

Penggalian Pola Kemampuan Peserta Ujian Berbasis Klaster untuk Penentuan Aturan Sistem Penilaian

Umi Laili Yuhana^{1,2}, Eko M. Yuniarno³, Supeno Mardi S. Nugroho⁴, Siti Rochimah⁵, Mauridhi Hery Purnomo⁶

Abstract—Determination of initial ability of examinees is one of the important stages in the adaptive assessment system. The accuracy of the examinee's ability level prediction will influence the appropriateness of chosen item difficulty level for each examinee. This paper discusses the patterns mining of cognitive ability based on cluster using K-Means. The K-means method is utilized to mine the examinees' ability pattern from examinees' pretest answers. The patterns are used for developing rules to determine examinee's ability level in the adaptive assessment system. The addition of this method is proposed to improve the performance of the prediction methods to predict the examinees' ability level. Extraction of graduation value at each level is done before the pattern excavation process. Patterns found become the basis of making the rules as well as replace the rules from the experts in previous studies. The prediction of participants' ability is done by implementing rule based method classifier. A total of 140 data were used for the experiment. Based on the results of the experiment, it can be concluded that the cluster-based pattern mining using K-means can be utilized to determine the cognitive ability level of examinee. The application of this method to student pretest data shows the performance improvement of all the prediction methods used in this paper, i.e. Naive Bayes, MLP, SMO, Decision Table, JRIP, and J48. This method is suitable for adaptive assessment system where the rules can be adjusted along with the addition of the number of the data as well as the addition of the number of variations in the ability pattern of examinees.

Intisari—Penentuan kemampuan awal peserta ujian merupakan salah satu tahapan penting dalam sistem penilaian adaptif. Akurasi prediksi tingkat kemampuan peserta mempengaruhi ketepatan pemilihan tingkat kesulitan soal yang sesuai bagi tiap peserta. Makalah ini membahas penggalian pola kemampuan kognitif peserta ujian berbasis klaster menggunakan *K-means* pada sistem penilaian adaptif. Metode *K-means* digunakan untuk menggali pola kemampuan peserta ujian dari jawaban tes awal (*pretest*) peserta. Pola yang diperoleh digunakan untuk membangun aturan penentuan tingkat kemampuan peserta ujian pada sistem penilaian adaptif. Tujuan implementasi metode ini adalah meningkatkan kinerja metode

prediksi untuk memprediksi tingkat kemampuan peserta ujian.

Ekstraksi fitur nilai kelulusan peserta ujian pada tiap tingkatan dilakukan sebelum proses penggalian pola. Berdasarkan hasil uji coba, dapat disimpulkan bahwa penggalian pola berbasis klaster menggunakan *K-means* dapat digunakan untuk menentukan tingkat kemampuan peserta ujian. Prediksi kemampuan peserta dilakukan dengan mengimplementasikan metode pengklasifikasi berbasis aturan. Pola yang ditemukan dijadikan dasar pembuatan aturan sekaligus menggantikan aturan dari pakar pada penelitian sebelumnya. Sebanyak 140 data digunakan untuk uji coba. Penerapan metode ini pada data ujian awal siswa terbukti dapat menunjukkan peningkatan kinerja semua metode prediksi yang digunakan dalam makalah ini, yakni metode *Naive Bayes*, MLP, SMO, *Decision Table*, JRIP, dan J48. Metode ini cocok untuk sistem penilaian adaptif dengan aturan dapat disesuaikan seiring penambahan jumlah data ujian maupun penambahan jumlah variasi pola kemampuan peserta ujian.

Kata Kunci— sistem penilaian, penentuan tingkat kemampuan peserta, *K-means*, metode pengklasifikasi berbasis aturan.

I. PENDAHULUAN

Saat ini teknologi informasi telah banyak digunakan untuk membantu proses pembelajaran. Di tingkat sekolah menengah atas, *e-learning* berbasis *web* dibangun sesuai kebutuhan sekolah dengan memanfaatkan LMS [1]. Bahkan, sebuah *e-learning* berbasis konteks pengguna telah dibangun untuk menyediakan materi belajar yang sesuai dengan gaya belajar siswa [2] maupun kondisi personal pengguna saat belajar [3].

Selain membantu proses pembelajaran, teknologi informasi juga diterapkan untuk penilaian pembelajaran. Penilaian pembelajaran merupakan proses evaluasi capaian pengetahuan, pemahaman, dan keterampilan pembelajar [4]. Penilaian dapat digunakan untuk beberapa kebutuhan, antara lain ujian masuk, dasar perancangan materi pembelajaran, pengukuran kompetensi, pengukuran capaian pembelajaran, serta penentuan kelulusan [5]. Sistem penilaian berbasis komputer mulai banyak digunakan. Ujian Nasional Berbasis Komputer (UNBK) sangat cocok digunakan untuk ujian masuk, merancang materi pembelajaran, maupun ujian kelulusan [6]. Namun, untuk mengukur kompetensi ataupun capaian pembelajaran, diperlukan sebuah sistem yang dapat memilih dan menyajikan soal yang sesuai dengan kemampuan peserta ujian. Sistem penilaian adaptif dapat digunakan untuk memenuhi kebutuhan ini karena sistem ini dapat memilih dan menyajikan soal sesuai kemampuan peserta ujian [7].

Sistem penilaian adaptif diekspresikan sebagai pendekatan interaktif untuk menilai pembelajar dalam sistem pembelajaran [8]. Tahapan yang dilakukan dalam pengembangan sistem ini antara lain penentuan bank soal, penentuan tingkat kemampuan awal peserta ujian, pemilihan

¹ Mahasiswa, Departemen Teknik Elektro, Fakultas Teknologi Elektro, Institut Teknologi Sepuluh Nopember, Jln. Teknik Mesin, Gedung B, C, dan AJ, Kampus ITS, Sukolilo, Surabaya, Jawa Timur, 60111 Indonesia (telp: 031- 5994251; fax: 031-5931237; e-mail: yuhana@if.its.ac.id)

^{3, 4, 6} Dosen, Departemen Teknik Komputer, Fakultas Teknologi Elektro, Institut Teknologi Sepuluh Nopember, Gedung B & C, Jln. Teknik Mesin, Kampus ITS, Sukolilo, Surabaya, Jawa Timur, 60111 Indonesia (telp: 031- 5922936; e-mail: ekomulyanto@ee.its.ac.id, mardi@its.ac.id, hery@ee.its.ac.id)

^{2, 5} Dosen, Departemen Informatika, Fakultas Teknologi Informasi dan Komunikasi, Institut Teknologi Sepuluh Nopember, Jln. Teknik Kimia, Gedung Informatika, Kampus ITS, Sukolilo, Surabaya, Jawa Timur, 60111 Indonesia (telp: 031- 5939214; fax: 031-5913804; e-mail: yuhana@if.its.ac.id, siti@if.its.ac.id)

soal, penilaian, penghentian tes, dan pengambilan kesimpulan tentang kemampuan peserta ujian [9]. Penentuan tingkat kemampuan awal peserta ujian merupakan hal yang sangat penting karena ketepatannya sangat memengaruhi efektivitas pemilihan soal.

Sistem penilaian adaptif telah banyak dikembangkan. Namun, masih sedikit yang fokus pada penentuan tingkat kemampuan peserta ujian. Tingkat kemampuan peserta ujian dapat diklasifikasikan dalam beberapa kategori, antara lain berdasarkan konsep yang sedang dipelajari, kompetensi yang dimiliki, tingkat *bloom*, maupun kelas yang sedang ditempuh. Pada sebuah penelitian, metode *Naive Bayes* digunakan untuk menentukan tingkat kemampuan peserta ujian pada konsep yang dipelajari [10]. Penelitian lain mengklasifikasikan kemampuan peserta ujian berdasarkan taksonomi *bloom* [11]. Dalam sebuah penelitian yang lain, tipe populasi dijadikan sebagai dasar dalam menentukan soal ujian [12]. Pada populasi yang homogen dan tidak terdapat informasi tentang tingkat kemampuan peserta ujian, dipilih soal dengan tingkat kesulitan medium sebagai soal awal. Pada populasi yang heterogen dan terdapat informasi tentang tingkat kelas peserta, dipilih soal dengan tingkat kesulitan medium sesuai tingkatan kelasnya.

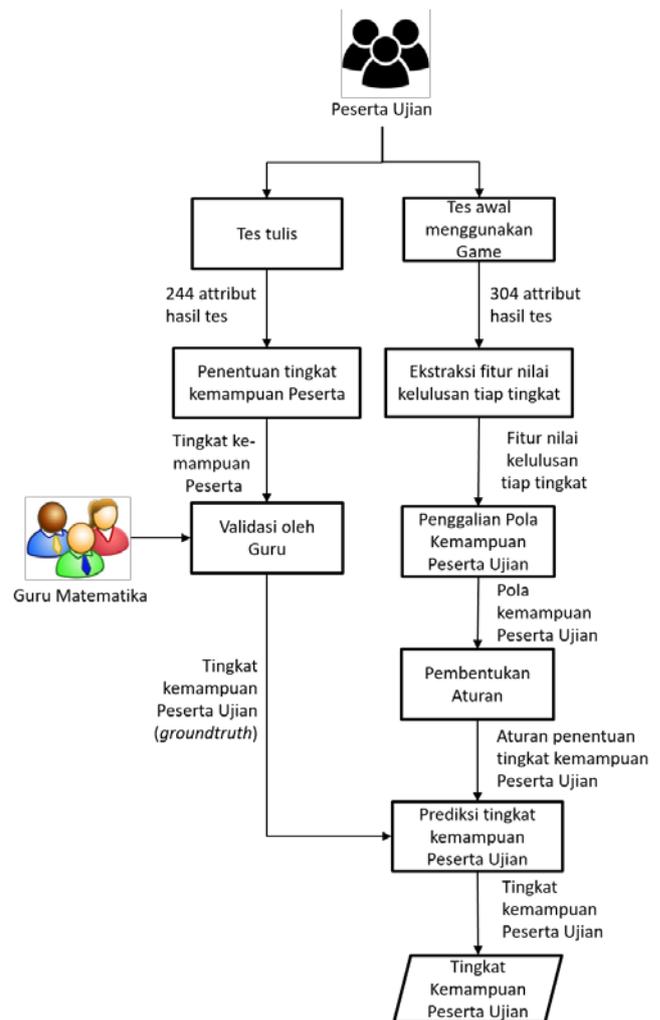
Metode berbasis aturan digunakan untuk mengekstrak informasi pada laporan hasil pemeriksaan BPK RI [13]. Pada penelitian sebelumnya, metode berbasis aturan digabung dengan teknik pembelajaran mesin diusulkan untuk menilai tingkat kemampuan siswa reguler dan siswa berkebutuhan khusus [14]. Teknik pembelajaran mesin yang digunakan adalah *Naive Bayes*, *Multilayer Perceptron*, *SMO*, *Decision Tree*, *JRIP*, dan *J48*. Hasil akurasi terbaik dicapai dengan menggunakan metode berbasis aturan *JRIP* sebesar 64,12. Aturan untuk penentuan tingkat kemampuan dibentuk berdasarkan pendapat pakar. Kemampuan peserta ujian yang bervariasi dan jumlah data yang berkembang melatarbelakangi perlunya pembentukan aturan secara dinamis. Penemuan pola pada data ujian peserta dapat digunakan sebagai dasar pembentukan aturan menggantikan pakar serta meningkatkan akurasi prediksi. Diperlukan penggalian pola tersebut sehingga dapat digunakan untuk pembentukan aturan penentuan kemampuan awal peserta ujian.

Makalah ini membahas penggalian pola kemampuan peserta ujian berbasis kluster dengan *K-means* untuk penentuan aturan pada sistem penilaian adaptif. Penambahan penggunaan *K-means* ditujukan untuk menggali pola kelompok kemampuan peserta ujian dari jawaban tes awal (*pretest*) peserta dan memperbaiki akurasi penentuan tingkat kemampuan peserta ujian. Pola yang ditemukan dijadikan dasar pembuatan aturan sekaligus menggantikan aturan dari pakar pada penelitian sebelumnya. Metode yang diusulkan dibahas pada bagian II. Uraian hasil dan analisis dibahas pada bagian III. Bagian IV berisi kesimpulan dan penelitian lebih lanjut.

II. USULAN METODE

Tahapan untuk penentuan tingkat kemampuan peserta ujian yang diusulkan ditunjukkan pada Gbr. 1. Tingkat kemampuan peserta ujian ditentukan berdasarkan data hasil tes awal. Tes awal dilakukan untuk mendapatkan data kemampuan peserta

ujian. Setiap peserta ujian melakukan dua jenis tes, secara tertulis dan menggunakan *game* matematika. Soal yang diberikan pada tes tulis dan tes dengan *game* adalah sama.



Gbr. 1 Tahapan penentuan tingkat kemampuan peserta

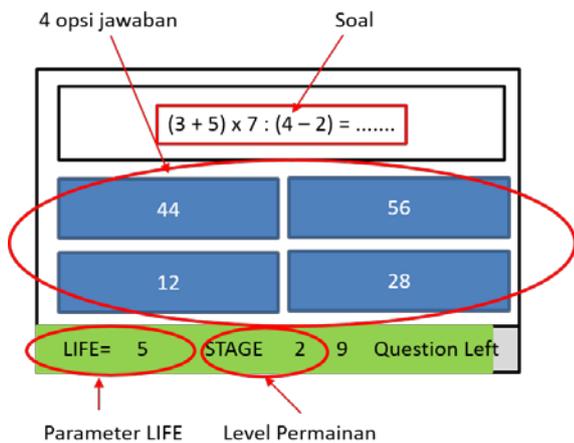
Terdapat 140 *record* hasil tes tulis dan 140 *record* hasil tes menggunakan *game*. Setiap satu *record* data merepresentasikan satu peserta tes. Tiap *record* hasil tes tulis terdiri atas 244 atribut. Atribut pada tes tulis adalah nama, umur, kelas, sekolah, soal yang dijawab (60 soal), jawaban benar (60 jawaban), kompetensi yang terkait dengan soal (60 kompetensi), jawaban responden (60 jawaban), dan status jawaban responden (60 status). Status jawaban bernilai 0 atau 1, dengan 0 untuk jawaban salah dan 1 untuk jawaban benar. Hasil tes menggunakan *game* terdiri atas 304 atribut, 240 atribut sama dengan atribut tes tulis, ditambah dengan atribut waktu yang digunakan untuk menjawab (60 waktu).

Hasil tes tulis diolah untuk mendapatkan tingkat kemampuan peserta ujian dan dijadikan *groundtruth* setelah divalidasi oleh guru matematika [14]. Data hasil tes awal menggunakan *game* disimpan dalam *game log*. Ekstraksi fitur nilai kelulusan tiap tingkat dilakukan sebelum proses penggalian pola. Selanjutnya, proses penggalian pola dilakukan dengan proses pengklusteran untuk menemukan

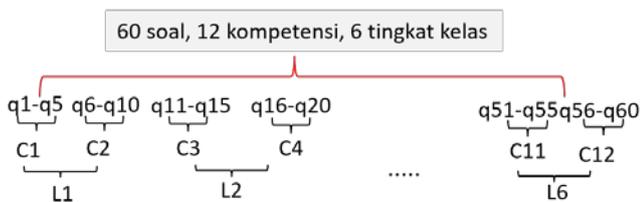
pola jawaban siswa. Pola yang ditemukan kemudian dianalisis untuk pembentukan aturan penentuan tingkat kemampuan peserta ujian. Proses detail dari tiap tahapan, terutama pengolahan data tes dengan *game*, dibahas dalam bagian ini.

A. Tes Awal Menggunakan Game

Game yang digunakan untuk tes awal adalah *game MathBharata* [15]. *Game MathBarata* adalah sebuah *game* kuis matematika yang dirancang untuk menilai kemampuan siswa SD kelas 1 sampai kelas 6. Setiap level/tingkatan diwakili oleh satu *stage*. Gbr. 2 menunjukkan contoh tampilan salah satu soal pada *game MathBarata*. Soal tersebut berada pada *stage* 2 yang berarti berasal dari kelas 2 SD. Terdapat empat opsi jawaban dalam setiap soal. Peserta ujian diminta memilih jawaban yang dianggap benar. Waktu untuk menjawab setiap soal maksimal 120 detik atau 2 menit. Parameter *LIFE* digunakan untuk menentukan *game* berakhir atau tidak. Tiap jawaban benar akan meningkatkan nilai *LIFE*, sebaliknya tiap jawaban salah akan mengurangi nilai *LIFE*. *Game* berakhir jika parameter *LIFE* bernilai 0.



Gbr. 2 Ilustrasi tampilan pada *game MathBarata* [15].



Gbr. 3 Ilustrasi hubungan soal dengan kompetensi dan tingkatan kelas.

Setiap tingkat terdiri atas beberapa kompetensi yang sesuai dengan aturan standar kompetensi [16], [17]. Setiap peserta memulai *game* dengan pertanyaan pada tingkat kelas 1. Gbr. 3 menunjukkan hubungan soal dengan kompetensi dan tingkatan kelas, dengan *q* menyatakan soal, *C* menyatakan kompetensi, dan *L* menyatakan tingkat kelas. Untuk tes awal, terdapat 60 soal yang berhubungan dengan 12 kompetensi dari enam tingkat kelas. Setiap kompetensi terdiri atas lima soal dan setiap tingkat diwakili oleh dua kompetensi. Sebagai contoh, soal pertama sampai kelima berhubungan dengan kompetensi pertama yakni C1. Soal keenam sampai ke-10 berhubungan dengan kompetensi C2. C1 dan C2 berhubungan

dengan capaian pembelajaran yang dapat dicapai oleh siswa pada tingkat kelas 1 atau L1.

TABEL I
JENIS KOMPETENSI MATEMATIKA SD YANG DIGUNAKAN [18]

Kode Kompetensi – Tingkat Kelas	Standar Kompetensi	Kompetensi Dasar
C1-1	Melakukan penjumlahan dan pengurangan bilangan sampai 20	Membilang banyak benda
C2-1	Melakukan penjumlahan dan pengurangan bilangan sampai 20	Melakukan penjumlahan dan pengurangan bilangan sampai 20
C3-2	Melakukan perkalian dan pembagian sampai dua angka	Melakukan perkalian bilangan yang hasilnya bilangan dua angka
C4-2	Melakukan perkalian dan pembagian sampai dua angka	Melakukan operasi hitung campuran
C5-3	Melakukan operasi hitung bilangan sampai tiga angka	Melakukan penjumlahan dan pengurangan tiga angka
C6-3	Melakukan operasi hitung bilangan sampai tiga angka	Melakukan perkalian yang hasilnya bilangan tiga angka dan pembagian bilangan tiga angka
C7-4	Memahami dan menggunakan sifat-sifat operasi hitung bilangan dalam pemecahan masalah	Melakukan operasi hitung campuran
C8-4	Menggunakan pecahan dalam pemecahan masalah	Menjumlahkan pecahan
C9-5	Menggunakan pecahan dalam pemecahan masalah	Mengubah pecahan ke bentuk persen dan desimal serta sebaliknya
C10-5	Menggunakan pecahan dalam pemecahan masalah	Mengalikan dan membagi berbagai bentuk pecahan
C11-6	Melakukan operasi hitung bilangan bulat dalam pemecahan masalah	Melakukan operasi hitung bilangan bulat dalam pemecahan masalah
C12-6	Menyelesaikan masalah yang berkaitan dengan data	Menentukan rata-rata hitung dan modus sekumpulan data

Tabel I menyajikan jenis kompetensi matematika sekolah dasar yang digunakan. Daftar kompetensi yang digunakan diambil dari Standar Kompetensi dan Kompetensi Dasar dari Badan Standar Nasional Pendidikan [18]. Kompetensi dibedakan menjadi dua jenis, standar kompetensi dan kompetensi dasar. Standar kompetensi adalah deskripsi

pengetahuan, keterampilan, dan sikap yang harus dikuasai setelah siswa mempelajari mata pelajaran tertentu pada jenjang pendidikan tertentu.

Kompetensi dasar adalah pengetahuan, keterampilan dan sikap minimal yang harus dicapai oleh siswa untuk menunjukkan bahwa siswa telah menguasai standar kompetensi yang telah ditetapkan. Kompetensi dasar merupakan penjabaran dari standar kompetensi.

Sejumlah 140 siswa SD dari dua sekolah dasar di Jawa Timur menjadi responden untuk tes awal. Jumlah siswa dari SD pertama adalah 79 siswa atau sebesar 44%. Jumlah siswa dari SD kedua sebanyak 61 siswa atau 56%. Berdasarkan jenis kelamin, terdapat 75 siswa laki-laki dan 65 siswa perempuan. Paling banyak responden berasal dari kelas 5, dengan 34 responden berumur 10 tahun dan 65 responden berumur 11 tahun.

B. Ekstraksi Fitur Nilai Kelulusan Tiap Tingkat

Ekstraksi fitur nilai kelulusan tiap tingkat dilakukan untuk mendapatkan fitur nilai kelulusan tiap tingkat dari setiap peserta ujian. Ada tiga langkah yang dilakukan, yaitu menghitung nilai tiap kompetensi, menentukan nilai ketuntasan per kompetensi, dan menghitung nilai kelulusan tiap tingkat. Ketiga proses tersebut secara rinci dijelaskan sebagai berikut.

1) *Menghitung Nilai Tiap Kompetensi:* Nilai tiap kompetensi, N_i^C , dihitung dengan (1). N_i^C menyatakan nilai kompetensi $C = i$ untuk $i = 1$ sampai 12 (dengan jumlah kompetensi yang digunakan adalah 12). Nilai tersebut diperoleh dari rata-rata nilai status jawaban soal Q ke-1 sampai n dan n menyatakan jumlah soal per kompetensi (dalam penelitian ini n bernilai 5).

$$N_i^C = \frac{Q_1 + Q_2 + \dots + Q_n}{n} \tag{1}$$

2) *Menentukan Nilai Ketuntasan per Kompetensi:* Nilai ketuntasan per kompetensi, T_i^C , menyatakan status ketuntasan tiap kompetensi yang dihitung dengan (2). Jika nilai tingkat ke- i kurang dari batas lulus, maka T_i^C bernilai 0, jika tidak maka T_i^C bernilai 1. Pada penelitian ini, minimal batas ketuntasan (b) mengikuti kriteria ketuntasan minimal dari sebagian besar sekolah dasar di Indonesia, yaitu 0,7. Persamaan ini juga dapat digunakan untuk mendeteksi kompetensi yang telah dan belum dikuasai oleh peserta. Contoh hasil pemrosesan nilai ketuntasan per kompetensi ditunjukkan pada Gbr. 4. Pada gambar tersebut terdapat contoh nilai ketuntasan 12 kompetensi dari 18 siswa. Nilai 1 pada C_p menunjukkan peserta tuntas pada C_p tersebut, sebaliknya nilai 0 menunjukkan peserta belum tuntas pada kompetensi tersebut.

$$T_i^C = \begin{cases} 0, & T_i^C < b \\ 1, & \text{else} \end{cases} \tag{2}$$

dengan $i = 1 \dots 6$ adalah tingkat kelas.

Siswa	Cp1	Cp2	Cp3	Cp4	Cp5	Cp6	Cp7	Cp8	Cp9	Cp10	Cp11	Cp12
1	1	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	0	0	0	0	0	0
3	1	1	1	0	0	1	0	0	0	0	0	0
4	1	1	0	0	0	1	0	0	0	0	0	1
5	1	1	1	1	1	1	0	0	1	0	0	0
6	1	1	1	1	1	1	0	0	0	0	0	0
7	1	1	1	0	1	1	0	0	0	0	0	0
8	1	1	0	0	0	0	0	0	0	0	0	0
9	1	1	1	0	0	0	0	0	0	0	0	0
10	1	1	1	0	1	0	0	0	0	0	0	0
11	1	1	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	0	0	0	0	0	0	0	0
13	1	1	1	0	1	1	0	0	0	0	0	0
14	1	1	0	1	1	0	0	0	0	0	0	0
15	1	1	0	0	0	0	0	0	0	0	0	0
16	1	1	0	0	0	1	0	0	0	0	0	0
17	1	1	1	0	1	0	0	0	0	0	0	0
18	1	1	0	0	1	1	0	0	0	0	0	0

Gbr. 4 Contoh hasil pemrosesan ketuntasan kompetensi peserta.

3) *Menghitung Nilai Kelulusan Tiap Tingkat Kelas:* Nilai kelulusan tiap tingkat kelas, N_i^L , dihitung dengan (3). N_i^L menyatakan nilai kelulusan tiap tingkat kelas $L = i$ untuk i dari 1 sampai 6. N_i^L diperoleh dari status nilai ketuntasan seluruh kompetensi per tingkatan. Persamaan (3) menyatakan bahwa jika peserta ujian dapat menuntaskan semua kompetensi dalam tingkat L , maka peserta dinyatakan lulus pada tingkat L . Sebaliknya, jika ada satu kompetensi yang tidak tuntas pada tingkat L , maka peserta ujian dianggap belum lulus pada tingkat L . Fitur N_i^L selanjutnya digunakan dalam proses pengklasteran.

$$N_i^L = \begin{cases} 1, & T_1^C \text{ AND } T_2^C \text{ AND } \dots \text{ AND } T_n^C \\ 0, & \text{else} \end{cases} \tag{3}$$

C. Penggalian Pola Kemampuan Siswa

Penggalian pola kemampuan siswa dilakukan dengan menggunakan pengklasteran. Proses pengklasteran dilakukan untuk menemukan pola kelompok untuk tiap tingkatan. Metode *K-Means* digunakan dengan nilai $k = 6$. Pemilihan nilai k didasarkan pada asumsi bahwa ada enam kelompok tingkatan sesuai tingkat kelas. Fitur yang digunakan dalam proses pengklasteran ada enam fitur, yakni $N_1^L, N_2^L, N_3^L, N_4^L, N_5^L$, dan N_6^L . Fitur ini menyatakan nilai status ketuntasan tiap peserta ujian pada tingkat kelas 1, 2, 3, 4, 5, dan 6. Fitur ini dipilih berdasarkan hasil studi sebelumnya bahwa fitur yang paling berpengaruh terhadap penentuan tingkat kemampuan peserta ujian adalah fitur status kelulusan nilai dari masing-masing tingkat [11]. Hasil dan analisis klaster yang terbentuk ditunjukkan pada Tabel II.

Pola pertama, 100000, bermakna peserta ujian telah lulus tingkat 1 karena memenuhi ketuntasan semua kompetensi pada tingkat 1. Karenanya, peserta tersebut dapat dikategorikan dalam pola kemampuan kelas 1. Pola kedua, 111000, bermakna peserta ujian telah lulus tingkat 1, tingkat 2, dan tingkat 3. Pola ini dapat dikategorikan dalam pola kemampuan kelas 3. Pola ketiga, 110000, bermakna peserta ujian lulus tingkat 1 dan tingkat 2, sehingga peserta ujian

dalam klaster ini dapat dikelompokkan dalam pola kemampuan kelas 2.

TABEL II
HASIL DAN ANALISIS KLAS TERBENTUK

Klaster	Nilai tengah	Jumlah peserta tes	Persentase	Pola kelas
Klaster 0	100000	61	44%	1
Klaster 1	111000	32	23%	3
Klaster 2	110000	15	11%	2
Klaster 3	101000	24	17%	1
Klaster 4	000000	4	3%	0
Klaster 5	111100	4	3%	4

Pola keempat, 101000, memiliki arti peserta ujian lulus tingkat 1 dan tingkat 3, namun tidak lulus tingkat 2, 4, 5, dan 6. Pola ini dapat dikategorikan setara kemampuan kelas 1. Pada penelitian berikutnya, pola ini perlu diteliti lebih jauh karena terdapat 24 peserta di dalamnya. Banyaknya peserta dengan pola ini tidak dapat dianggap sebagai *noise*. Pola kelima, 000000, memiliki arti peserta ujian tidak lulus tingkat manapun, sehingga tidak dapat dikategorikan di kelas manapun atau dapat disebut tingkat 0. Pola keenam, 111100, berarti peserta ujian lulus tingkat 1, 2, 3, dan 4, sehingga peserta pada kelompok ini dapat dikategorikan memiliki kemampuan setara dengan kelas 4.

D. Pembentukan Aturan

Aturan dibentuk dari pola kelas hasil pengklasteran, ditunjukkan pada Gbr. 5.

$If (N_1^L, N_2^L, N_3^L, N_4^L, N_5^L, N_6^L) = (0,0,0,0,0,0)$
 then $L = 0$
 $Else if (N_1^L, N_2^L, N_3^L, N_4^L, N_5^L, N_6^L)$
 $= (1,0,0,0,0,0) or (N_1^L, N_2^L, N_3^L, N_4^L, N_5^L, N_6^L)$
 $= (1,0,1,0,0,0) then L = 1$
 $Else if (N_1^L, N_2^L, N_3^L, N_4^L, N_5^L, N_6^L) = ((1,1,0,0,0,0) L = 2$
 $Else if (N_1^L, N_2^L, N_3^L, N_4^L, N_5^L, N_6^L) = (1,1,1,0,0,0) L = 3$
 $Else if (N_1^L, N_2^L, N_3^L, N_4^L, N_5^L, N_6^L) = ((1,1,1,1,0,0) L = 4$

Gbr. 5 Aturan penentuan tingkat kemampuan peserta ujian.

Aturan pada Gbr. 5 memiliki makna sebagai berikut.

1. Jika pola status kelulusan peserta ujian seperti nilai tengah klaster kelima (000000), maka peserta ujian dikategorikan memiliki kemampuan setara dengan kelas 0.
2. Jika pola kelulusan peserta ujian seperti nilai tengah klaster pertama (100000) atau klaster keempat (101000), maka peserta ujian dikategorikan memiliki kemampuan setara dengan kelas 1.
3. Jika pola kelulusan peserta ujian seperti nilai tengah klaster ketiga (110000), maka peserta ujian dikategorikan memiliki kemampuan setara dengan kelas 2.

4. Jika pola kelulusan peserta ujian seperti nilai tengah klaster kedua (111000), maka peserta ujian dikategorikan memiliki kemampuan setara dengan kelas 3.
5. Jika pola kelulusan peserta ujian seperti nilai tengah klaster keenam (111100), maka peserta ujian dikategorikan memiliki kemampuan setara dengan kelas 4.

Aturan ini digunakan sebagai logika pengambilan keputusan penentuan tingkat kemampuan peserta ujian. Aturan berkembang secara dinamis sesuai dengan penemuan pola baru akibat pertumbuhan data. Perkembangan aturan ini merupakan salah satu kemampuan adaptif dari sistem penilaian adaptif.

E. Prediksi Tingkat Kemampuan Peserta Ujian

Prediksi tingkat kemampuan peserta ujian dilakukan dengan mengimplementasikan metode *Naive Bayes*, *Multi Layer Perceptron* (MLP), SMO, *Decision Table* (DT), JRIP, dan J48 seperti pada 140 data pada penelitian sebelumnya [14]. Hasil prediksi kemudian dibandingkan dengan hasil prediksi pada penelitian sebelumnya dan dianalisis untuk mendapatkan kesimpulan, metode yang diusulkan memiliki kinerja yang lebih baik atau tidak, dibandingkan pendekatan tanpa pengklasteran.

III. HASIL DAN ANALISIS

Bagian ini berisi hasil uji coba untuk prediksi tingkat kemampuan peserta ujian. Teknik evaluasi model *cross validation* dan *percentage split* digunakan untuk mengamati dan menganalisis hasil pengukuran kinerja prediksi menggunakan *Naive Bayes* (NB), MLP, SMO, DT, JRIP, dan J48. Langkah pengujian adalah sebagai berikut.

1. Uji dilakukan dengan menerapkan enam metode prediksi menggunakan *fold* 10, 15, 20, 25, dan 30 serta *percentage split* 75%, 80%, 85%, dan 90% pada data yang diolah menggunakan metode berbasis aturan berdasarkan pakar. Hasil disajikan dalam NB, MLP, SMO, DT, JRIP, dan J48.
2. Uji dilakukan dengan menerapkan enam metode prediksi menggunakan *fold* 10, 15, 20, 25, dan 30 serta *percentage split* 75%, 80%, 85%, dan 90% pada data yang diolah menggunakan metode usulan, yakni metode berbasis aturan berdasarkan pengklasteran (diistilahkan sebagai *Rule from Clustering* (RC)). Hasil disajikan dalam RC_NB, RC_MLP, RC_SMO, RC_DT, RC_JRIP, dan RC_J48.
3. Analisis dilakukan dengan membandingkan hasil prediksi pada nomor 1 dan 2 menggunakan tujuh ukuran pengklasifikasi, yaitu *Kappa*, *Mean Absolute Error* (MAE), *Precision*, *Recall*, *F-Measure*, ROC, dan akurasi.

Tabel III dan Tabel IV menunjukkan hasil kinerja prediksi menggunakan *Naive Bayes*. Ukuran kinerja yang paling optimal adalah menggunakan *fold* 15, 20, 25, dan 30 dan menggunakan *percentage split* 90% pada RC_NB. Pada *fold* 15, 20, 25, dan 30 ini *Naive Bayes* mampu mencapai nilai

maksimal $Kappa = 0,96$, $MAE = 0,01$, $Precision = 0,74$, $Recall = 0,73$, $F-Measure = 0,74$, dan $ROC = 0,97$.

$MAE = 0,02$, $Precision = 0,74$, $Recall = 0,73$, $F-Measure = 0,74$, dan $ROC = 0,97$.

TABEL III
UKURAN KINERJA PREDIKSI MENGGUNAKAN NAIVE BAYES DENGAN CROSS VALIDATION

Metode Naive Bayes (NB)						
Fold	Kappa	MAE	Precision	Recall	F-Measure	ROC
10	0,29	0,14	0,35	0,40	0,36	0,73
15	0,29	0,14	0,35	0,40	0,36	0,71
20	0,3	0,14	0,35	0,40	0,36	0,71
25	0,3	0,14	0,35	0,40	0,36	0,71
30	0,29	0,14	0,35	0,40	0,36	0,70
Metode RC_Naive Bayes (RC_NB)						
10	0,96	0,08	0,74	0,73	0,74	0,97
15	0,96	0,01	0,74	0,73	0,74	0,97
20	0,96	0,01	0,74	0,73	0,74	0,97
25	0,96	0,01	0,74	0,73	0,74	0,97
30	0,96	0,01	0,74	0,73	0,74	0,97

TABEL IV
UKURAN KINERJA PREDIKSI MENGGUNAKAN NAIVE BAYES DENGAN PERCENTAGE SPLIT

Metode Naive Bayes (NB)						
PS	Kappa	MAE	Precision	Recall	F-Measure	ROC
75%	0,36	0,13	0,35	0,38	0,36	0,69
80%	0,27	0,15	0,33	0,35	0,33	0,67
85%	0,26	0,16	0,33	0,44	0,34	0,67
90%	0,21	0,17	0,11	0,34	0,28	1,00
Metode RC_Naive Bayes (RC_NB)						
75%	0,95	0,01	0,74	0,73	0,74	0,97
80%	0,94	0,01	0,61	0,58	0,58	0,94
85%	0,92	0,02	0,48	0,45	0,46	0,92
90%	1,00	0,00	0,50	0,50	0,50	1,00

Sementara dengan *percentage split* 90%, Naive Bayes mampu mencapai nilai maksimal $Kappa = 1$, $MAE = 0$ dan $ROC = 1$, meskipun nilai *precision*, *recall*, maupun *F-Measure* lebih baik ketika menggunakan *percentage split* 75% dibandingkan *percentage split* 90%. Dari hasil tersebut, dapat disimpulkan bahwa pembentukan aturan berdasarkan pengklasteran dapat meningkatkan kinerja prediksi menggunakan Naive Bayes. Penggunaan *fold* dan *percentage split* sama-sama menunjukkan peningkatan hasil yang baik dibandingkan dengan metode pembentukan aturan berdasarkan pendapat pakar.

Hasil kinerja prediksi menggunakan MLP ditunjukkan pada Tabel V dan Tabel VI. Penggunaan pengklasteran untuk membangun aturan terbukti dapat meningkatkan kinerja metode MLP, seperti yang terlihat pada hasil RC_MLP. Penggunaan *fold* 10, 15, 20, 25, dan 30 pada RC_MLP dapat meningkatkan kinerja dengan rata-rata nilai $Kappa = 0,96$,

TABEL V
UKURAN KINERJA PREDIKSI MENGGUNAKAN MLP DENGAN CROSS VALIDATION

Metode MLP						
Fold	Kappa	MAE	Precision	Recall	F-Measure	ROC
10	0,34	0,14	0,39	0,41	0,40	0,66
15	0,32	0,15	0,35	0,32	0,32	0,64
20	0,33	0,14	0,38	0,41	0,39	0,65
25	0,38	0,14	0,44	0,46	0,44	0,71
30	0,29	0,14	0,34	0,43	0,37	0,73
Metode RC_MLP						
10	0,96	0,02	0,74	0,73	0,74	0,98
15	0,96	0,02	0,74	0,74	0,74	0,97
20	0,96	0,02	0,74	0,73	0,74	0,97
25	0,96	0,02	0,74	0,73	0,74	0,98
30	0,96	0,02	0,74	0,73	0,74	0,97

TABEL VI
UKURAN KINERJA PREDIKSI MENGGUNAKAN MLP DENGAN PERCENTAGE SPLIT

Metode MLP						
PS	Kappa	MAE	Precision	Recall	F-Measure	ROC
75%	0,42	0,14	0,44	0,44	0,44	0,77
80%	0,29	0,15	0,21	0,25	0,23	0,80
85%	0,37	0,15	0,33	0,44	0,34	0,67
90%	0,28	0,17	0,35	0,38	0,36	0,81
Metode RC_MLP						
75%	0,95	0,02	0,62	0,58	0,59	0,94
80%	0,94	0,02	0,61	0,58	0,59	0,94
85%	0,92	0,02	0,48	0,45	0,46	0,92
90%	1,00	0,01	0,50	0,50	0,50	1,00

Peningkatan kinerja juga terjadi pada RC_MLP dengan penggunaan *percentage split*. Sama seperti pada RC_NB, nilai $Kappa$ dan MAE terbaik dicapai ketika digunakan *percentage split* 90% yaitu $Kappa = 1$, $MAE = 0$. Namun, nilai *Precision*, *Recall*, dan *F-Measure* terbaik dicapai ketika digunakan *percentage split* 75%, dengan nilai masing-masing $Precision = 0,74$, $Recall = 0,73$, dan $F-Measure = 0,74$.

Tabel VII dan Tabel VIII menunjukkan kinerja prediksi menggunakan SMO. Dengan penggunaan *fold* 10, 15, 20, 25, dan 30 pada RC_SMO, nilai $Kappa$ meningkat menjadi 0,95, $Precision$ menjadi 0,71, $Recall$ menjadi 0,73, $F-Measure$ menjadi 0,72, dan ROC menjadi 0,98, dibandingkan dengan SMO. Namun, nilai MAE cenderung kurang baik, menjadi 0,2, masih lebih baik jika tidak menggunakan pengklasteran.

Penggunaan *percentage split* pada RC_SMO juga meningkatkan nilai kinerja $Kappa$, $Precision$, $Recall$, $F-Measure$, dan ROC . Sama seperti penggunaan *fold*, nilai MAE cenderung naik dibandingkan penggunaan SMO untuk data

yang diproses dengan aturan berdasar pakar. Penggunaan *percentage split* 75%, 80%, 85%, dan 90% pada metode RC_SMO mampu mencapai rata-rata nilai $Kappa = 0,95$, $Precision = 0,55$, $Recall = 0,53$, $F-Measure = 0,54$, dan $ROC = 0,97$. Namun demikian, ukuran kinerja MAE menurun karena nilai MAE naik menjadi 0,23.

TABEL VII
UKURAN KINERJA PREDIKSI MENGGUNAKAN SMO DENGAN CROSS VALIDATION

Metode SMO						
Fold	Kappa	MAE	Precision	Recall	F-Measure	ROC
10	0,34	0,14	0,22	0,40	0,24	0,68
15	0,32	0,15	0,22	0,28	0,24	0,67
20	0,33	0,14	0,22	0,28	0,24	0,68
25	0,38	0,14	0,22	0,28	0,24	0,67
30	0,29	0,14	0,22	0,28	0,24	0,66
Metode RC_SMO						
10	0,95	0,20	0,71	0,72	0,72	0,97
15	0,95	0,20	0,71	0,73	0,72	0,98
20	0,95	0,20	0,71	0,73	0,72	0,98
25	0,95	0,20	0,71	0,73	0,72	0,98
30	0,95	0,20	0,71	0,72	0,71	0,98

TABEL VIII
UKURAN KINERJA PREDIKSI MENGGUNAKAN SMO DENGAN PERCENTAGE SPLIT

Metode SMO						
PS	Kappa	MAE	Precision	Recall	F-Measure	ROC
75%	0,42	0,14	0,21	0,27	0,24	0,62
80%	0,29	0,15	0,20	0,27	0,23	0,55
85%	0,37	0,15	0,17	0,24	0,24	0,58
90%	0,28	0,17	0,12	0,20	0,15	0,65
Metode RC_SMO						
75%	0,95	0,20	0,62	0,58	0,59	0,96
80%	0,94	0,20	0,61	0,58	0,59	0,96
85%	0,92	0,20	0,48	0,45	0,46	0,95
90%	1,00	0,30	0,50	0,50	0,50	1,00

Tabel IX dan Tabel X menunjukkan kinerja prediksi menggunakan DT. Ukuran kinerja yang paling optimal adalah pada *fold* 15, 20, 25, dan 30 pada RC_DT. Pada *fold* 15, 20, 25, dan 30 ini, DT mampu mencapai nilai maksimal $Kappa = 0,96$, $MAE = 0,06$, $Precision = 0,74$, $Recall = 0,73$, $F-Measure = 0,74$, dan $ROC = 0,97$. Sementara, dengan *percentage split* 90% nilai $Kappa = 1$, $MAE = 0,06$, dan $ROC = 1$. Nilai $Precision$, $Recall$, dan $F-Measure$ optimal pada *percentage split* 80%, yaitu $Precision = 0,61$, $Recall = 0,58$, dan $F-Measure = 0,59$.

Tabel XI dan Tabel XII menunjukkan kinerja prediksi menggunakan metode JRIP. Ukuran kinerja optimal adalah pada semua *fold* dengan RC_DT dengan nilai $Kappa = 0,95$, $MAE = 0,02$, $Precision = 0,71$, $Recall = 0,73$, $F-Measure =$

0,72, dan $ROC = 0,97$. Sementara, untuk *percentage split* menggunakan RC_JRIP juga menunjukkan kinerja yang lebih baik dibandingkan JRIP. Pada *percentage split* 90% nilai $Kappa = 1$, $MAE = 0,01$, dan $ROC = 1$.

TABEL IX
UKURAN KINERJA PREDIKSI MENGGUNAKAN DECISION TABLE DENGAN CROSS VALIDATION

Metode DT						
Fold	Kappa	MAE	Precision	Recall	F-Measure	ROC
10	0,34	0,173	0,21	0,27	0,24	0,69
15	0,3	0,17	0,21	0,26	0,23	0,71
20	0,27	0,17	0,20	0,25	0,23	0,74
25	0,32	0,17	0,21	0,26	0,23	0,73
30	0,27	0,17	0,20	0,25	0,22	0,71
Metode RC_DT						
10	0,96	0,07	0,74	0,73	0,74	0,97
15	0,96	0,06	0,74	0,73	0,74	0,97
20	0,96	0,06	0,74	0,73	0,74	0,97
25	0,96	0,06	0,74	0,73	0,74	0,97
30	0,96	0,06	0,74	0,73	0,74	0,97

TABEL X
UKURAN KINERJA PREDIKSI MENGGUNAKAN DECISION TABLE DENGAN PERCENTAGE SPLIT

Metode DT						
PS	Kappa	MAE	Precision	Recall	F-Measure	ROC
75%	0,29	0,17	0,21	0,26	0,23	0,65
80%	0,21	0,18	0,19	0,24	0,21	0,64
85%	0,12	0,19	0,16	0,21	0,18	0,61
90%	0,00	0,21	0,11	0,16	0,13	0,59
Metode RC_DT						
75%	0,84	0,08	0,46	0,45	0,45	0,89
80%	0,94	0,07	0,61	0,58	0,59	0,92
85%	0,92	0,07	0,48	0,45	0,46	0,89
90%	1,00	0,06	0,50	0,50	0,50	1,00

Nilai $Precision$, $Recall$ dan $M-Measure$ paling optimal dengan *percentage split* 70%, yaitu $Precision = 0,62$, $Recall = 0,62$, dan $F-Measure = 0,59$.

Tabel XIII dan Tabel XIV menunjukkan kinerja prediksi menggunakan metode RC_J48. Penggunaan metode RC_J48 juga menunjukkan kinerja yang lebih baik dibandingkan J48. Ukuran kinerja menggunakan *fold* 10, 15, 20, 25, dan 30 menunjukkan nilai yang sama untuk $Kappa$, MAE , $Precision$, $Recall$, $F-Measure$, dan ROC , yaitu $Kappa = 0,96$, $MAE = 0,01$, $Precision = 0,74$, $Recall = 0,73$, $F-Measure = 0,74$, dan $ROC = 0,97$.

Ukuran kinerja menggunakan *percentage split* juga menunjukkan bahwa RC_J48 lebih baik dibandingkan J48. Pada *percentage split* 90%, nilai $Kappa = 1$, $MAE = 0,01$, dan $ROC = 1$. Nilai $Precision$, $Recall$, dan $F-Measure$ terbaik pada *percentage split* 75%, yaitu $Precision = 0,62$, $Recall = 0,58$, dan $F-Measure = 0,59$.

TABEL XI
UKURAN KINERJA PREDIKSI MENGGUNAKAN JRIP DENGAN CROSS VALIDATION

Metode JRIP						
Fold	Kappa	MAE	Precision	Recall	F-Measure	ROC
10	0,35	0,15	0,27	0,30	0,28	0,56
15	0,29	0,16	0,21	0,26	0,23	0,57
20	0,29	0,16	0,21	0,26	0,23	0,58
25	0,27	0,16	0,20	0,25	0,22	0,59
30	0,31	0,16	0,22	0,26	0,23	0,55
Metode RC_JRIP						
10	0,95	0,02	0,71	0,73	0,72	0,97
15	0,95	0,02	0,71	0,73	0,72	0,97
20	0,95	0,02	0,71	0,73	0,72	0,97
25	0,95	0,02	0,71	0,73	0,72	0,97
30	0,95	0,02	0,71	0,73	0,72	0,97

TABEL XII
UKURAN KINERJA PREDIKSI MENGGUNAKAN JRIP DENGAN PERCENTAGE SPLIT

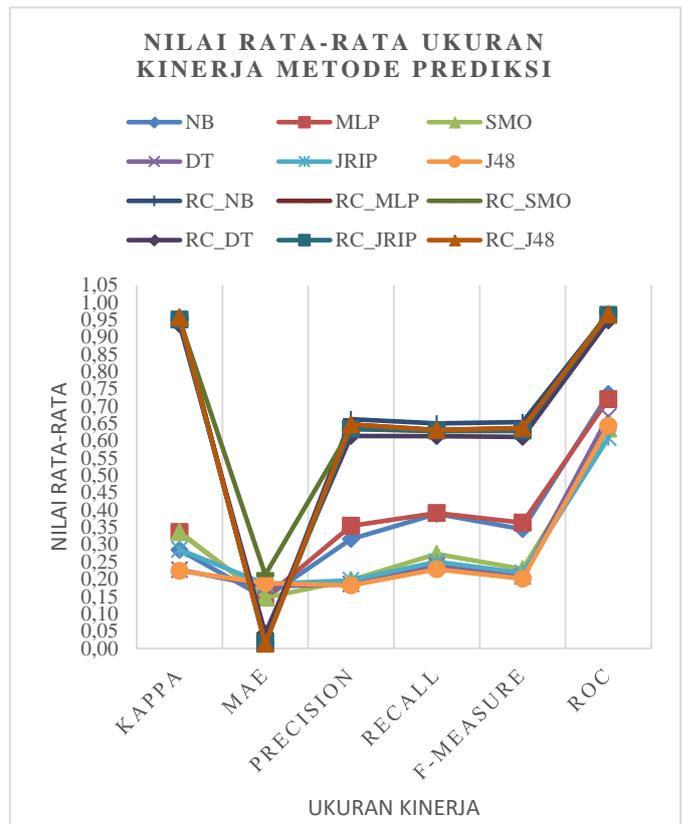
Metode JRIP						
PS	Kappa	MAE	Precision	Recall	F-Measure	ROC
75%	0,35	0,16	0,21	0,27	0,24	0,67
80%	0,30	0,16	0,20	0,27	0,23	0,65
85%	0,22	0,17	0,17	0,24	0,20	0,63
90%	0,21	0,37	0,11	0,16	0,13	0,65
Metode RC_JRIP						
75%	0,95	0,02	0,62	0,58	0,59	0,95
80%	0,94	0,02	0,61	0,58	0,59	0,95
85%	0,92	0,02	0,48	0,45	0,46	0,93
90%	1,00	0,01	0,50	0,50	0,50	1,00

TABEL XIII
UKURAN KINERJA PREDIKSI MENGGUNAKAN J48 DENGAN CROSS VALIDATION

Metode J48						
Fold	Kappa	MAE	Precision	Recall	F-Measure	ROC
10	0,29	0,15	0,21	0,26	0,23	0,61
15	0,24	0,16	0,21	0,24	0,22	0,61
20	0,24	0,17	0,20	0,24	0,22	0,62
25	0,19	0,16	0,19	0,22	0,21	0,65
30	0,22	0,16	0,20	0,23	0,21	0,63
Metode RC_J48						
10	0,96	0,01	0,74	0,73	0,74	0,97
15	0,96	0,01	0,74	0,73	0,74	0,97
20	0,96	0,01	0,74	0,73	0,74	0,97
25	0,96	0,01	0,74	0,73	0,74	0,97
30	0,96	0,01	0,74	0,73	0,74	0,97

TABEL XIV
UKURAN KINERJA PREDIKSI MENGGUNAKAN J48 DENGAN PERCENTAGE SPLIT

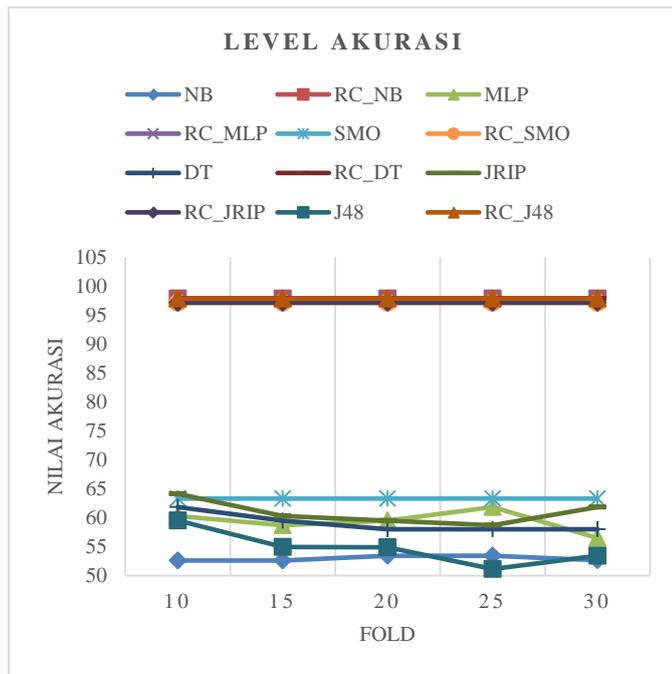
Metode J48						
PS	Kappa	MAE	Precision	Recall	F-Measure	ROC
75%	0,32	0,15	0,20	0,26	0,23	0,69
80%	0,21	0,17	0,19	0,24	0,21	0,66
85%	0,12	0,18	0,16	0,21	0,18	0,64
90%	0,20	0,37	0,11	0,16	0,13	0,65
Metode RC_J48						
75%	0,95	0,01	0,62	0,58	0,59	0,95
80%	0,94	0,02	0,61	0,58	0,59	0,95
85%	0,92	0,02	0,48	0,45	0,46	0,93
90%	1,00	0,01	0,50	0,50	0,50	1,00



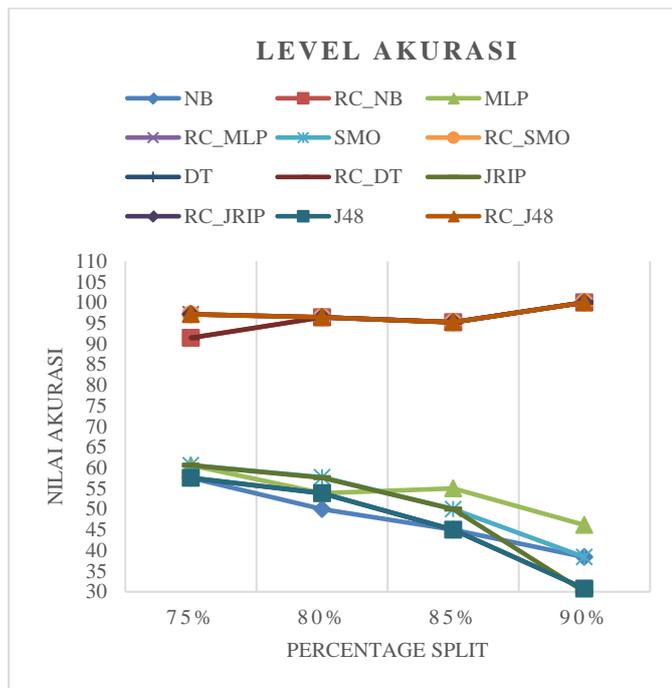
Gbr. 6 Nilai rata-rata ukuran kinerja semua metode.

Nilai rata-rata ukuran kinerja dari semua metode ditunjukkan pada Gbr. 6. Semua metode prediksi mengalami peningkatan kinerja dengan adanya pembentukan aturan berdasarkan pengklasteran.

Metode RC_NB, RC_MLP, RC_SMO, RC_DT, RC_JRIP, dan RC_J48 memiliki kinerja lebih baik dibandingkan metode NB, MLP, SMO, DT, JRIP, dan J48 secara berurutan. Hal ini ditunjukkan dengan peningkatan nilai Kappa, Precision, Recall, F-Measure, dan ROC, serta penurunan nilai MAE, kecuali pada RC_SMO, nilai MAE naik dibandingkan pada SMO.



Gbr. 7 Level akurasi dari semua metode klasifikasi dengan cross validation



Gbr. 8 Level akurasi dari semua metode klasifikasi dengan percentage split.

Hasil kinerja klasifikasi berupa level akurasi disajikan pada Gbr. 7 dan Gbr. 8.

Kedua gambar menunjukkan bahwa hasil akurasi pada metode NB, MLP, SMO, DT, JRIP, dan J48 pada data yang diolah menggunakan metode berbasis aturan yang dibentuk dari penggalian pola kemampuan siswa dengan pengklasteran (RC_NB, RC_MLP, RC_SMO, RC_DT, RC_JRIP, dan RC_J48) lebih baik dibandingkan akurasi pada metode NB, MLP, SMO, DT, JRIP, dan J48 terhadap data yang diolah dengan menggunakan metode berbasis aturan dari pakar.

IV. KESIMPULAN

Berdasarkan hasil yang diperoleh, dapat diambil kesimpulan bahwa penggalian pola berbasis kluster menggunakan *K-means* dapat digunakan untuk menentukan tingkat kemampuan peserta ujian. Penerapan metode ini pada data ujian siswa terbukti dapat meningkatkan kinerja semua metode prediksi yang digunakan dalam makalah ini, yakni metode *Naive Bayes*, MLP, SMO, *Decision Table*, JRIP, dan J48. Peningkatan kinerja ditunjukkan dengan adanya peningkatan nilai *Kappa*, *Precision*, *Recall*, *F-Measure*, dan ROC, serta penurunan nilai MAE, kecuali pada metode SMO, penerapan metode ini cenderung menyebabkan kenaikan nilai MAE. Metode ini cocok untuk sistem penilaian adaptif dengan aturan dapat disesuaikan seiring penambahan jumlah data ujian maupun penambahan jumlah variasi pola kemampuan peserta ujian.

Selain dapat digunakan untuk memprediksi tingkat kemampuan peserta ujian, hasil dari penelitian ini juga dapat digunakan untuk mendeteksi kompetensi yang telah dan belum dikuasai oleh peserta ujian. Dari Gbr. 4 terlihat ada pola bahwa sebagian besar data peserta ujian yang tuntas pada *Cp2* juga tuntas pada *Cp1*. Begitu juga peserta yang tuntas pada *Cp6* rata-rata juga tuntas pada *Cp2*. Ada kemungkinan terdapat hubungan asosiasi antara sebuah kompetensi dengan kompetensi lain. Hal ini dapat diamati dan ditemukan dengan jumlah data yang besar dan mewakili seluruh kelas. Hubungan ini menarik untuk diteliti pada penelitian berikutnya, dengan terlebih dahulu menambah *data set* yang digunakan untuk pelatihan dan pengujian.

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada LPDP atas dukungan dana yang telah diberikan dalam penelitian ini.

REFERENSI

- [1] R. Yunis and K. Telaumbanua, "Pengembangan E-Learning Berbasis LMS untuk Sekolah, Studi Kasus SMA/SMK di Sumatera Utara," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, vol. 6, no. 1, pp. 32–36, 2017.
- [2] S. Schiaffino, P. Garcia, and A. Amandi, "eTeacher: Providing personalized assistance to e-learning students," *Computers and Education*, vol. 51, no. 4, pp. 1744–1754, 2008.
- [3] D. Syarif and S. Sahid, "Modeling the Flow Experience for Personalized Context Aware E-learning," in *Proceeding of 8th International Conference on Information Technology and Electrical Engineering (ICITEE)*, 2016.
- [4] S. Q. Authority, *Guide to Assessment*. Scottish Qualifications Authority, 2014.
- [5] K. J. Kennedy and J. C. K. Lee, *The changing Roles of Schools in Asian Societies: Schools for Knowledge Society*. Routledge London and New York, 2008.
- [6] Kementerian Pendidikan dan Kebudayaan RI, "Ujian Nasional Berbasis Komputer." [Online]. Available: <https://ubk.kemdikbud.go.id/>. [Accessed: 31-Jul-2017].
- [7] K. Rajamani and V. Kathiravan, "An adaptive assessment system to compose serial test sheets using item response theory," *Proceeding of the 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, PRIME 2013*, pp. 120–124, 2013.
- [8] J. Linacre, S. Chae, U. Kang and E. Jeon, "Computer-Adaptive Testing : A Methodology Whose Time Has Come". MESA Memorandum, no. 69, 2000.

- [9] N. a. Thompson and D. J. Weiss, "A framework for the development of computerized adaptive tests," *Practical Assessment, Research and Evaluation*, vol. 16, no. 1, pp. 1–9, 2011.
- [10] S. Agarwal, N. Jain, and S. Dholay, "Adaptive Testing and Performance Analysis using Naive Bayes Classifier," *Procedia - Procedia Computer Science*, vol. 45, pp. 70–75, 2015.
- [11] I. N. Sukajaya, "Klasifikasi Domain Kognitif Pembelajar Matematika Menggunakan Serious Game Berbasis Taksonomi Bloom", *Buku Sidang Terbuka Promosi Doktor Teknik Elektro ITS*, 2016.
- [12] C. K. Hu lin, C. L., Drasgow, F., & Parsons, "Item response theory application to psychological measurement," *Homewood, Dow Jones-Irwin*, 1983.
- [13] A. Ismaya, "Algoritma Ekstraksi Informasi Berbasis Aturan," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, vol. 3, no. 4, pp. 242–247, 2014.
- [14] U. L. Yuhana, R. G. Mangowal, S. Rochimah, E. M. Yuniarno, and M. H. Purnomo, "Predicting Math Performance of Children with Special Needs Based on Serious Game," in *The 5th IEEE Conference on Serious Game and Application for Health*, 2017.
- [15] R. G. Mangowal, U. L. Yuhana, E. M. Yuniarno, and M. H. Purnomo, "MathBharata : A Serious Game for Motivating Disabled Students to Study Mathematics," in *The 5th IEEE Conference on Serious Game and Application for Health*, 2017.
- [16] Menteri Pendidikan dan Kebudayaan RI, "Permendikbud No 24 Tahun 2016," 2016.
- [17] Menteri Pendidikan dan Kebudayaan RI, "Lampiran Permendikbud No 24 Tahun 2016," 2016.
- [18] Badan Standar Nasional Pendidikan, *Standar Kompetensi dan Kompetensi Dasar SD/MI*. 2006.