

EDUCATIONAL DATA MINING UNTUK MENINGKATKAN PENDUKUNG KEPUTUSAN DALAM SISTEM INFORMASI SEKOLAH

M. Mahaputra Hidayat^{1*}

¹Program Studi Teknik Informatika, Fakultas Teknik, Universitas Bhayangkara Surabaya
Jl. A. Yani 114, Surabaya Telp. 031 - 8285602

*Email: mahaputra@ubhara.ac.id

Abstrak

Educational Data Mining adalah disiplin yang muncul berkaitan dengan pengembangan metode untuk mendapatkan informasi unik dari dataset yang berasal dari pendidikan. Informasi unik tersebut kemudian dievaluasi agar dapat digunakan untuk mengatasi permasalahan-permasalahan dalam dunia pendidikan seperti mengukur kinerja siswa, memprediksi kegagalan sekolah dan sebagainya. Pada paper ini diusulkan sebuah model konseptual untuk meningkatkan integrasi data dan memasukkan data mining dalam statistik pendidikan. Dalam model konseptual ini, data pendidikan periodik yang dihasilkan dari sistem informasi sekolah akan digunakan sebagai temuan awal untuk proses evaluasi pengambilan keputusan. Sebagai prototipe dari sebuah implementasi dari konsep yang diajukan ini, prediksi keberhasilan sekolah menggunakan Jaringan Saraf Tiruan (JST) diamati dan diuji menggunakan struktur yang diusulkan. Hasil dari penelitian ini menunjukkan bahwa penggunaan variabel siswa, guru dan sekolah memiliki pengaruh yang signifikan terhadap keberhasilan sekolah dibanding hanya menggunakan variabel siswa saja. Sehingga temuan ini dapat digunakan oleh pengambil keputusan untuk merekomendasikan perbaikan sistem pendidikan berdasarkan fakta dari data.

Kata kunci: *Educational data mining, sistem informasi sekolah, sistem pendukung keputusan.*

1. PENDAHULUAN

Pendidikan merupakan proses berkelanjutan yang berlangsung beberapa tahun untuk setiap siswa. Selama waktu tersebut, informasi pendidikan yang terkait disimpan secara teratur. Pada awal semester, misalnya, setiap siswa mengambil daftar mata pelajaran yang sudah ditentukan sesuai kurikulum yang berlaku. Kemudian, pada akhir semester, tahap evaluasi diberikan untuk setiap mata pelajaran. Pada semester berikutnya, proses ini diulang dengan rincian yang berbeda. Untuk menyimpan kegiatan tersebut, diperlukan suatu sistem *database* yang mampu mendukung penanganan data pendidikan, yang dikenal sebagai sistem informasi sekolah.

Kebanyakan sistem informasi sekolah yang ada saat ini, hanya memberikan informasi data pendidikan yang bersifat individu seperti rincian masing-masing siswa (data pribadi), rincian guru, rincian fasilitas sekolah, rincian substansi pembelajaran dan rincian proses pembelajaran. Fungsi utama dari sistem informasi sekolah yang ada adalah: menyimpan data, pengolahan data, dan laporan. Sistem ini menyediakan berbagai jenis laporan transaksional pendidikan dan statistik periodik yang hanya berakhir sebagai laporan saja. Untuk itu diperlukan suatu perbaikan dengan memasukkan kecerdasan ke dalam sistem, dimana proses pendidikan dimodelkan, dan bersama-sama dengan pendekatan *data mining*, dirancang sebuah sistem cerdas *monitoring* pendidikan.

Paper ini mengusulkan sebuah model konseptual untuk meningkatkan integrasi data dan memasukkan *data mining* dalam sistem informasi sekolah. Dalam model ini, data statistik pendidikan periodik yang dihasilkan sistem akan digunakan sebagai *input* untuk memulai proses evaluasi dalam pengambilan keputusan. Hasil dari proses tersebut dapat digunakan oleh pengambil keputusan untuk merekomendasikan perbaikan sistem pendidikan berdasarkan fakta dari data.

Peningkatan kualitas pendidikan merupakan salah satu tanggung jawab pemerintah. Tingkat kegagalan sekolah adalah salah satu di antara banyak parameter yang dapat digunakan untuk mengevaluasi kualitas pendidikan nasional. Banyak peneliti telah mengidentifikasi berbagai faktor yang dapat mempengaruhi kinerja rendah dari sekolah. Faktor-faktor tersebut dapat berasal dari aspek siswa, aspek guru, dan lingkungan sekitar siswa. Sebuah laporan dari statistik pendidikan dapat digunakan sebagai kondisi awal untuk mengembangkan penyelidikan yang menggabungkan lebih banyak variabel dan metode sehingga dapat meningkatkan pendukung keputusan yang lebih baik.

Berbagai penelitian telah dilakukan untuk topik serupa tentang memprediksi kegagalan sekolah. Araque dkk., (2009) telah mengidentifikasi beberapa faktor yang mempengaruhi rendahnya kinerja siswa sehingga dapat menyebabkan kegagalan siswa pada tingkat pendidikan yang berbeda. Dua tahun berikutnya, Vera dkk., (2011) menerapkan teknik *Educational Data Mining* (EDM) untuk memprediksi penurunan dan kegagalan sekolah. Penelitian terbaru yang dilakukan Hidayat dkk., (2013) menerapkan teknik EDM untuk memprediksi mahasiswa *drop-out* berdasarkan perilaku sosialnya. Penelitian-penelitian ini telah menunjukkan hasil yang menjanjikan sehubungan dengan karakteristik sosiologis, ekonomi atau pendidikan yang mungkin lebih relevan dalam prediksi kinerja akademik yang rendah.

Pada paper ini digunakan contoh tentang bagaimana data periodik pendidikan dapat memicu temuan evaluasi penyelidikan tingkat keberhasilan sekolah. Model statistik yang dihasilkan menggunakan parameter dari aspek siswa. Meskipun model ini sudah cukup memberikan hasil yang spesifik, tapi diselidiki juga apakah aspek guru dan aspek lingkungan terlibat dalam keberhasilan sekolah. Jaringan Saraf Tiruan (JST) digunakan sebagai metode untuk membuktikan hipotesis tentang pentingnya hubungan multivariat antar variabel yang mempengaruhi kegagalan atau keberhasilan sekolah. JST dipilih karena memiliki kemampuan untuk memperoleh informasi berdasarkan hasil pelatihan *output* data, walaupun tidak memiliki pelatihan data riil sebagai masukan. Dari hasil tersebut, diharapkan beberapa tindakan dapat diambil untuk sistem peringatan dini sebagai upaya untuk meningkatkan pengambilan keputusan dalam sistem pendidikan.

2. METODOLOGI

Sebuah sekolah dapat diklasifikasikan sebagai sekolah yang baik, bila memenuhi beberapa kriteria, yaitu: (1) aspek siswa: kualitas lulusan harus diakui oleh lembaga lain, (2) aspek guru: guru merencanakan subyek yang memadai untuk belajar, jumlah guru yang memadai, (3) aktivitas organisasi: sekolah memiliki libur nasional dan program perayaan agama, (4) orang tua menerima laporan hasil belajar anak-anak mereka dengan baik. Orang tua memiliki pilihan sekolah untuk mengirimkan anak-anak mereka ke sekolah (Lipsitz, 1984). Sedangkan menurut Departemen Pendidikan Nasional, sekolah dapat dikategorikan sebagai sekolah yang baik, bila memenuhi delapan kriteria berikut: (1) kualitas tinggi dalam pendaftaran siswa. Siswa akan dievaluasi berdasarkan prestasi akademik, tes psikologi, dan tes fisik, (2) pemenuhan fasilitas untuk proses belajar mengajar, (3) suasana yang ideal untuk kegiatan belajar, (4) Tingkat profesionalisme guru yang tinggi, (5) perbaikan kurikulum untuk memenuhi siswa yang memiliki motivasi tinggi dalam belajar, (6) jam belajar siswa yang panjang, (7) kualitas yang lebih baik dalam proses pembelajaran, (8) memiliki manfaat bagi lingkungan sekitar sekolah (Depdikbud, 1994).

2.1. Jenis Data Pendidikan

Pada bagian ini dibahas desain konseptual data pendidikan berdasarkan proses pendidikan. Ada tiga karakteristik utama data pendidikan yang dianggap sebagai bentuk yang paling penting dari model pendidikan, yaitu:

a. Relasional

Proses pendidikan atau proses pembelajaran memiliki empat entitas utama yaitu: siswa, guru, sekolah, dan substansi pembelajaran. Hubungan antara guru dan substansi belajar menghasilkan kompetensi guru. Ini adalah contoh dari korelasi atau hubungan antara dua entitas. Banyak hubungan bisa dibangun dengan memilih dua entitas, tiga entitas, atau empat entitas.

b. Longitudinal

Proses pendidikan merupakan proses jangka panjang. Dalam pendidikan formal, proses dicatat dari *playgroup* atau TK (Taman Kanak-kanak), diikuti oleh enam tahun proses belajar di sekolah dasar, tiga tahun di sekolah menengah pertama, tiga tahun di sekolah menengah atas, dan diakhiri dengan proses belajar di sebuah universitas/pendidikan tinggi.

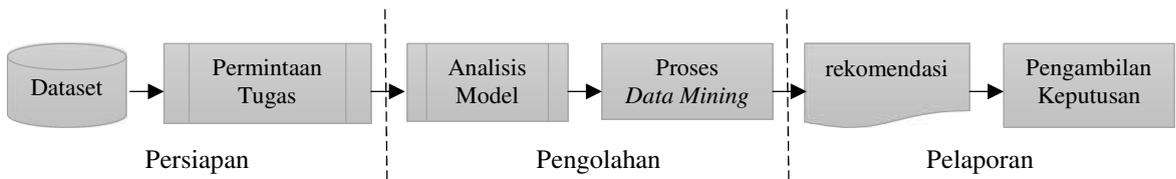
c. Individu

Istilah individu yang dimaksud disini adalah kebalikan dari data agregat. Dalam data agregat, jumlah total diberikan kepada variabel, biasanya didasarkan pada survei, misalnya jumlah guru di sekolah, atau jumlah total siswa untuk tahun pertama di sekolah. Data agregat ini membatasi kemungkinan analisis data. Daripada hanya menyimpan data agregat, detail dari setiap individu juga harus dicatat dalam *database*.

2.2. Sistem Cerdas *Monitoring Pendidikan*

Untuk mengotomatisasi dan mengatur sistem operasional di lingkungan pendidikan, dibuatlah sistem *database* dalam bentuk sistem informasi sekolah. Sistem ini mendukung sebagian besar pengolahan data yang dibutuhkan untuk menghasilkan berbagai laporan. Sebuah sistem informasi sekolah yang baik, seharusnya dapat sekaligus berfungsi sebagai sistem pendukung keputusan dengan mengintegrasikan kecerdasan bisnis ke dalam sistem. Integrasi kecerdasan bisnis mengubah sistem *database* transaksional menjadi sistem data *warehouse* yang mendukung analisis *data mining* yang lebih canggih.

Dengan integrasi kecerdasan bisnis dalam sistem informasi sekolah, proses pendidikan di lingkungan tertentu dapat dimodelkan dan dipahami lebih baik. Model konseptual yang diusulkan pada paper ini ditunjukkan oleh Gambar 1 berikut.



Gambar 1. Diagram Model Konseptual dalam EDM

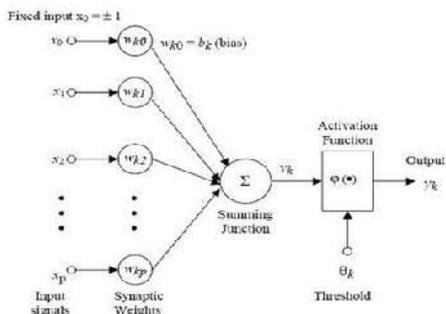
Dalam model konseptual tersebut, dataset statistik pendidikan dihasilkan dari laporan utama pendidikan atau laporan periodik yang didapat dari sistem informasi sekolah. Setiap tugas analisis didorong oleh model pendidikan yang menyediakan sistem dengan daftar struktur data yang diperlukan dan spesifikasi input data (lingkup data). Pengolahan data dibagi menjadi tiga fase yang serupa dengan *data mining*, yaitu persiapan, pengolahan, dan pelaporan.

2.3. Jaringan Saraf Tiruan

Neural Networks (Artificial Neural Networks atau Jaringan Saraf Tiruan) merupakan sebuah metode *softcomputing* atau *data mining* yang banyak digunakan untuk melakukan pengklasifikasian dan prediksi. Jaringan Saraf Tiruan (JST) pertama kali dikembangkan oleh McCulloch dan Pitts pada tahun 1943, dan sekarang ini telah banyak dikembangkan menjadi bentuk yang bermacam-macam. Perubahan bentuk ini dapat berupa perubahan *activation function*, *topology*, *learning algorithm* dan lain-lain. Pada penelitian ini digunakan Jaringan Saraf Tiruan *backpropagation*.

2.3.1. Algoritma *Backpropagation*

Salah satu algoritma yang sering digunakan dalam melakukan learning terhadap JST adalah *backpropagation algorithm*. Algoritma ini dikembangkan oleh Rumelhart dan McClelland pada tahun 1986. Sebuah *neural networks* umumnya terdiri dari *input*, *output* dan *hidden layer*. Dalam algoritma *Backpropagation*, proses *learning* dilaksanakan, pertama melakukan proses *feed forward* dengan mengirimkan sinyal ‘forward’. Adapun proses yang dilakukan dalam *feed forward* ini untuk sebuah neuron dijabarkan dalam gambar 2.



Gambar 2. Activation of single neuron

Algoritma *backpropagation* terdiri dari arah maju dan arah balik seperti berikut ini:

Tahap 0: pembobotan awal (set ke nilai random serendah mungkin), set harga *error* minimal.

Tahap 1 : Ketika kondisi *stop, false*, lakukan tahap 2-9.

Tahap 2 : Untuk setiap pasangan *training*, lakukan tahap 3-8.

Feedforward :

Tahap 3 : tiap unit masukan ($X_i, i = 1, \dots, n$), menerima sinyal x_i dan menyebarkan sinyal ke seluruh lapis tersembunyi (*hidden layer*).

Tahap 4 : tiap unit tersembunyi ($Z_j, j = 1, \dots, p$), jumlahkan bobot sinyal inputnya,

$$Z_{in_j} = v_{oj} + \sum_{i=1}^p x_i v_{ij} \quad (1)$$

Terapkan fungsi aktivasi untuk menghitung sinyal keluarannya, $Z_j = f(z_{in_j})$,

Tahap 5 : tiap unit keluaran ($Y_k, k = 1, \dots, m$), jumlahkan bobot sinyal keluarannya,

$$Y_{in_k} = w_{ok} + \sum_{j=1}^m x_j w_{jk} \quad (2)$$

$$Y_{in_k} = w_{ok} + \sum_{j=1}^m x_j w_{jk} + \sum_{i=1}^f x_i w_{ij} \quad (3)$$

Terapkan fungsi aktivasi untuk menghitung sinyal keluarannya,

$$y_k = f(z_{in_k}) \quad (4)$$

Backpropagation Error:

Tahap 6 : tiap unit keluaran ($Y_k, k = 1, \dots, m$), menerima pola target dan mengacu ke target masukan, hitung kesalahannya

$$\delta_k = (t_k - y_k) f'(y_{in_k}) \quad (5)$$

Hitung koreksi bobot,

$$\Delta w_{jk} = \alpha \delta_k x_j \quad (6)$$

Hitung koreksi terhadap bias,

$$\Delta w_{ok} = \alpha \delta_k \quad (7)$$

Dan mengirimkan sinyal tersebut ke lapis sebelumnya (mundur).

Tahap 7 : tiap unit tersembunyi ($Z_j, j = 1, \dots, p$), menjumlahkan delta masukan dari lapis di atasnya.

$$\delta_{in_j} = \sum_{k=1}^m \delta_k w_{jk} \quad (8)$$

Dan hitung bobot koreksinya,

$$\Delta v_{ij} = \alpha \delta_{in_j} x_i \quad (9)$$

Dan hitung koreksi biasnya,

$$\Delta v_{oj} = \alpha \delta_{in_j} \quad (10)$$

Update bobot dan bias:

Tahap 8 : tiap unit keluaran ($Y_k, k = 1, \dots, m$), update bias dan bobot-bobotnya ($j = 0, \dots, p$):

$$w_{jk}(\text{new}) = w_{jk}(\text{old}) + \Delta w_{jk} \quad (11)$$

Tiap unit tersembunyi ($Z_j, j = 1, \dots, p$), update bias dan bobotnya ($i = 0, \dots, n$):

$$v_{ij}(\text{new}) = v_{ij}(\text{old}) + \Delta v_{ij} \quad (12)$$

Tahap 9 : Uji kondisi berhenti.

if $\delta_k <$ harga *error* set awal *then* "stop training".

3. HASIL DAN PEMBAHASAN

Pada bagian ini dibahas mengenai contoh implementasi dari model konseptual yang diusulkan. Data primer untuk implementasi ini didapatkan dari dataset sistem informasi sekolah. Pada penyelidikan permintaan tugas didefinisikan sebagai "memprediksi keberhasilan sekolah menggunakan jaringan saraf tiruan". Penelitian ini dimulai dari temuan statistik periodik saat menganalisis keberhasilan sekolah berdasarkan variabel siswa saja.

Untuk lebih memahami pengaruh lingkungan dalam keberhasilan proses belajar mengajar, satu set variabel dari entitas siswa, entitas guru, dan entitas sekolah akan diamati. Model ini secara otomatis mengatur kebutuhan inputan yang merupakan daftar variabel dari dataset.

3.1. Tahap I : Persiapan Data

Dalam *data mining*, persiapan data terdiri dari 3 proses: seleksi data, *preprocessing*, dan transformasi data. Tujuan dari *preprocessing* adalah untuk mengubah dataset sehingga isi informasi yang akan diproses sesuai dengan format yang dibutuhkan. Persiapan data dapat menghabiskan cukup banyak sumber daya manusia seperti, waktu pemrosesan, tenaga dan biaya.

3.1.1. Seleksi Data

Dalam proses seleksi data, diidentifikasi semua variabel yang akan digunakan dalam penelitian ini. Seperti terlihat pada Tabel 1, ada lima kolom: Parameter, deskripsi, relasi variabel, data individual, dan data longitudinal. Parameter adalah sesuatu yang akan diukur milik variabel siswa, guru, dan sekolah. Beberapa parameter dapat diperoleh dari hubungan antara variabel. Data individu adalah data non agregat. Dalam statistik, data agregat menggambarkan data gabungan dari beberapa pengukuran. Longitudinal adalah pengumpulan data sampel dari populasi yang lebih besar dengan jangka waktu tertentu.

3.1.2. Pembersihan Data

Proses ini digunakan untuk menghapus data ganda, memeriksa data yang tidak konsisten, penanganan terhadap nilai data yang hilang dan merapikan dari data yang bersifat *outlier*.

Tabel 1. Seleksi Variabel

Parameter	Deskripsi	Relasi variabel	Data Individual	Data Longitudinal
Nama	Nama siswa	Siswa	Ya	Tidak
Kelas	Kelas siswa	Siswa – Sekolah	Ya	Ya
JenKel	Jenis kelamin siswa	Siswa	Ya	Tidak
RataNilai	Nilai rata-rata siswa	Siswa – Sekolah	Tidak	Ya
Alamat	Alamat rumah siswa	Siswa	Ya	Ya
Jarak	Jarak rumah ke sekolah	Siswa – Sekolah	Ya	Ya
StatTinggal	Status tinggal siswa	Siswa	Ya	Ya
PendidikanIbu	Pendidikan ibu siswa	Siswa	Ya	Ya
PekerjaanIbu	Pekerjaan ibu siswa	Siswa	Ya	Ya
PendidikanAyah	Pendidikan ayah siswa	Siswa	Ya	Ya
PekerjaanAyah	Pekerjaan ayah siswa	Siswa	Ya	Ya
AnggotaKeluarga	Jumlah anggota keluarga	Siswa	Ya	Ya
HubKeluarga	Hubungan intern keluarga	Siswa	Ya	Ya
AlasanPilih	Alasan memilih sekolah	Siswa – Sekolah	Ya	Tidak
WaktuBelajar	Waktu belajar seminggu	Siswa	Tidak	Ya
Ekstrakurikuler	Kegiatan ekstrakurikuler	Siswa – Sekolah	Tidak	Ya
DukunganSekolah	Dukungan dari sekolah	Siswa – Sekolah	Tidak	Tidak
DukunganKelu	Dukungan dari keluarga	Siswa	Ya	Ya
InternetRumah	Internet di rumah	Siswa	Ya	Ya
InternetSekolah	Internet di sekolah	Siswa – Sekolah	Ya	Ya
Handphone	Menggunakan handphone	Siswa	Ya	Ya
LanjutSekolah	Ingin lanjut kuliah	Siswa	Ya	Ya
Pacaran	Status pacaran	Siswa	Ya	Ya
WaktuBermain	Waktu untuk bermain	Siswa	Tidak	Ya
WaktuLuang	Waktu luang	Siswa	Tidak	Ya
Kesehatan	Status kesehatan	Siswa	Ya	Ya
Absensi	Jumlah absen	Siswa	Tidak	Ya
Akreditasi	Status akreditasi	Sekolah	Ya	Ya
LamaBerdiriSek	Usia sekolah	Sekolah	Tidak	Ya
StatusSertifikasi	Status sertifikasi guru	Guru	Ya	Ya
LamaMengajar	Lama mengajar guru	Guru	Tidak	Ya
UsiaGuru	Usia guru	Guru	Ya	Ya

3.1.3. Transformasi Data

Pada fase ini, data dapat diubah atau direnovasi berdasarkan kebutuhan model data pendidikan. Dalam percobaan ini, sebagai contoh parameter jarak rumah siswa ke sekolah diubah ke dalam tiga kategori (peringkat) seperti yang ditunjukkan pada Tabel 2 dan 3. Langkah ini memastikan setiap baris berisi data seragam yang cocok untuk skala klasifikasi.

Tabel 2. Skala Peringkat

Kategori	Jarak (kilometer)
1	> 20
2	11 – 20
3	0 – 10

Tabel 3. Contoh Nilai Parameter Jarak

Jarak	Nilai
13.9	2
2.6	3
7.2	3
11.9	2
13.8	2
6.7	3

3.1.4. Normalisasi Data

Sebelum menerapkan input data dan data target ke dalam Jaringan Saraf Tiruan (JST), pertama input data harus dinormalisasi. Normalisasi adalah salah satu dari beberapa teknik dalam data transformasi, yang mengubah data asli ke berbagai data antara 0 dan 1 sesuai dengan fungsi aktivasi yang akan digunakan. Fungsinya adalah untuk mengenali data yang akan menjadi bobot input.

$$normalisasi = (A - A_{min}) / (A_{max} - A_{min}) \tag{13}$$

dimana A adalah variabel nilai input, A_{min} adalah nilai minimum dari input, dan A_{max} adalah nilai maksimum input.

3.2. Tahap II : Pengolahan

Pada fase pengolahan ini digunakan JST sebagai metode untuk mengevaluasi hubungan multivariat antara variabel yang mempengaruhi keberhasilan sekolah. Jumlah data yang dikumpulkan adalah sebanyak 117. Dimana 75 data digunakan untuk proses pelatihan, dan 42 data sisanya untuk proses pengujian. Langkah pertama adalah proses pelatihan menggunakan bagian dataset yang telah ditentukan. Kemudian, proses pengujian dengan sampel data sisanya dijalankan. Proses ini dilatih dan diuji menggunakan *toolbox* JST dari *software* MATLAB dengan variasi arsitektur yang ditentukan pada Tabel 4. Hasil dari proses ini dapat dilihat pada tabel 5 dan 6 yang kemudian akan dibandingkan rata-rata perbedaan tingkat *error*-nya.

Tabel 4. Struktur JST yang Digunakan

Karakteristik	Spesifikasi
Arsitektur	1 <i>hidden layer</i>
<i>Neuron input</i>	Iterasi pertama 30 <i>node</i>
<i>Hidden node</i>	5, 10, 15
Fungsi pelatihan	<i>Resilent backpropagation, levenberg-marquardt</i>
<i>Neuron layer output</i>	2
Toleransi <i>error</i>	0,001
<i>Learning rate</i>	0.01, 0.05, 0.1, 0.5
Maksimum <i>epoch</i>	2000

Tabel 5. Hasil Uji Coba dengan menggunakan variabel Siswa

No	Siswa	Target	Output	Perbedaan
1	0.62	0.428	0.4303	0.0023
2	0.08	0.192	0.1925	0.0005
3	0.69	0.552	0.5609	0.0089
4	0.46	0.128	0.1166	0.0114
5	0.31	0.24	0.2365	0.0035
6	0.31	0.68	0.6738	0.0062
7	0.08	0.24	0.2365	0.0035
8	0.85	0.3	0.2941	0.0059
9	0.15	0.508	0.5075	0.0005
10	0.54	0.252	0.2614	0.0094
11	0.38	0.3	0.2941	0.0059
12	0.85	0.492	0.495	0.003
13	0.85	0.668	0.6298	0.0382
14	0.46	0.372	0.3661	0.0059
15	0.23	0.56	0.5635	0.0035
16	0.77	0.628	0.6249	0.0031
17	0.46	0.368	0.3686	0.0006
18	0.77	0.628	0.6249	0.0031
19	1.00	0.8	0.7509	0.0491
20	0.77	0.632	0.6307	0.0013
21	0.62	0.5	0.4983	0.0017
22	0.38	0.3	0.2941	0.0059
23	0.46	0.36	0.3535	0.0065
24	0.54	0.42	0.4147	0.0053
25	0.46	0.36	0.3535	0.0065
Rata-rata perbedaan error				0.009511412

Tabel 6. Hasil Uji Coba dengan menggunakan variabel Siswa, Guru dan Sekolah

No	Siswa	Guru + Sekolah	Target	Output	Perbedaan
1	0.62	0.5	0.248	0.258	0.01
2	0.08	0.625	0.76	0.746	0.014
3	0.69	0.875	0.44	0.435	0.005
4	0.46	0.75	0.486	0.484	0.002
5	0.31	0.75	0.528	0.529	0.001
6	0.31	0.75	0.672	0.673	0.001
7	0.08	0.5	0.71	0.724	0.014
8	0.85	0.5	0.476	0.477	0.001
9	0.15	0.75	0.438	0.434	0.004
10	0.54	0.5	0.248	0.247	0.001
11	0.38	0.75	0.672	0.673	0.001
12	0.85	0.625	0.528	0.529	0.001
13	0.85	0.5	0.476	0.477	0.001
14	0.46	0.75	0.574	0.573	0.001
15	0.23	0.5	0.286	0.292	0.006
16	0.77	0.625	0.526	0.526	0
17	0.46	0.25	0.2	0.21	0.01
18	0.77	0.625	0.626	0.623	0.003
19	1.00	0.5	0.576	0.577	0.001
20	0.77	0.875	0.724	0.721	0.003
21	0.62	0.75	0.528	0.529	0.001
22	0.38	1	0.672	0.673	0.001
23	0.46	0.625	0.71	0.724	0.014
24	0.54	1	0.576	0.577	0.001
25	0.46	0.75	0.85	0.851	0.001
Rata-rata perbedaan error				0.004581231	

3.3. Tahap III : Pelaporan

Hasil pengolahan dan analisis di fase terakhir kemudian akan dikirim ke pihak sekolah yang berwenang dalam mengambil keputusan. Dari tahap pengolahan, hasil pengujian yang hanya berdasarkan variabel siswa disajikan secara sederhana seperti yang ditunjukkan pada Tabel 5. Sedangkan hasil pengujian berdasarkan variabel siswa, guru dan sekolah disajikan pada Tabel 6. Dalam paper ini, tabel pengujian data yang ditunjukkan hanya 25 baris pertama untuk efisiensi halaman.

Nilai target dalam dua tabel tersebut, diperoleh dari pengolahan variabel siswa dan variabel guru + sekolah. Prioritas untuk masing-masing variabel: yaitu variabel siswa sebesar 60% dan variabel Guru + Sekolah sebesar 40%. Kolom terakhir di setiap tabel menunjukkan perbedaan *error* yang diperoleh dari selisih antara target dan *output*.

Rata-rata perbedaan *error* dalam pengujian berdasarkan faktor siswa, guru dan sekolah adalah 0.004581231. Sementara pengujian berdasarkan faktor siswa saja memiliki rata-rata *error* yang lebih besar, yaitu 0.009511412. Hal ini menunjukkan bahwa variabel siswa, guru dan sekolah memiliki pengaruh yang signifikan terhadap keberhasilan sekolah dibanding variabel siswa saja. Sehingga dapat disimpulkan keberhasilan sekolah tidak hanya bergantung dari satu aspek saja, namun sangat bergantung terhadap 3 aspek yaitu siswa, guru, dan sekolah.

Langkah terakhir dari analisis hasil tersebut adalah membuat laporan baik berupa laporan statistik maupun laporan yang bersifat rekomendasi. Laporan rekomendasi tersebut berisi saran maupun tindakan-tindakan yang dapat diambil oleh pengambil kebijakan untuk meningkatkan pengambilan keputusan yang lebih tepat dan bersifat preventif.

4. KESIMPULAN

Penelitian ini memperkenalkan konsep pengorganisasian data pendidikan yang mendukung sistem *database* transaksional dan sistem cerdas menggunakan pendekatan *data mining*. Dalam konsep struktur *database*, penelitian ini menekankan pentingnya karakteristik data pendidikan yaitu individu, relasional, dan longitudinal. Dalam proses analisis *data mining*, paper ini mengusulkan sebuah model konseptual yang menentukan rekomendasi kepada pengambil keputusan. Dalam desain yang diusulkan ini, analisis didorong oleh model pendidikan, dan fase pengolahan dibagi dengan menggunakan tiga fase yang umum digunakan dalam pendekatan *data mining*.

DAFTAR PUSTAKA

- Araque F., Roldan C., Salguero A., (2009), *Factors Influencing University Drop Out Rates*. Computers & Education, 53, 563–574.
- Depdikbud, (1994), *Pengembangan Sekolah Unggul* Jakarta, Depdikbud.
- Hidayat M. M., Purwitasari D., dan Ginardi H., (2013), *Analisis Prediksi DO Mahasiswa dalam Educational Data Mining menggunakan Jaringan Syaraf Tiruan*, Jurnal IPTEK Vol 17 No.2 Desember 2013.
- Lipsitz, J. (1984), *Successful Schools For Young Adolescents*. New Brunswick, NJ: Transaction.
- Vera, C. Marquez, Romero C., Ventura, S., (2011), *Predicting School Failure Using Data Mining*. Proceedings of the Fourth International Conference on Data Mining, 271-275.
- www.educationaldatamining.org, diakses terakhir: 12 Agustus 2015, jam 22.30.