
	2.75 – 3.49	B
	2.10 – 2.74	C
	0.00 – 2.00	K
IPK	3.50 – 4.00	CM
	2.75 – 3.49	SM
	2.10 – 2.74	M
	0.00 – 2.00	K
Totsks	0 – 17	<18
	18 – 24	>18

	25 – 43	>24
	44 – 64	>44
	65 – 74	>64
	75 – 99	>74
	100 – 148	>99

Pembentukan Classifier dan Prediksi Model

Untuk membentuk model klasifikasi dalam paper ini menggunakan algoritma induksi decision tree (Id3).

PROCEDURE Id3(Examples, TargetAttribute, Attributes)

Examples are the training examples. Target-attribute is the attribute whose value is to be predicted by the tree. Attributes is a list of other attributes that may be tested by the learned decision tree. Returns a decision tree that correctly classifies the given Examples.

- ♦ Create a Root node for the tree
- ♦ If all Examples are positive, Return the single-node tree Root, with label = +
- ♦ If all Examples are negative, Return the single-node tree Root, with label = -
- ♦ If Attributes is empty, Return the single-node tree Root, with label = most common value of Target_attribute in Examples
- ♦ Otherwise Begin
 - A <--- the attribute from Attributes that best* classifies Examples
 - The decision attribute for Root <--- A
 - For each possible value, v_i , of A,
 - Add a new tree branch below Root, corresponding to the test $A = v_i$
 - Let $Examples_{v_i}$ be the subset of Examples that have value v_i for A
 - If $Examples_{v_i}$ is empty
 - Then below this new branch add a leaf node with label = most common value of Target_attribute in Examples
 - Else below this new branch add the subtree
 - Call ID3 ($Examples_{v_i}$, Target_attribute, Attributes - {A})
- ♦ End
- ♦ Return Root

6 Untuk memilih attribute sebagai root, didasarkan pada nilai gain tertinggi dari attribute-attribute yang ada, seperti yang tertera dalam persamaan 1 berikut.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times Entropy(S_i) \quad (1)$$

Sedangkan untuk perhitungan nilai entropy dapat dilihat pada persamaan 2 berikut.

$$Entropy(S) = \sum_{i=1}^c -p_i \cdot \log_2 p_i \quad (2)$$

Evaluation

Evaluasi diperlukan untuk mengukur akurasi dan error dari sebuah model berdasarkan akurasi, kehandalan, dan kegunaan.

a. Akurasi

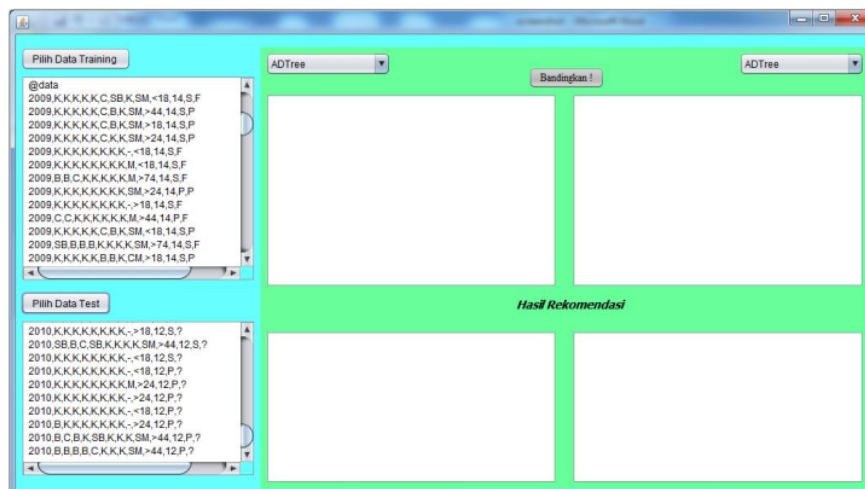
- Mengukur seberapa baik model mengkorelasikan antara hasil dengan feature dalam data yang telah disediakan
- Ada beberapa model akurasi, tetapi semua model akurasi tergantung pada data yang digunakan.

- b. Kehandalan
- Mengukur model data mining yang diterapkan pada dataset yang berbeda
 - Model data mining dapat diandalkan jika menghasilkan pola yang sama dari data testing yang disediakan
- c. Kegunaan
- Kegunaan ini meliputi berbagai metric yang mengukur apakah model tersebut memberikan informasi yang berguna.

Keseimbangan dari ketiganya sangat diperlukan karena belum tentu model yang akurat adalah handal, dan yang handal atau akurat belum tentu berguna. Pada paper ini *evaluation* yang digunakan berdasarkan hasil *confusion matrix*.

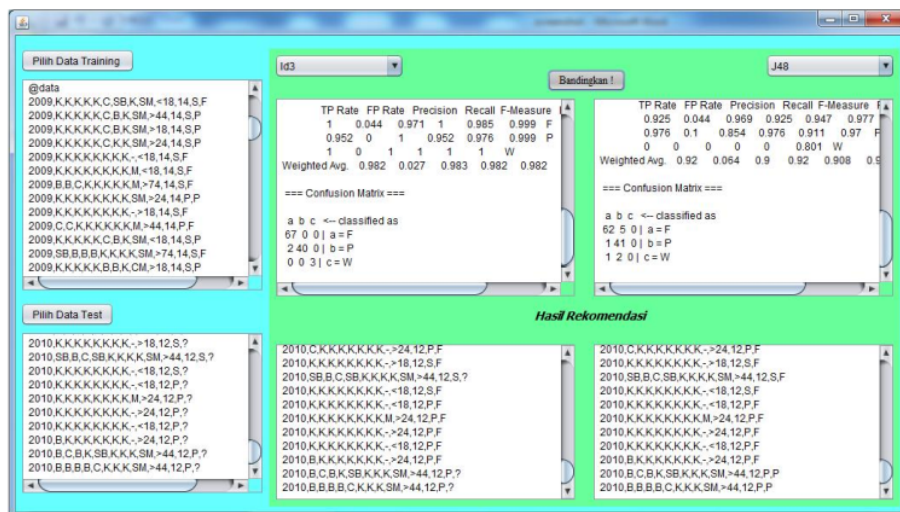
3. HASIL DAN PEMBAHASAN

Pada eksperimen ini proses yang dilakukan adalah upload data yang akan dilakukan proses mining, kemudian proses training. Proses training ini merupakan proses yang harus dilakukan oleh mesin (komputer) agar mesin ini memiliki kemampuan untuk memprediksi. Pada proses training sejumlah data akan dibagi menjadi dua 70% untuk training dan 30% untuk testing.



Gambar 2. Proses Data training dan testing

Pada proses pengujian ini menggunakan metode induksi decision tree, tetapi sebagai pembandingan sehingga paper ini menambahkan metode C4.5. Dari hasil rekomendasi pada gambar 3. Menunjukkan bahwa algoritma Id3 lebih baik dari sisi true positif rate (TP) yaitu sebesar 98%, sedangkan pada algoritma C4.5 sebesar 94% seperti yang terlihat pada hasil *confusion matrix* pada gambar 3 berikut.



Gambar 3. Hasil Rekomendasi dari sistem

4. KESIMPULAN

Dari hasil dan pembahasan yang dapat disimpulkan sebagai berikut.

1. Metode Id3 sistem mampu menebak benar dengan nilai rata-rata (TP-Rate) sebesar 98%, lebih baik dibandingkan dengan metode C4.5 sebesar 94%.
2. Hasil rekomendasi menunjukkan ada beberapa mahasiswa yang bermasalah untuk dilakukan treatmen oleh DPAM.
3. Sistem ini lebih baik lagi bila hasil dari rekomendasi mampu memberikan reminder atau notifikasi secara otomatis pada dosen DPAM.

5. REFERENSI

- [1] BAN-PT, "Matriks Penilaian Instrumen Akreditasi Program Studi Sarjana", Buku 6, 2008: <http://sipma.ui.ac.id>. Diakses tanggal 16 November 2015.
- [2] Jiawei Han, Micheline Kamber, "Data Mining Concepts and Techniques" 2th Edition, Elsevier. 2006
- [3] Tom. Mitchell, "Machine Learning", The McGraw-Hill, Inc. International Edition, 1997, halaman 56
- [4] Slamet Kacung, Edi Prihartono, "Sistem Cerdas Untuk Mendeteksi Dini Penyakit Jantung Dengan Decision Tree", INFORM Vol. 1 No.2, 2016



Laporan Penelitian

ORIGINALITY REPORT

22%

SIMILARITY INDEX

20%

INTERNET SOURCES

5%

PUBLICATIONS

6%

STUDENT PAPERS

PRIMARY SOURCES

1

docplayer.info

Internet Source

6%

2

slideplayer.info

Internet Source

6%

3

Submitted to Program Pascasarjana
Universitas Negeri Yogyakarta

Student Paper

2%

4

idoc.pub

Internet Source

2%

5

adoc.pub

Internet Source

1%

6

anzdoc.com

Internet Source

1%

7

pesquisa.bvsalud.org

Internet Source

1%

8

Submitted to City University of Hong Kong

Student Paper

1%

9

Submitted to Universitas Putera Batam

Student Paper

1%

10

pt.scribd.com

Internet Source

1 %

11

www.scribd.com

Internet Source

1 %

12

hal.archives-ouvertes.fr

Internet Source

1 %

13

tunasbangsa.ac.id

Internet Source

1 %

14

sandbox.ijcaonline.org

Internet Source

<1 %

Exclude quotes On

Exclude matches Off

Exclude bibliography On

Laporan Penelitian

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6